Mathematics 502 probability and statistics
California State University, Fullerton
Fall 2007

Nasser M. Abbasi

spring 2007 Compiled on October 14, 2025 at 5:19pm [public]

Contents

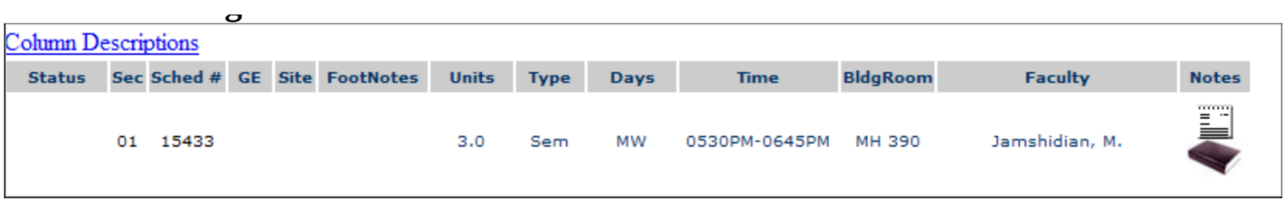
Chapter 1

Introduction

Local contents

This course part of my Masters degree in Applied Mathematics at California State University, Fullerton.

1.1 Class meet time



Column Descriptions

Status	Sec	Sched #	GE	Site	FootNotes	Units	Type	Days	Time	BldgRoom	Faculty	Notes
	01	15433				3.0	Sem	MW	0530PM-0645PM	MH 390	Jamshidian, M.	

Figure 1.1: class schedule time

1.2 Syllabus

Math 502AB – Probability and Statistics I & II

Fall 2007 - Section 1, MH 390, MW 5:30-6:45 and 7:00-8:15

Instructor: Mortaza (Mori) Jamshidian, Ph.D. **Office:** MH 180 **Phone:** 714-278-2398

Office Hours: Mon. 2:20 – 3:30, Wed. 4:30-5:25 p.m., or *by appointment*

Homepage: http://math.fullerton.edu/mori **E-mail:** *mori@fullerton.edu*

Text: *Mathematical Statistics and Data Analysis* by John Rice, Third Edition, Thompson/Brooks/Cole, 2007.

Software: We will use R, Matlab, and SAS for the projects and homework assignments. Instructions for use of these packages will be given, as necessary. R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To download R, please choose your preferred CRAN mirror. SAS and Matlab are available to students in the Mathematics Department Computing Laboratory in MH 452. Please download and install R on your computers as soon as possible.

Your e-mail address wanted: You are ***required*** to fill out the "Student Information Form" (click here) and submit it to mori@fullerton.edu no later than **Saturday, August 25**. **Do not** save the pdf file and attach to an e-mail. The file that I need is an XML file. You need to **use the submit button** on the form and follow the instructions. I will send various communiqué, **take home quizzes**, and last minute announcements about our class through e-mail. Please provide an e-mail address that you check frequently. I will send a "test e-mail" on Sunday August 26 to everyone. If you do not receive this test e-mail, please see me on Monday to resolve any problems there may be. ***Note:*** Any credits that you lose due to not establishing your e-mail connection with me on time will be your responsibility.

Course Description: This course has two parts. In the first part we learn fundamentals of probability theory, including random variables, joint and conditional distributions, expected values, major probability limit theorems, and some well-known distributions. The objective in the second part is to utilize the probability theory learnt in the first part mainly for statistical inference. We will learn topics including survey sampling methods, parameter estimation specially maximum likelihood and method of moments, Bayesian estimation, properties of estimators, test of hypothesis and goodness of fit, exploratory data analysis, analysis of variance, regression analysis, and analysis of categorical data.

Course requirements and Grading Policy: Homework/projects (**30%**) will be assigned and graded. I often give a quiz related to the homework problems, and use the quiz grade instead of the homework grade. Two midterm exams (**40%**) and a final exam (**30%**) will be given. Portions of the exams may be take-home. For in-class exams you will be allowed to bring in one page of information during each midterm exam and two pages of information during the final exam. Letter grades will be assigned according to the distribution of the overall grades. Plus-minus grading will be used.

The exam dates are as follows:

Figure 1.2: page 1

Exam I	*Exam II*	*Final Exam*
Wednesday, Oct. 3	Wednesday, Nov. 7	Monday, Dec. 10, 5:00-6:50 p.m.
		Wednesday, Dec. 12, 7:30-9:20 p.m.

Late homework/projects will not be accepted. Make-up exams will be given only in extreme instances and only with advance permission of the instructor. Any student who does not take an exam at the scheduled time without prior consent of the instructor will receive a grade of zero on that exam. If any student feels that a sudden illness is sufficiently extreme to warrant a make-up exam, the instructor must be provided with documentation prepared by an appropriate authority.

Academic Integrity: Students who violate university standards of academic integrity are subject to disciplinary sanctions, including failure in the course and suspension from the university. Since dishonesty in any form harms the individual, other students and the university, policies on academic integrity are strictly enforced. I expect that you will familiarize yourself with the academic integrity guidelines found in the current student handbook (see http://www.fullerton.edu/deanofstudents/judicial/policies.htm).

Examples of actions that constitute academic dishonesty include, but are not limited to:

1. Unacceptable examination behavior – communicating with fellow students, copying material from another student's exam or allowing another student to copy from an exam, possessing or using unauthorized materials, or any behavior that defeats the intent of an exam.
2. Plagiarism – taking the work of another and offering it as one's own without giving credit to that source, whether that material is paraphrased or copied in verbatim or near-verbatim form.
3. Unauthorized collaboration on a project, homework or other assignment.
4. Documentary falsification including forgery, altering of campus documents or records, tampering with grading procedures, fabricating lab assignments, or altering medical excuses.

Emergency Evacuation: In the event of an emergency such as earthquake or fire:

- Take all your personal belongings and leave the classroom. Use the stairways located at the east, west, or center of the building.
- Do not use the elevator. They may not be working once the alarm sounds.
- Go to the lawn area towards Nutwood Avenue. Stay with class members for further instruction.
- For additional information on exits, fire alarms and telephones, **Building Evacuation Maps** are located near each elevator.
- Anyone who may have difficulty evacuating the building, please see the instructor.

Some Important dates:

September 4 (Tuesday): Last day for students to drop without a grade of "W". Students drop using Titan.
September 28 (Friday): Last day the Math Department will be flexible on the approval of late withdrawal requests. Beginning Monday, October 1, students must have a serious and compelling reason for withdrawing (e.g. medical reason) and must provide supporting documentation for their reason. Please encourage students who are considering withdrawing to do so BY September 28.

November 9 (Friday): Last day to withdraw with a truly serious and compelling reason that is beyond the student's control. Students must document their reason.

Figure 1.3: page 2

1.3 Instructor

Instructor and course official web site is `http://math.fullerton.edu/mori/`

Professor Mortaza Jamshidian
Department of Mathematics
800 N. State College Blvd.
Fullerton, CA 92834
E-mail: mori@fullerton.edu

Phone: 714-278-2398 (office)
714-278-3631 (Dept.)
Fax:714-278-1431

1.4 Textbook

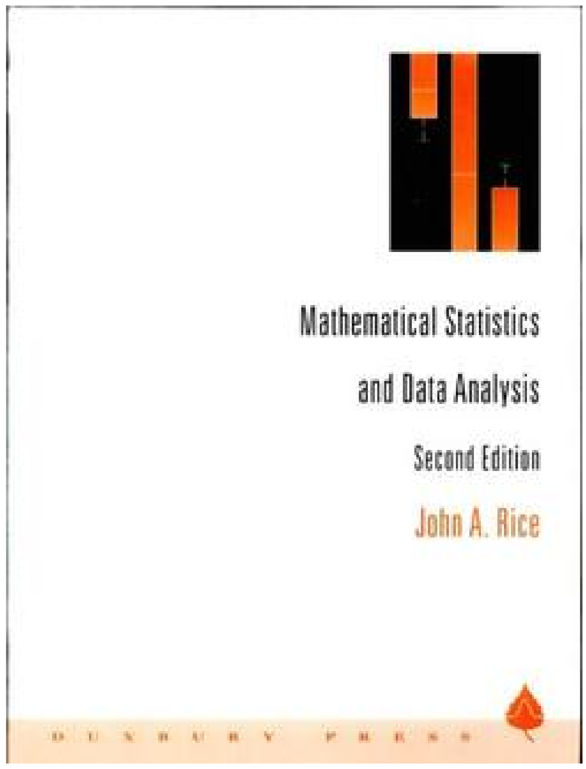


Figure 1.4: Book

Amazon web page for the textbook is here

1.4.1 Table of content of textbook

vii

Figure 1.5: page 1

Figure 1.6: page 2

Figure 1.7: page 3

Figure 1.8: page 4

Figure 1.9: page 5

Figure 1.10: page 6

Figure 1.11: page 7

Chapter 2

Projects

Local contents

2.1 List of Projects

1. Generating pseudo random numbers from a given distribution.
2. Accept/Reject Algorithm.
3. The Bivariate Normal Distribution, Due Wed. October 25.
4. Maximum likelihood and Bootstrap.

2.2 Figures

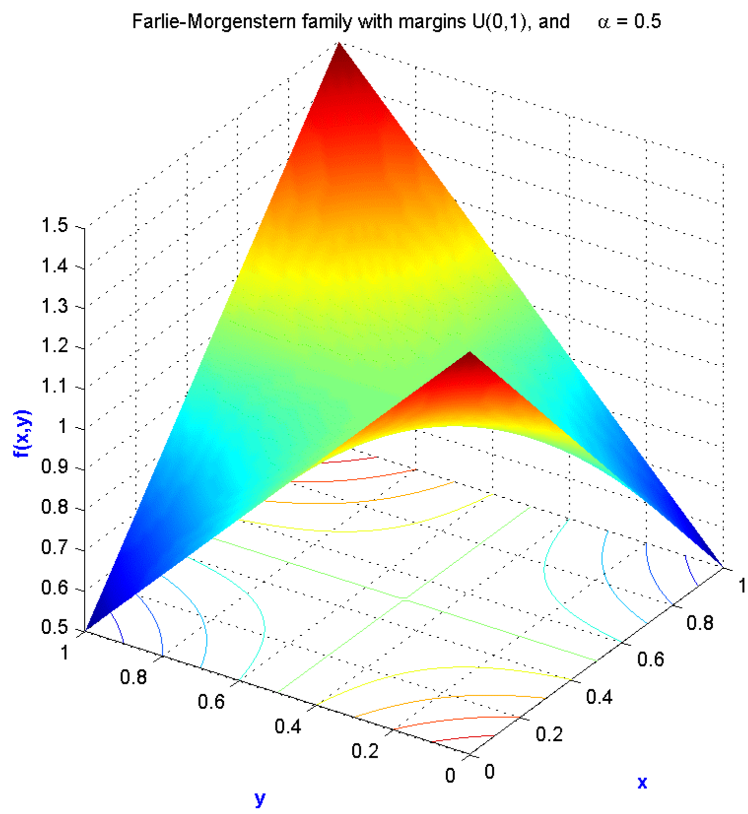


Figure 2.1: A bivariate distribution with uniform marginals

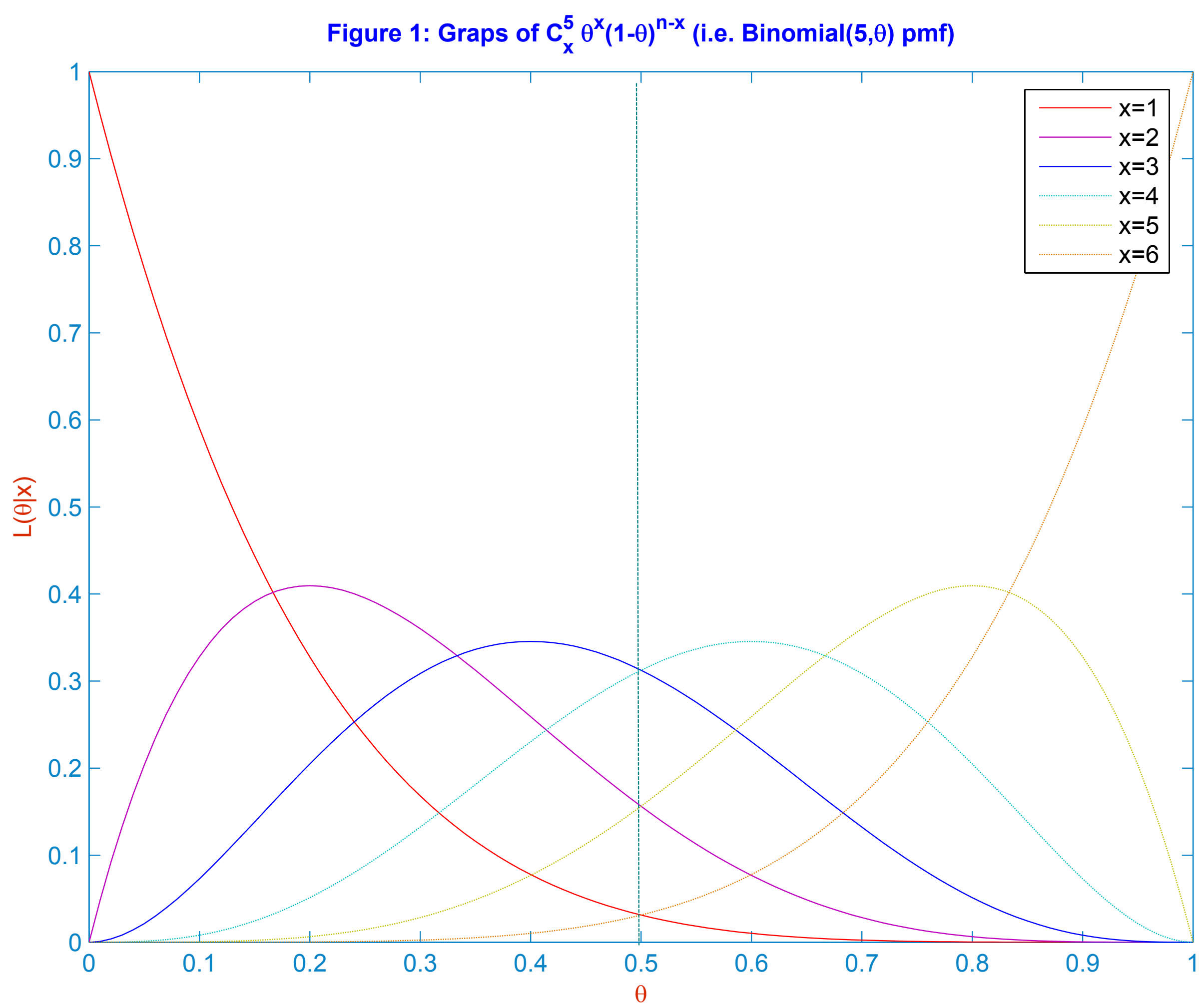


Figure 2.2: $Binomial(5, \theta)$ pmf

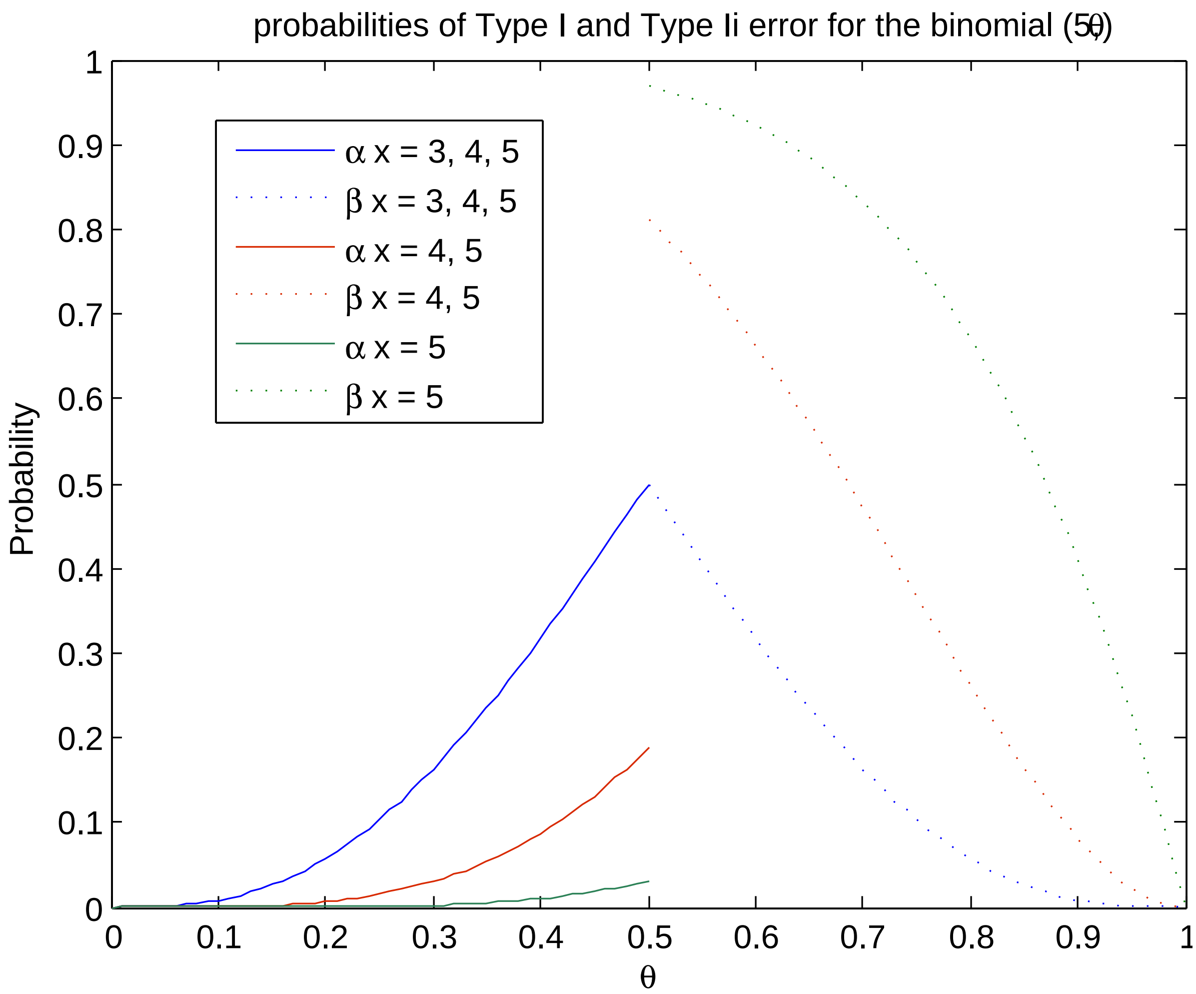


Figure 2.3: type I and type II errors for the $binomial(5, \theta)$ example

2.3 Project 1

2.3.1 Introduction

Project handout given to us by Professor Jamshidian. The following describes the project.

Project 1: Generating pseudo random numbers from various distributions

Instructions: Submit a hard copy of your project, including the code. Send a softcopy of your code by e-mail with subject line "code for Project 1 – Your name". Please choose filenames that clearly indicate which program belongs to which (part of a) problem [e.g. "prob3bc.R" for problem 3 section c]. I will accept code in Matlab, R, or Mathematica. However, I will provide help only to those that use R, just to encourage you to use R. In fact the project is written with references to the R language, but you can use equivalent commends in Matlab or Mathematica.

1. `[10 points]` The distribution function for the exponential distribution with parameter λ is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{Otherwise.} \end{cases}$$

a. Write an `R` program that uses F^{-1} and the uniform random number generator in `R` (`runif`) to generate n numbers from the exponential distribution with λ. Use `set.seed(`*your birthday MMDDYY*`)`.

b. Generate $n = 10000$ values from the exponential random variable with parameter $\lambda = 2$, using your program (don't give me the numbers generated!). Use the `hist` command in R to plot the *density* histogram of the *relative frequencies* for the generated data. Overlay this histogram by the graph of the density of the exponential random variable with parameter $\lambda = 2$. The commands `dexp` and `lines`, or `curves` can be used. Select an appropriate number of bins for your histogram. Briefly comment on the relationship between the histogram and the density curve.

2. `[20 points]` Let $\phi_0(x)$ and $\phi_1(x)$ denote the density functions for two normal random variables with $\sigma = 1$, and respective means 0 and 3. The density for a mixture distribution of these is defined as

$$f(x) = \gamma\phi_0(x) + (1-\gamma)\phi_1(x) \qquad -\infty < x < \infty.$$

for a given admixture parameter $0 < \gamma < 1$.

a. Write an `R` program, using the normal random generator `rnorm` and the uniform random number generator `runif`, to generate n random numbers from this mixture. The inputs to your program should be n, and γ. Generate two sets of 10,000 data points; one with $\gamma = .75$, and another with $\gamma = .25$.

b. For each of the data sets generated, graph the density histogram and superimpose it by the density $f(x)$ defined in the equation above. Make sure to choose the number of bins for the histogram appropriately. In each case explain why the shape of the density that you obtain is expected.

3. `[30 points]` Problem 42 on page 111 of your text gives the pdf for the double exponential density with parameter λ. It also suggests a method to generate random numbers from the double exponential family using two random variables W and T, described in the problem.

a. Write a program that generates random numbers from the double exponential family. The input to the program should be the parameter λ, and the number of random numbers to be generated, n. The output should be n pseudo random numbers from the double exponential with parameter λ. You are only allowed to use `runif` in your program for random number generation.

b. Write a program that uses the Accept/Reject algorithm efficiently to generate n observations from the standard normal density $N(0,1)$, using random numbers that are generated from the uniform(0,1) and the double exponential random variate with parameter $\lambda = 1$ [your program in part (a)]. Your program should also count and report the proportion of values that are rejected. Give the density histogram of $n = 10,000$ numbers generated from your program and superimpose it by the standard normal pdf.

The project contains 3 problems. To make it easier to run each problem (due to the use of Dynamic UI in the project), I implemented each one in a separate Mathematica notebook.

Below are the links to each problem. There are 3 links to each problem, one for the PDF file report, second for the HTML version of the report, and one for the Mathematica notebook itself to run the code.

In all 3 cases, to run the Mathematica code, please do the following: Download the Mathematica notebook. Open it using Mathematica. Click on the Evaluation menu option at the top, and then select Evaluate Notebook. This will run all the code. Scroll to the bottom of the notebook and the GUI should be UP and ready to use.

2.3.2 Report and code links

2.3.2.1 problem 1

Project one. Problem one. Mathematics 502 Probability and Statistics

Nasser Abbasi, September 26,2007. California State University, Fullerton

1. `[10 points]` The distribution function for the exponential distribution with parameter λ is given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{Otherwise.} \end{cases}$$

a. Write an `R` program that uses F^{-1} and the uniform random number generator in `R` (`runif`) to generate n numbers from the exponential distribution with λ. Use `set.seed(`*your birthday MMDDYY*`)`.

b. Generate $n = 10000$ values from the exponential random variable with parameter $\lambda = 2$, using your program (don't give me the numbers generated!). Use the `hist` command in R to plot the *density* histogram of the *relative frequencies* for the generated data. Overlay this histogram by the graph of the density of the exponential random variable with parameter $\lambda = 2$. The commands `dexp` and `lines`, or `curves` can be used. Select an appropriate number of bins for your histogram. Briefly comment on the relationship between the histogram and the density curve.

Problem 1 part (a)

The CDF given is defined as $F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & o.w. \end{cases}$ To find F^{-1} we need to solve for x in the equation $1 - e^{-\lambda x}$ for $x \geq 0$ Hence we write

$$\begin{aligned} y &= 1 - e^{-\lambda x} \\ e^{-\lambda x} &= 1 - y \\ -\lambda x &= \ln(1 - y) \\ x &= \frac{-1}{\lambda} \ln(1 - y) \end{aligned}$$

Therefore

$$F^{-1}(y) = \frac{-1}{\lambda} \ln(1 - y)$$

Now to generate random numbers which belongs to an exponential distribution, we will now generate random numbers from $U(0, 1)$ and for each such number generated, we will apply the above function F^{-1} on it, and the result will be a random number which belongs to the exponential distribution. For example, if $\lambda = 2$ and a uniform random number is say 0.4, then we evaluate $F^{-1}(0.4) = \frac{-1}{2} \ln(1 - 0.4) = 0.25541$

And so this is the idea to implement. We need to first seed the uniform random number generator before we start.

Algorithm

Input: λ: parameter, n: number of random numbers to generate

output: a list of n random numbers from the probability density function ~ $F(x)$ given above.

1. Seed the uniform random number generator with (010101).
2. initialize the array d of size n which will contain the list of random numbers generated below.

This loop below is just an algorithmic view. In actual code, a 'vector' operation Table[] in used for speed.

3. For i in 1..n LOOP

Generate y_k which is a random generated from uniform distribution using the build in function RandomReal[0,1]

$d[i]= F^{-1}(y_k)$ using input λ.

END LOOP

4. Find histogram of d. Select an appropriate number of bins. Let f_a be the histogram found.
5. Now find the relative frequency f_r by dividing set f_a by the number of observations n. Hence histogram now is $f_r = \frac{f_a}{n}$
6. Now scale the histogram such that it is density. Total area is 1. Do this by finding total area under histogram, and divide each bin count by this area.
7. Plot the histogram and the exponential distribution $\lambda e^{-\lambda x}$ on the same plot.

Code Implementation

Define the function F^{-1}which was derived earlier. This is the inverse of the CDF of the exponential density function $\lambda e^{-\lambda x}$

In[156]:=
```
Remove["Global`*"];
gDebug = False;
```

In[158]:=
```
inverseCDFofExponentialDistribution[λ_, n_] := Module[{}, -1/λ Log[1 - n]]
```

Function below is called to generate N random numbers using the above F^{-1} function (User needs to seed before calling

In[159]:=
```
getRandomNumbersFromExponential[λ_, nRandomVariables_] := Module[{i},
  Table[ inverseCDFofExponentialDistribution[λ, RandomReal[]], {i, nRandomVariables}]]
```

Problem 1 part(b)

Generate n = 10000 for $\lambda = 2$ and overlay with relative frequency, use appropriate number of bins. See appendix for the function postProcessForPartOne[] which generate the plots. Removed below to reduce code clutter in the main report.

This function makes a histogram which is scaled to be used to overlay density plots, or other functions.
Input: originalData: this is an array of numbers which represents the data to bin
nBins: number of bins
output: the histogram itself but scaled such that area is ONE

In[160]:=

```
Needs["BarCharts`"]
nmaMakeDensityHistogram[originalData_, nBins_] :=
 Module[{freq, binSize, from, to, scaleFactor, j, a, currentArea},
  to = Max[originalData];
  from = Min[originalData];
  binSize = (to - from) / nBins;
  freq = BinCounts[originalData, binSize];
  currentArea = Sum[binSize * freq[[i]], {i, nBins}];
  freq = freq / currentArea;
  a = from;
  Table[{a + (j - 1) * binSize, freq[[j]], binSize}, {j, 1, nBins}]
 ]
```

This function to overlay the histogram and the PDF. It is used by the simulation program as well (that is why it is a little larger than needed)

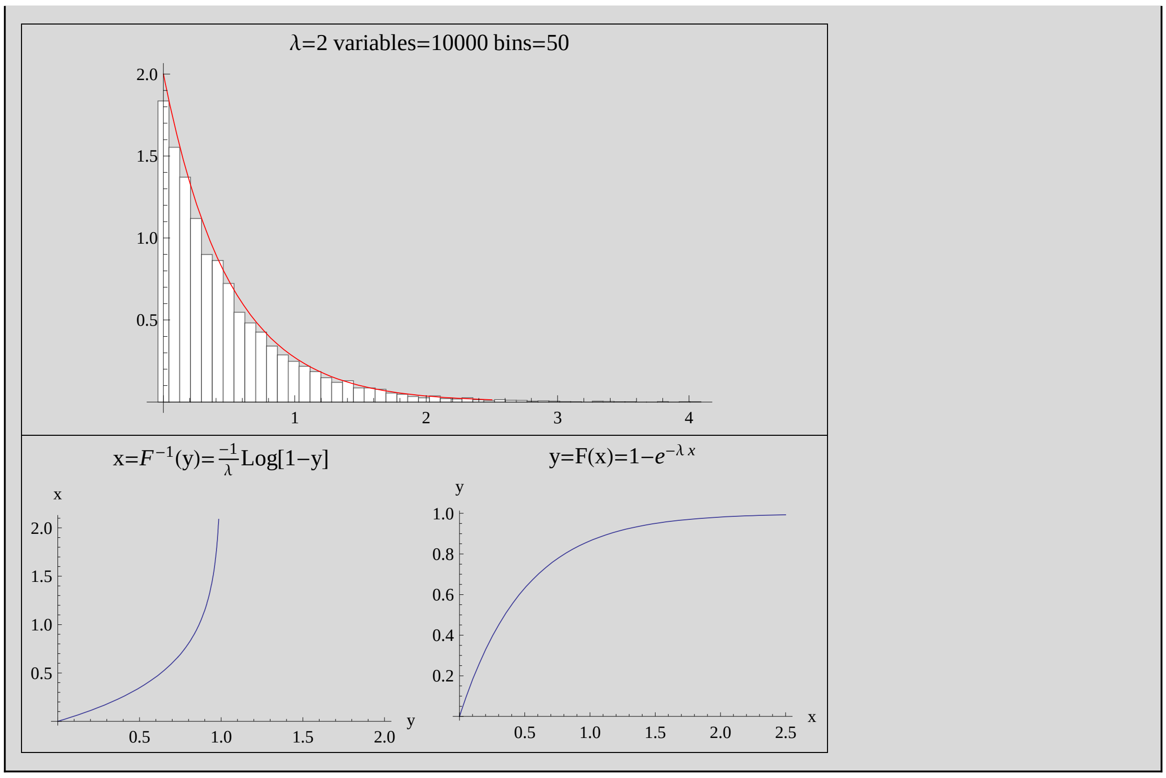
In[162]:=
```
postProcessForPartOne[randomNumbers_, nBins_, λ_, nRandomVariables_, fromX_, toX_] :=
 Module[{frequency, relativeFrequency, p1, p2,
   x, factor, max, imSize = 300, dx, p, pCDF, pinvCDF, gz, g},

  (*find dx which is bin size, needed by Mathematica BinCount function*)
  dx = (toX - fromX) / nBins;

  (*generate frequency count using the above bin size*)
  frequency = BinCounts[randomNumbers, {fromX, toX, dx}];

  (*now normalize by total number of observation to obtain the relative frequencies*)
  relativeFrequency = N[frequency / nRandomVariables];

  (*Now divide by scale factorλ, to scale it *)
  max = Max[relativeFrequency];
  factor = λ / max;
  relativeFrequency = relativeFrequency * factor;

  gz = nmaMakeDensityHistogram[randomNumbers, nBins];
  p1 = GeneralizedBarChart[gz, BarStyle → White,
    ImageSize → imSize, PlotLabel → "λ=" <> ToString[λ] <> " variables=" <>
      ToString[nRandomVariables] <> " bins=" <> ToString[nBins]];

  p2 = Plot[PDF[ExponentialDistribution[λ], x], {x, fromX, toX}
    , PlotRange → All, Frame → True, PlotStyle → {Red(*,Thick*)}, ImageSize → imSize];

  p = Show[{p1, p2}];

  pinvCDF = Plot[inverseCDFofExponentialDistribution[λ, y], {y, 0, λ}
    , PlotLabel → "x=F^-1(y)=-1/λ Log[1-y]", ImageSize → 200, AxesLabel → {"y", "x"}];

  pCDF = Plot[1 - Exp[-λ x], {x, fromX, toX}
    , PlotLabel → "y=F(x)=1-e^(-λ x)", ImageSize → 200, AxesLabel → {"x", "y"}];

  g = Grid[{{pinvCDF, pCDF}}];
  Grid[ { {p}, {g}}, Alignment → {{Center}, {Center}}, Frame → All]
 ]
```

now generate the needed outout for N = 10000

In[163]:=

```
nBins = 50; λ = 2; numberOfVariables = 10 000; fromX = 0; toX = 2.5;
postProcessForPartOne[getRandomNumbersFromExponential[λ, numberOfVariables],
 nBins, λ, numberOfVariables, fromX, toX]
```

Out[164]=



Comment and analysis

Below I show snap shots of few plots of the density overlaid with the histogram for different values of n which is the number of random variables.

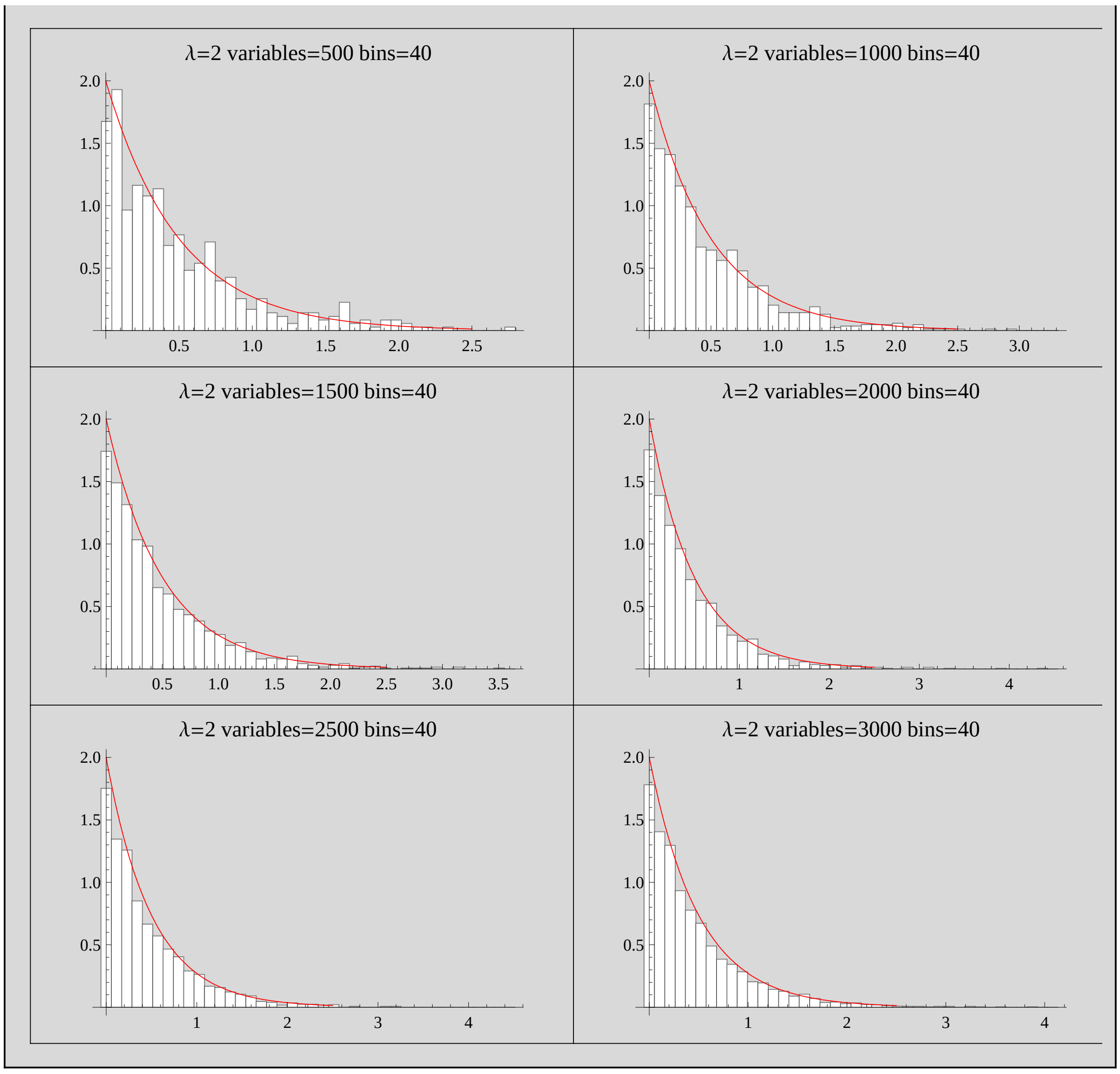
We see from the plots below, that for a fixed number of bins, fixed λ, that as more random variables are generated, the histogram overlaid on top of the actual PDF becomes closer and closer to the PDF curve. The error between the histogram and the PDF curve becomes smaller the larger the number of random variables used. This indicates that this method of finding random numbers for density function will converge to the density function. We need to select an appropriate bin size to see this more clearly. The smaller the bin size the more clear this will become (but too small a bin size will make the histogram itself not too clear).

Please see appendix for additional GUI based simulation for this part of the project.

In[165]:=

```
nBins = 40; λ = 2; fromX = 0; toX = 2.5;
SeedRandom[010 101];
p = Table[postProcessForPartOne[getRandomNumbersFromExponential[λ, nRandomVariables],
    nBins, λ, nRandomVariables, fromX, toX], {nRandomVariables, 500, 6 * 500, 500}];
GraphicsGrid[{ {First[p[[1, 1, 1]]], First[p[[2, 1, 1]]]}, {First[p[[3, 1, 1]]],
        First[p[[4, 1, 1]]]}, {First[p[[5, 1, 1]]], First[p[[6, 1, 1]]]}},
 Frame → All, ImageSize → 600]
```

Out[168]=

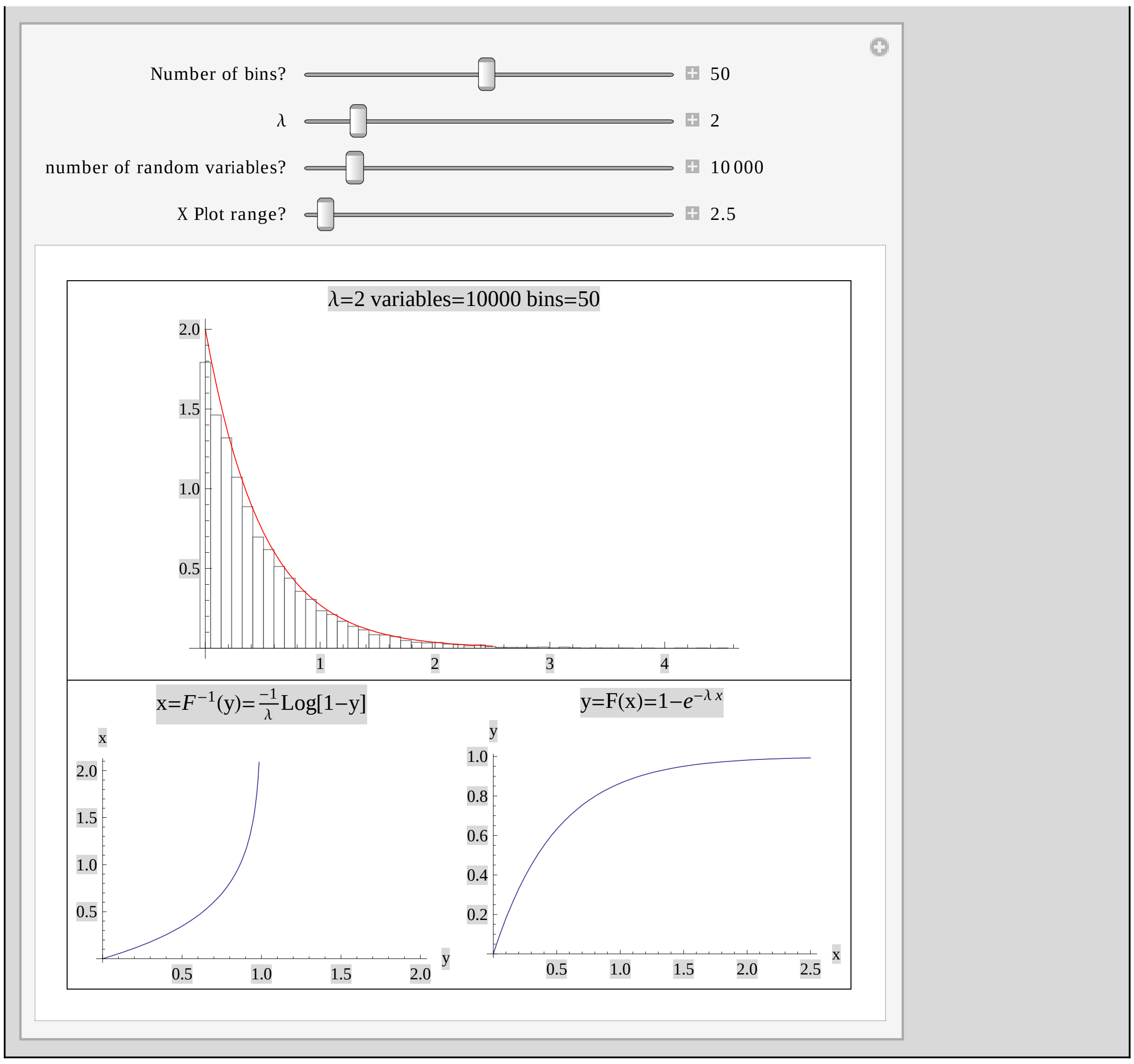


Problem 1 simulation

Define function which accepts a list of random variables from exponential distribution, and λ and generates a plot of the histogram overlaid by the exponential density plot.

In[169]:=
```
m = Manipulate[(SeedRandom[010101];
   postProcessForPartOne[getRandomNumbersFromExponential[λ, n], nBins, λ, n, 0, maxX]),
  {{nBins, 50, "Number of bins?"}, 1, 100, 1, ContinuousAction → True,
   Appearance → "Labeled"},
  {{λ, 2, "λ"}, 1, 10, .01, ContinuousAction → True, Appearance → "Labeled"},
  {{n, 10000, "number of random variables?"},
   10, 100000, ContinuousAction → True, Appearance → "Labeled"},
  {{maxX, 2.5, "X Plot range?"}, 1, 100, 1, ContinuousAction → True, Appearance → "Labeled"},
  AutorunSequencing → {{1, 15}, {2, 20}, {3, 15}}
 ]
```

Out[169]=



2.3.2.2 problem 2

Project one. Problem 2. Mathematics 502 Probability and Statistics

Nasser Abbasi, September 26,2007. California State University, Fullerton

2. [20 points] Let $\phi_0(x)$ and $\phi_1(x)$ denote the density functions for two normal random variables with $\sigma = 1$, and respective means 0 and 3. The density for a mixture distribution of these is defined as

$$f(x) = \gamma\phi_0(x) + (1-\gamma)\phi_1(x) \quad -\infty < x < \infty.$$

for a given admixture parameter $0 < \gamma < 1$.

a. Write an R program, using the normal random generator `rnorm` and the uniform random number generator `runif`, to generate n random numbers from this mixture. The inputs to your program should be n, and γ. Generate two sets of 10,000 data points; one with $\gamma = .75$, and another with $\gamma = .25$.

b. For each of the data sets generated, graph the density histogram and superimpose it by the density $f(x)$ defined in the equation above. Make sure to choose the number of bins for the histogram appropriately. In each case explain why the shape of the density that you obtain is expected.

Part 2 (a)

The mixed distribution is

$$f(x) = \gamma\phi_0(x) + (1-\gamma)\,\phi_1(x) \quad -\infty \le x \le \infty$$

Where ϕ_0 is the density function for normal distribution with σ=1,μ=0, and ϕ_1 is the density function for normal distribution with σ=1,μ=3

We need to generate random numbers from the above density function.

The following is the idea of how to solve this problem. Let us consider the case for γ=75% .Generate a r.v. from a uniform distribution, which will be between [0, 1]. Let this number be called ζ. If $\zeta < .76$ then now we will generate a random number from the above $\phi_0(x)$ normal distribution otherwise we will generate a random number from $\phi_1(x)$ distribution. Hence this is the algorithm

Algorithm

Input: γ,σ_0, μ_0, σ_1, μ_1, n (where n is number of random number to generate for mixture $f(x)$)
output: n random numbers that belong to mixture $f(x)$

seed the random number generator (010101)
Initialize array d of the size of the random numbers generated

For i in 1..n LOOP
  ζ = generate a r.v. from $U[0, 1]$
  IF $\zeta < \gamma$ THEN
    d[i]= generate a random number from $\phi_0 \sim N(\sigma_0, \mu_0)$
  ELSE
    d[i]= generate a random number from $\phi_1 \sim N(\sigma_1, \mu_1)$
  ENDIF
END LOOP

output d which will now contain n random variables drawn from the above probability density function.

Function to implement the mixture random variable algorithm

In[190]:=
```
Remove["Global`*"];
gDebug = False;
```

In[192]:=
```
(*this function below imlements the above algorithm*)
processPartTwo[γ_, μ0_, σ0_, μ1_, σ1_, n_] := Module[{d = {}, k, ξ},
  ξ = Table[RandomReal[], {k, 1, n}];
  d = Table[ If[ξ[[k]] < γ, RandomReal[NormalDistribution[μ0, σ0]],
     RandomReal[NormalDistribution[μ1, σ1]]], {k, 1, n}]
 ]
```

Now generate 2 sets of numbers each 10000 long, one for γ = .75 and the second for γ = .25

In[193]:=
```
μ0 = 0; σ0 = 1; μ1 = 3; σ1 = 1; n = 10 000;
SeedRandom[010 101];
γ = .75;  setA = processPartTwo[γ, μ0, σ0, μ1, σ1, n];
γ = .25;  setB = processPartTwo[γ, μ0, σ0, μ1, σ1, n];
```

Part 2 (b)

Generate 2 plots, one for γ = .75 and one for γ=.25 for number of random variables=10000 generated in part (a) overlaid by histogram.
First define the mixture density function (the true density). Please see appendix for the code that overlays the histogram and the mixture function called postprocessPartTwo[]. Moved below to the appendix to reduce code clutter in the main report.

In[197]:=
```
mixtureDensity[x_, γ_, μ0_, σ0_, μ1_, σ1_] :=
  γ PDF[NormalDistribution[μ0, σ0], x] + (1 - γ) PDF[NormalDistribution[μ1, σ1], x];
```

Now call the above on the 2 sets of 10000 numbers generated in part (a) and display the result

This function makes a histogram which is scaled to be used to overlay density plots, or other functions.
Input: originalData: this is an array of numbers which represents the data to bin

nBins: number of bins
output: the histogram itself but scaled such that area is ONE

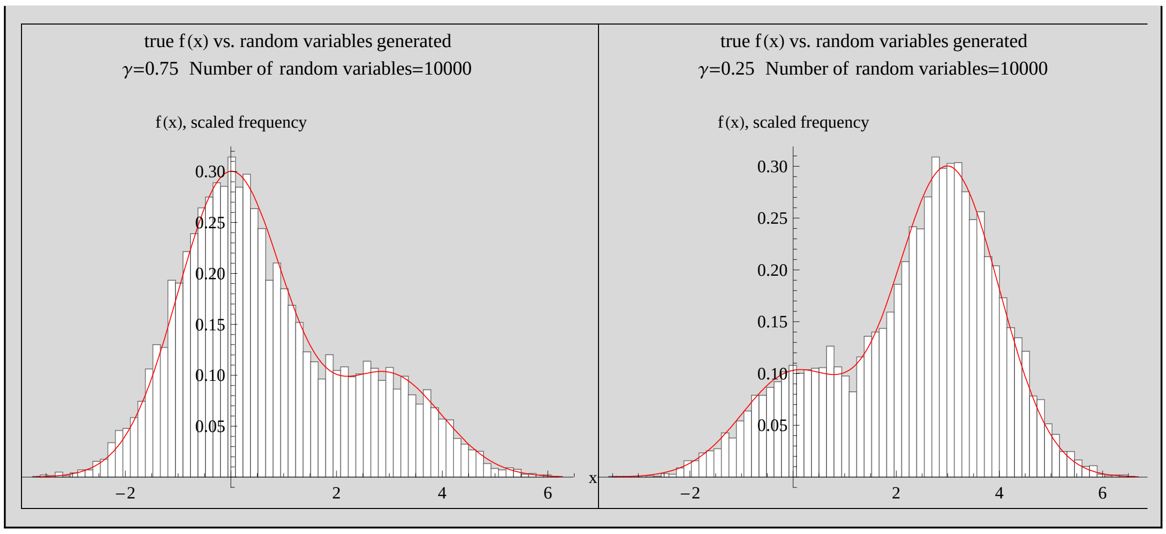
In[198]:=
```
Needs["BarCharts`"]
nmaMakeDensityHistogram[originalData_, nBins_] :=
 Module[{freq, binSize, from, to, scaleFactor, j, a, currentArea},
  to = Max[originalData];
  from = Min[originalData];
  binSize = (to - from) / nBins;
  freq = BinCounts[originalData, binSize];
  currentArea = Sum[binSize * freq[[i]], {i, nBins}];
  freq = freq / currentArea;
  a = from;
  Table[{a + (j - 1) * binSize, freq[[j]], binSize}, {j, 1, nBins}]
 ]
```

In[200]:=
```
postprocessPartTwo[d_, γ_, μ0_, σ0_, μ1_, σ1_, nRandomVariables_, imageSize_] :=
 Module[{freq, p, pList, xFrom, xTo, scaleFactor, maxSampled, sampled,
   gz = {}, fz0, fz1, x, maxBin, imSize = imageSize, nBins = 70, from},
  xFrom = Min[d];
  xTo = Max[d];
  gz = nmaMakeDensityHistogram[d, nBins];
  pList = GeneralizedBarChart[gz, BarStyle → White, ImageSize → imSize];

  p = Plot[mixtureDensity[x, γ, μ0, σ0, μ1, σ1], {x, xFrom, xTo},
    AxesOrigin → {0, 0}, PlotRange → All, PlotLabel → "Analytical plot of f(x)",
    ImageSize → imSize, (*PlotStyle→{Dashed,Red}*)PlotStyle → {Red}];

  Show[{pList, p},
   PlotLabel → Style["true f(x) vs. random variables generated\n" <> "γ=" <> ToString[γ] <>
      "  Number of random variables=" <> ToString[nRandomVariables] <> "\n", 10],
   AxesLabel → {"x", "f(x), scaled frequency"}]
 ]
```

In[201]:=

```
γ = .75;  p1 = postprocessPartTwo[setA, γ, μ0, σ0, μ1, σ1, n, 300];
γ = .25;  p2 = postprocessPartTwo[setB, γ, μ0, σ0, μ1, σ1, n, 300];
Grid[{{p1, p2}}, Frame → All, Spacings → 0, ItemSize → Full]
```

Out[203]=



Comment and analysis on result of part 2 (b) plots

In the left plot, γ=.75, hence 75% of the mixture comes from ϕ_0 which has a mean of 0, hence we would expect that at zero the bulk of the concentration, which what the plots shows to be the case (since both concentration have the same variance). Hence there should be more random numbers generated from this mixture around x=0 as well, and we see from the histogram that this is the case.
In the right plot, now γ=.25, hence 75% of the concentration will come the ϕ_1 distribution which has mean of 3. Hence again, we see that more random numbers are generated around 3 than anywhere else. These plots also show that the random numbers generated will have a probability density which will converge to the $f(x)$ given as more and more random variables are generated.

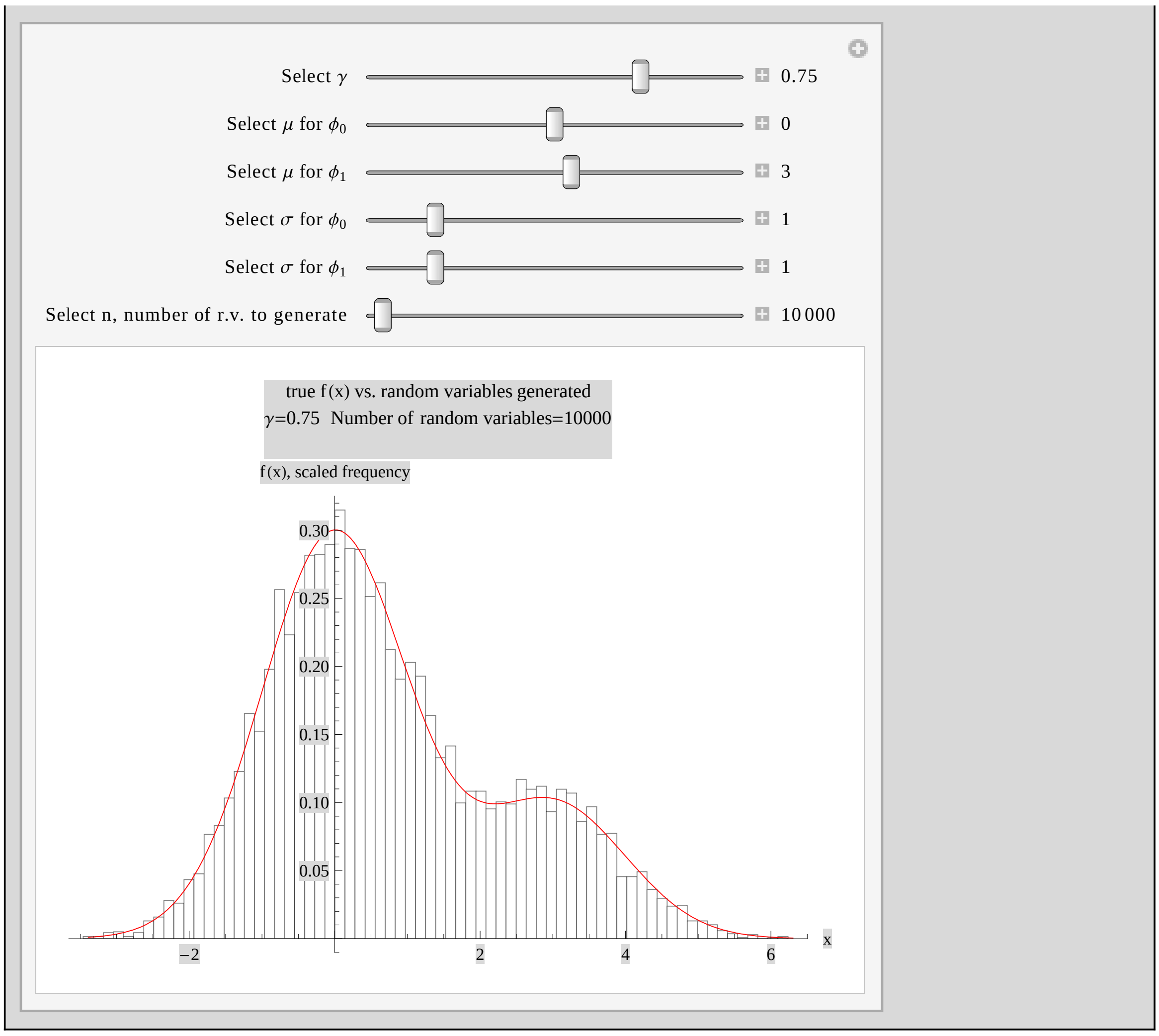
Simulation program for problem 2

GUI to simulate part 2(b) of the project

In[204]:=

```
m = Manipulate[
  postprocessPartTwo[processPartTwo[γ, μ0, σ0, μ1, σ1, n], γ, μ0, σ0, μ1, σ1, n, 400],
  {{γ, .75, "Select γ"}, 0, 1, .1, Appearance → "Labeled"}
  , {{μ0, 0, "Select μ for ϕ0"}, -30, 30, .1, Appearance → "Labeled"}
  , {{μ1, 3, "Select μ for ϕ1"}, -30, 30, .1, Appearance → "Labeled"}
  , {{σ0, 1, "Select σ for ϕ0"}, 0.1, 6, .1, Appearance → "Labeled"}
  , {{σ1, 1, "Select σ for ϕ1"}, 0.1, 6, .1, Appearance → "Labeled"}
  , {{n, 10 000, "Select n, number of r.v. to generate"}, 10 000, 100 000, 10 000,
   Appearance → "Labeled"}, AutorunSequencing → {{1, 15}, {2, 25}, {3, 25}, {4, 25}, {5, 25}}
 ]
```

Out[204]=



In[205]:=

```
(*Export["m.swf",m,"RepeatAnimation"→True,"CompressionMethod"->None]*)
```

Printed by Mathematica for Students

Mathematica notebook

2.3.2.3 problem 3

Project one. Problem Three. Mathematics 502 Probability and Statistics

Nasser Abbasi, September 26,2007. California State University, Fullerton

3. [30 points] Problem 42 on page 111 of your text gives the pdf for the double exponential density with parameter λ. It also suggests a method to generate random numbers from the double exponential family using two random variables W and T, described in the problem.

a. Write a program that generates random numbers from the double exponential family. The input to the program should be the parameter λ, and the number of random numbers to be generated, n. The output should be n pseudo random numbers from the double exponential with parameter λ. You are only allowed to use `runif` in your program for random number generation.

b. Write a program that uses the Accept/Reject algorithm efficiently to generate n observations from the standard normal density $N(0,1)$, using random numbers that are generated from the uniform(0,1) and the double exponential random variate with parameter $\lambda = 1$ [your program in part (a)]. Your program should also count and report the proportion of values that are rejected. Give the density histogram of $n = 10,000$ numbers generated from your program and superimpose it by the standard normal pdf.

Problem 3 part (a)

We are asked to generate R.V's from $f(x) = \frac{\lambda}{2} e^{-\lambda|x|}$. We note as shown in the problem itself, that R.V. X can be written as product of 2 R.V WT where $W$is ± 1 with probability $\frac{1}{2}$each. Hence to generate R.V. we do the following. We generate n R.V. from uniform distribution [0,1] using Mathematica random number generator. Then we check if each number is $< \frac{1}{2}$ or not, and we generate 1 or -1 as the case may be. We then generate n random variables from the exponential distribution, which we know how to do from part (a). Then we multiply the above 2 vectors, element wise, with each others. The first vector being the vector of 1's and -1's. And the second vector being the RV's from the exponential distribution. This is the algorithm

Algorithm
Input: λ,n (number of random variables to generate)
output: list of random numbers which belong to density $f(x) = \frac{\lambda}{2} e^{-\lambda|x|}$

Seed the random number generator with unique value for us.
A = Generate n random numbers from the exponential distribution with parameter λ (CALL problem 1 part(a) with the input λ,n) This uses F^{-1}method and uniform random number generator as well.
B = Generate n random numbers from uniform random number generator [0,1]

```
FOR i in 1..n LOOP  -- Note: This is algorithm view. In code 'vectored' operation is used.
    IF B(i)<.5 THEN
       B(i) = 1
     ELSE
       B(i) = -1
     END IF
 END LOOP

 result = B * A
```

Now generate a histogram from the result above.
The following function implements the above algorithm

Code Implementation

Define the function F^{-1}which was derived earlier. This is the inverse of the CDF of the exponential density function $\lambda e^{-\lambda x}$

```
Remove["Global`*"];
gDebug = False;
```

```
inverseCDFofExponentialDistribution[λ_, n_] := Module[{}, -1/λ Log[1 - n]]
```

Function below is called to generate N random numbers using the above F^{-1} function (User needs to seed before calling

```
getRandomNumbersFromExponential[λ_, nRandomVariables_] := Module[{i},
  Table[ inverseCDFofExponentialDistribution[λ, RandomReal[]], {i, nRandomVariables}]]
```

```
getRandomNumbersFromDoubleExponential[λ_, numberOfRandomVariables_] := Module[{W, T},
  W = getRandomNumbersFromExponential[λ, numberOfRandomVariables];
  T = Table[If[RandomReal[] < .5, 1, -1], {i, numberOfRandomVariables}];
  W T]
```

Test the above function by plotting the histogram generated for say $n = 10\,000$ overlaid by the true double exponential density function.
First, define the double exponential function

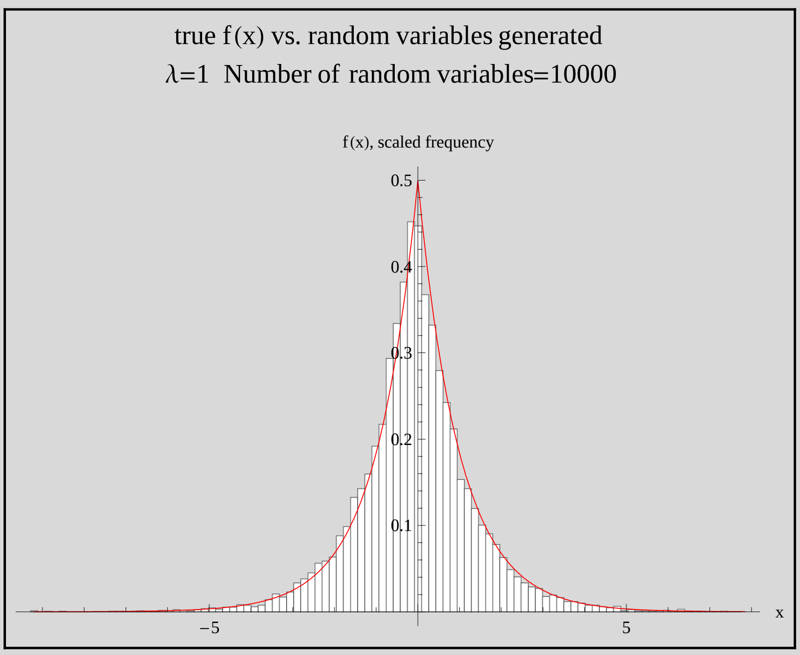
```
doubleExponential[λ_, x_] := λ/2 Exp[-λ Abs[x]]
```

Now do the overlay plot

This function makes a histogram which is scaled to be used to overlay density plots, or other functions.
Input: originalData: this is an array of numbers which represents the data to bin
nBins: number of bins
output: the histogram itself but scaled such that area is ONE

```
Needs["BarCharts`"]
nmaMakeDensityHistogram[originalData_, nBins_] :=
 Module[{freq, binSize, from, to, scaleFactor, j, a, currentArea},
  to = Max[originalData];
  from = Min[originalData];
  binSize = (to - from) / nBins;
  freq = BinCounts[originalData, binSize];
  currentArea = Sum[binSize * freq[[i]], {i, nBins}];
  freq = freq / currentArea;
  a = from;
  Table[{a + (j - 1) * binSize, freq[[j]], binSize}, {j, 1, nBins}]
 ]
```

```
SeedRandom[010101];
n = 10000; λ = 1; nBins = 100; imageSize = 400;

postprocessPartThreeA[listOfRandomNumbers_, λ_, nBins_, imageSize_] :=
 Module[{gz, pList, xFrom, xTo},
  xFrom = Min[listOfRandomNumbers];
  xTo = Max[listOfRandomNumbers];
  gz = nmaMakeDensityHistogram[listOfRandomNumbers, nBins];
  pList = GeneralizedBarChart[gz, BarStyle → White, ImageSize → imageSize];
  p = Plot[doubleExponential[λ, x], {x, xFrom, xTo}, AxesOrigin → {0, 0}, PlotRange → All,
    ImageSize → imageSize, (*PlotStyle→{Dashed,Red}*)PlotStyle → {Red}];

  Show[{pList, p},
   PlotLabel → Style["true f(x) vs. random variables generated\n" <> " λ=" <> ToString[λ] <>
      "  Number of random variables=" <> ToString[Length[listOfRandomNumbers]] <> "\n", 14],
   AxesLabel → {"x", "f(x), scaled frequency"}]
 ]

Framed[
 postprocessPartThreeA[getRandomNumbersFromDoubleExponential[λ, n], λ, nBins, imageSize]]
```



Problem 3 part(b)

In this part, we need to generate a list of r.v's that belong to normal distribution N(0,1), using uniform random number generator U[0,1] and using the random numbers generated from the double exponential density function in part (a) above. We are asked to use the accept/reject method.
First the method is explained, then the algorithm outlined, then the implementation shown and a test case given, then a GUI interface written to test the algorithm for different parameters values.

Accept/Reject algorithm
input: n (number of random variables to generate)
λ (the exponential density parameter)
$f(x)$ the density function for random variable X which we wish to generate random variables
$f_M(x)$ the density which we will use to help in generating the random variables from $f_X(x)$. This density is such that it is easy to generate random variables from. Much easier that from f(x) and that is why it was selected.
output: list of random numbers of length n from $f(x)$

Step 1: Find C. Where $c = \sup \forall x \frac{f_X(x)}{f_M(x)}$ To solve this, this is the algorithm

Algorithm for step 1: Let $f_M(x) = \frac{\lambda}{2} e^{-\lambda x}$ (since double exponential is symmetric, I'll use one sided version). Let $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$. Now find the ratio $r(x) = \frac{f_X(x)}{f_M(x)} = \frac{\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}}{\lambda e^{-\lambda x}}$ Now find where the maximum of this ratio is using normal calculus method: Take the derivative w.r.t. x and set it to zero. Solve the resulting equation for x. Evaluate the ratio at this root. This gives C. We find that $C = 1.31549$ The following few lines of code finds C:

```
λ = 1; fm = λ/2 Exp[- x]; fx = PDF[NormalDistribution[0, 1], x];
ratio = fx/fm;
root = First@Solve[D[ratio, x] == 0, x];
c = N[ratio /. root]
```

```
1.31549
```

Step 2: Now that we found C in step 1, then the envelop function becomes $C * f_M(x) = C \frac{\lambda}{2} e^{-\lambda|x|}$

Step 3: seed the number random generator

initialize an array d of size *n* to contain all the accepted random numbers generated

initialize counter number_accepted=0

WHILE number_accepted < n DO

generate r.v. from U[0,1] call it *u*.

Generate r.v. from double exponential density (using part(a)) call this x

IF $u * C * f_M(x) < f_X(x)$ THEN

d[i]=x

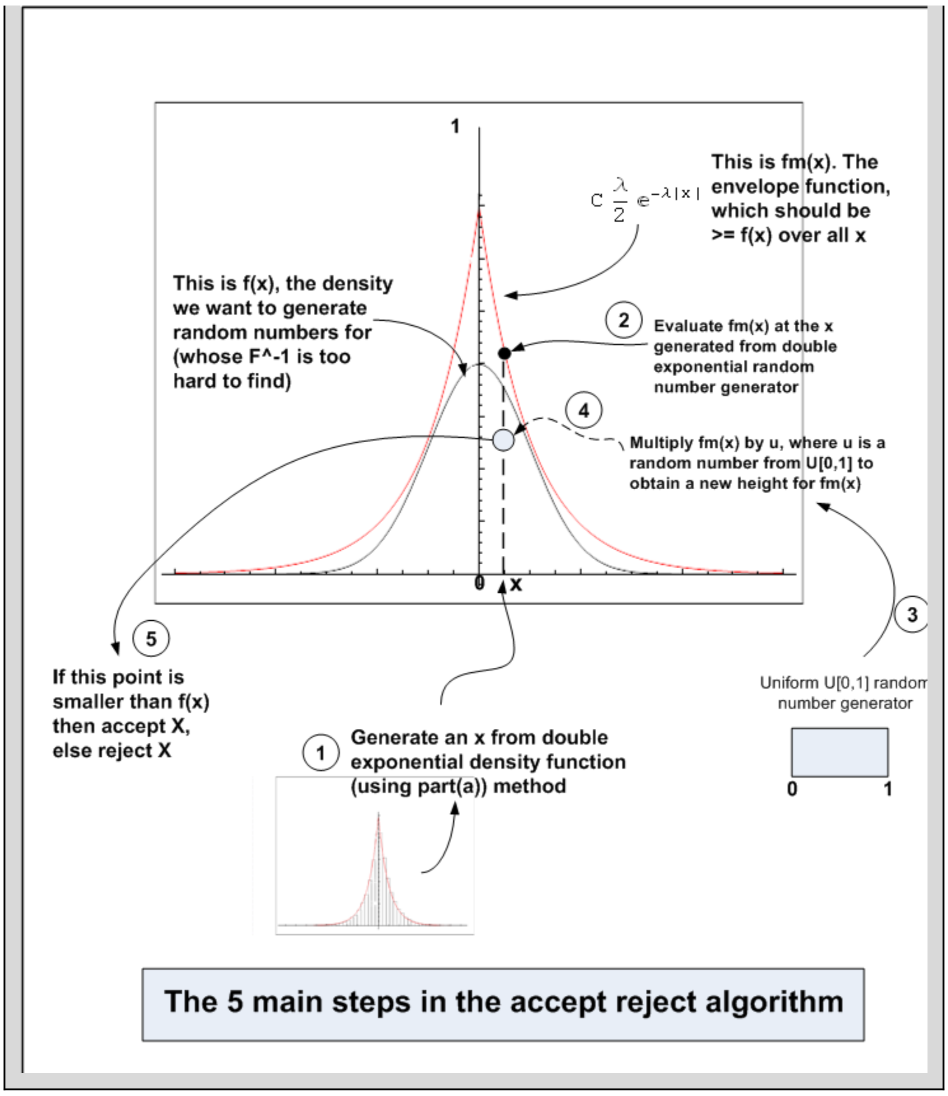
number_accepted++

END IF

END LOOP

Step 4: Now array d contains the n random numbers generated from the normal density N[0,1]. Make histogram and overlay it over N[0,1]

Diagram showing main steps in the algorithm



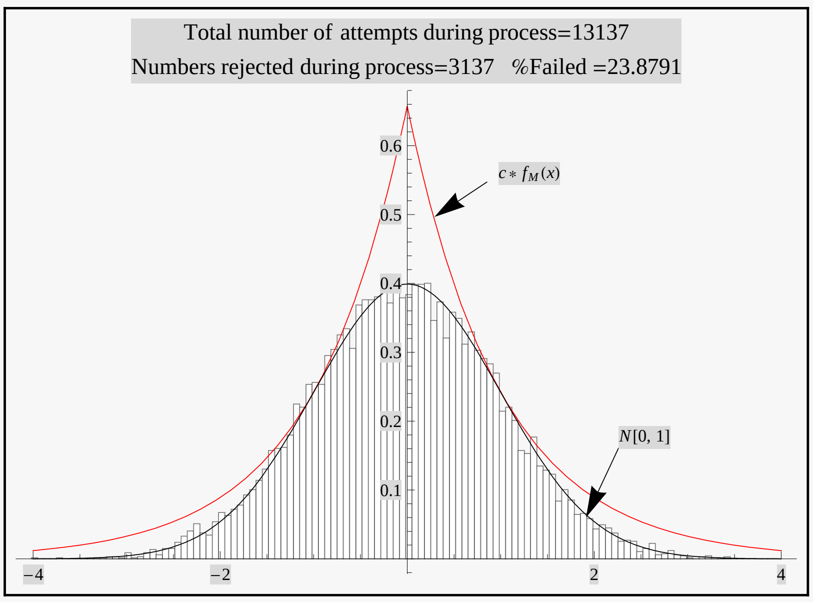
Accept Reject Algorithm Implementation

```
acceptReject[λ_ , numberOfRandomNumbersToGenerate _ , c_
  (*This is for scaling the envelope with so that envelope ≥ f(x) everywhere*)
  , μ_ (*mean of Normal Dist*), σ_ (*std of normal dist*)
 ] := Module[{nFailed = 0, nPassed = 0, y, x, d, i, maxEnvelope, fx, u},
  RandomSeed[010 101]; (*start from clean random number generator*)
  maxEnvelope = c * doubleExponential[λ, 0];
  d = Table[0, {i, numberOfRandomNumbersToGenerate}];

  While[nPassed < numberOfRandomNumbersToGenerate,
   {x = getRandomNumbersFromDoubleExponential[λ, 1][[1]];
    y = c * doubleExponential[λ, x] * RandomReal[{0, 1}];
    fx = PDF[NormalDistribution[μ, σ], x];
    If[y ≤  fx, {nPassed ++, d[[nPassed]] = x}, nFailed ++];}
  ];

  {d, nFailed}
 ]
```

Test case for n=10,000

Test the above function, and make a plot of histogram overlaid on top of density of N(0,1)

```
λ = 1; μ = 0; σ = 1; xFrom = -4 σ; xTo = 4 σ; n = 10 000;
c = 1.315489246958914; (*see algorithm above on how C was found*)
nBins = 120;
Clear[x];

{listOfNumbers, nFailed} = acceptReject[λ, n, c, μ, σ];

gz = nmaMakeDensityHistogram[listOfNumbers, nBins];
pList = GeneralizedBarChart[gz, BarStyle → White, ImageSize → 400, PlotRange → All];

p = Plot[{c * doubleExponential[1, x], PDF[NormalDistribution[0, 1], x]},
   {x, xFrom, xTo}, PlotRange → All, PlotStyle → {Red, Black}, ImageSize → 400];

Framed[Show[{pList, p},
  PlotLabel → "Total number of attempts during process=" <> ToString[n + nFailed] <>
    "\nNumbers rejected during process=" <> ToString[nFailed] <> "   %Failed =" <>
    ToString[nFailed / (n + nFailed) * 100.]]]
```



The above is a plot showing the histogram for random numbers generated using the accept - reject method for N=10,000. The random numbers are very close the N[0,1] which indicates this method is working well. The larger N is, the more closely the random numbers histogram will approach N[0,1] probability density.

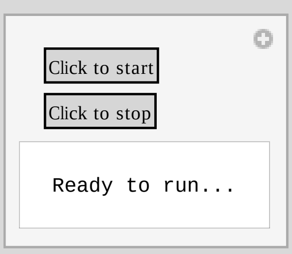
I have implemented a GUI based simulation as well for the above problem, please see the appendix below to run the simulation part.

Problem 3 simulation

```
Module[{gnTrialsSoFar = 0, gnRejectSoFar = 0, gnAcceptedSoFar = 0, gmaxEnvelope, gmultiplier,
  gλ, gμ, gσ, gAcceptedXSet, gnBins, gxFrom, gxTo, gAcceptedPointsCoordinates, gMaxAccepted},

 initializeSimulation[] := Module[{},
   RandomSeed[010 101];
   gmultiplier = 1.315489246958914;
   gnBins = 40;
   gλ = 1;
   gμ = 0;
   gσ = 1;
   gxFrom = -6 gσ;
   gxTo = 6 gσ;
   gMaxAccepted = 10 000;
   gnTrialsSoFar = 0; gnRejectSoFar = 0; gnAcceptedSoFar = 0;
   gAcceptedXSet = Table[0, {i, gMaxAccepted}];
   gAcceptedPointsCoordinates = Table[0, {i, gMaxAccepted}];
   gmaxEnvelope = gmultiplier * doubleExponential[gλ, 0]
  ];

 finalizeSimulation[] := Module[{gz, p, pList, x, res},
   If[gnTrialsSoFar > 0,
    {
     gz = nmaMakeDensityHistogram[gAcceptedXSet⟦1 ;; gnAcceptedSoFar⟧, gnBins];

     pList = GeneralizedBarChart[gz,
       BarStyle → White,
       ImageSize → 250,
       PlotRange → {{-4.5 gσ, 4.5 gσ}, {0, 1.2}}];

     p = Plot[PDF[NormalDistribution[gμ, gσ], x],
       {x, gxFrom, gxTo},
       PlotStyle → Red,
       PlotRange → All];

     res = Show[{pList, p},
       PlotLabel → "Total number of attempts during process=" <> ToString[gnTrialsSoFar] <>
         "\nNumbers accepted during process=" <> ToString[gnAcceptedSoFar] <>
         "   %Accepted =" <>
         ToString[gnAcceptedSoFar / (gnTrialsSoFar) * 100.] <>
         "\nNumbers rejected during process=" <>
         ToString[gnRejectSoFar] <> "   %Failed =" <>
         ToString[gnRejectSoFar / (gnTrialsSoFar) * 100.]
      ];
    }
    ,
    res = "Ready...";
   ];
   res
  ];

 processOneAcceptReject[] := Module[{x, y, fx, res, p, accepted, p2, pStats},
   gnTrialsSoFar ++;
   If[gnTrialsSoFar < 10, Return["Ready.."]];

   x = getRandomNumbersFromDoubleExponential[gλ, 1][[1]];
```

```
    y = gmultiplier * doubleExponential[gλ, x] * RandomReal[{0, 1}];
    fx = PDF[NormalDistribution[gμ, gσ], x];
    If[y ≤ fx
     , {gnAcceptedSoFar ++;
      accepted = True;
      gAcceptedXSet⟦gnAcceptedSoFar⟧ = x;
      gAcceptedPointsCoordinates⟦gnAcceptedSoFar⟧ = {x, y}
     },
     {gnRejectSoFar ++, accepted = False}
    ];

    p = Plot[{gmultiplier * doubleExponential[1, x], PDF[NormalDistribution[0, 1], x]},
      {x, gxFrom, gxTo}, PlotRange → All, PlotStyle → {Red, Black}, ImageSize → 250,
      PlotLabel → Row[{"Trial [", gnTrialsSoFar, "]\tc=[", gmultiplier, "]\n", If[y ≤ fx,
          Style["Accepted", Black], Style["Rejected", Red]], "\tPoint=(", x, " , ", y, ")"}],
      Epilog → {If[accepted, {PointSize[Large], Green, Point[{x, y}],
          {PointSize[Small], Gray, Point[gAcceptedPointsCoordinates⟦1 ;; gnAcceptedSoFar⟧]}},
         {PointSize[Large], Red, Point[{x, y}], {PointSize[Small], Gray,
           Point[gAcceptedPointsCoordinates⟦1 ;; gnAcceptedSoFar⟧]}}
        ]
       }
     ];
    p2 = finalizeSimulation[];
    pStats = Row[{"Trial [", gnTrialsSoFar, "]\n", If[y ≤ fx,
        Style["Accepted", Black], Style["Rejected", Red]], "\tPoint=(", x, " , ", y, ")"}];
    (*Grid[{ {pStats},{Grid[{ {p,p2}  }]}},Frame→All,Alignment→{Center}];*)
    Grid[{  {p, p2} }, Frame → All, Alignment → {Center}]
   ]
 ]
```

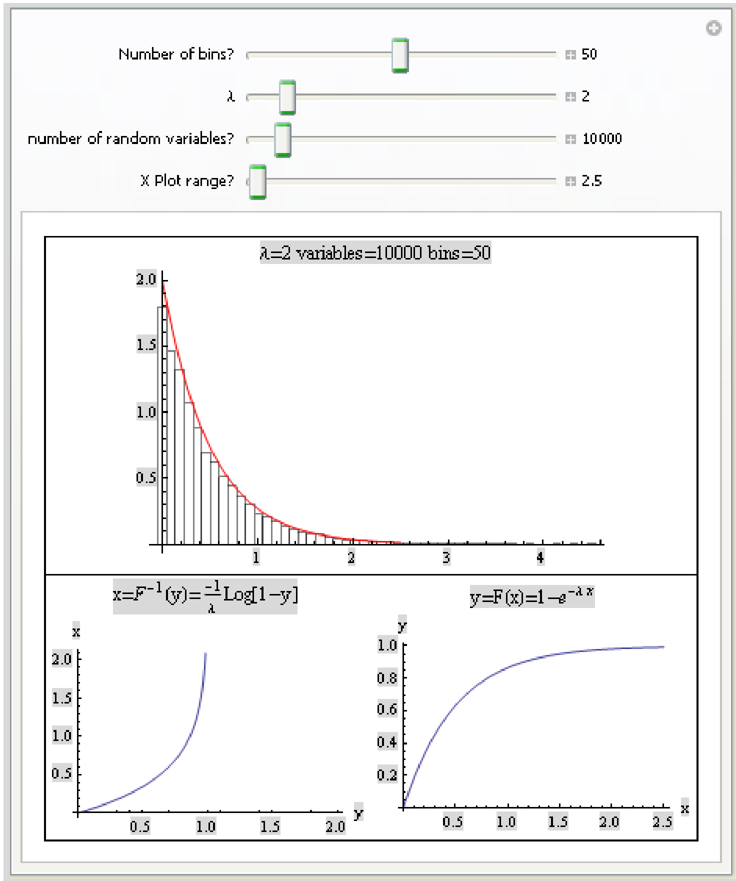
```
m = Manipulate[res = "Ready to run..."; runIt = False; i = 0;
  Dynamic[
   If[runIt && Not[stopIt] && i < 10 000,
    (i++; res = processOneAcceptReject[]),
    res
   ]
  ],
  {{runIt, True, ""}, Button[Style["Click to start", 10], {i = 0;
      initializeSimulation[]; stopIt = False; runIt = True}] &, ContinuousAction -> False},
  {{stopIt, False, ""}, Button[Style["Click to stop", 10], {stopIt = True; res}] &,
   ContinuousAction -> False}, AutorunSequencing → {{2, 120}}
 ]
```



Mathematica notebook

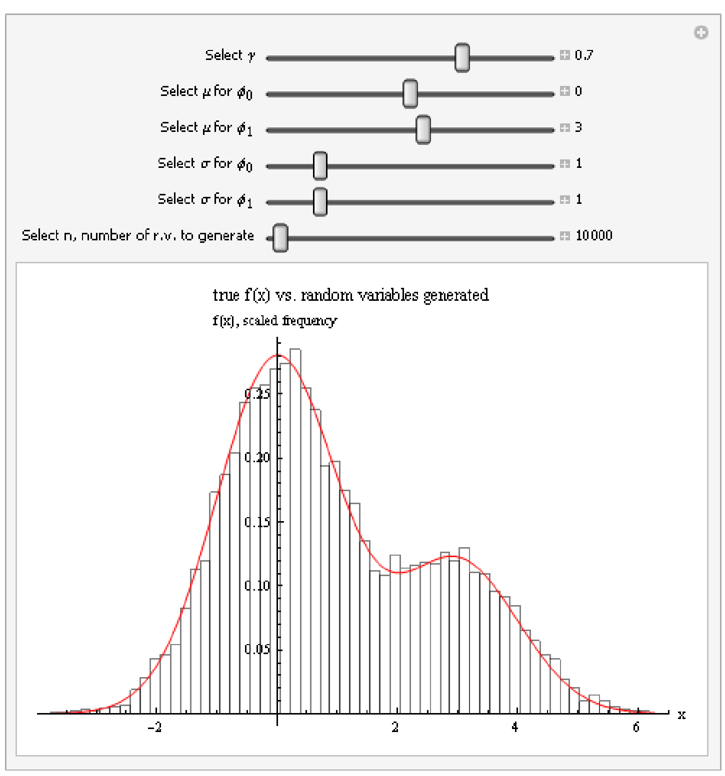
2.3.3 Simulation movies

2.3.3.1 Simulation problem 1. Random variables from exponential distribution using F^{-1} method



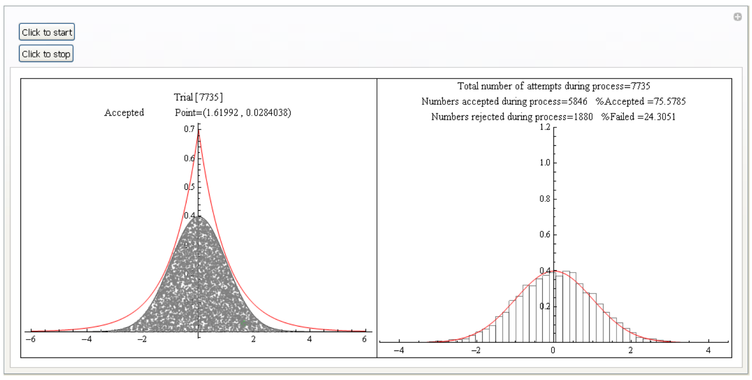
Movie swf

2.3.3.2 Simulation problem 2 (Random numbers from mixture)



Movie swf

2.3.3.3 Simulation Problem 3 (Accept/Reject)



Movie swf

2.4 Project 2

Problem: Simulation for estimator to estimate population size as sample size and number of samples taken is changed. Estimator for population size used in `2*sample_mean-1`

Chapter 3

HWs

Local contents

3.1 List of HWs to do

Math 502AB (Probability and Statistics) Homework - Fall 2007

From John Rice's book, Third Edition: Problems with an asterisk are new in the 3^{rd} Edition, as compared to the 2^{nd} Edition.

Ch.	Sec.	Exercise Numbers	Due Date
1	2-3	2, 6, 7, 9	
	4	11, 16, 18, 35b, 42	
	5	45, 53, 54, 64, 79	
2	1	6, 11, 19, 21, 27, 28, 31	
	2	34, 39, 44, 49	
	3	59, 60, 62, 66, 67, 70, 71, 72 (You may use Matlab or R)	
3	1, 2, 3	1, 3, 6, 8ab, 9a, 12ab, 15abcd	
	4	14a, 19, 15d	
	5	1b, 8c, 9b, 12c, 14b, 15e, 20, 21, 22, 24, 32*, 33a*, 34*, 37 (Problem 29 in 2^{nd} Ed., and there is an error in the expression of the density. Change the value 6 to 15/16), 38 (Problem 30 of 2^{nd} Ed.), 40 (don't do the expectation question)	
	6	11, 42a (Problem 32a in 2^{nd} Ed.), 44(Problem 34 in 2^{nd} Ed.), 47(Problem 37 in 2^{nd} Ed.), 52(Problem 42 in 2^{nd} Ed.), 55(Problem 45 in 2^{nd} Ed.), 58(Problem 48 in 2^{nd} Ed.), 64(Problem 54 in 2^{nd} Ed.)	
	7	65(Problem 55 in 2^{nd} Ed.), 68(Problem 58 in 2^{nd} Ed.), 70 (Problem 60 in 2^{nd} Ed.)	
4	1, 2	1, 2, 6, 8, 10, 12, 13, 16, 18, 22, 25, 26, 30, 32, 35, 42(Problem 38 in 2^{nd} Ed), 49(Problem 45 in 2^{nd} Ed), 50(Problem 46 in 2^{nd} Ed), 57(Problem 51 in 2^{nd} Ed)	
	3	43(Problem 39 in 2^{nd} Ed), 45(Problem 41 in 2^{nd} Ed), 46(Problem 42 in 2^{nd} Ed), 60(Problem 54 in 2^{nd} Ed)	

Figure 3.1: page 1

	4, 5	20, 63(Problem 57 in 2nd Ed), 66(Problem 60 in 2nd Ed), 67(Problem 61 in 2nd Ed), 68(Problem 62 in 2nd Ed), 70(Problem 64 in 2nd Ed), 75(Problem 69 in 2nd Ed), 77(Problem 71 in 2nd Ed), 83(Problem 77 in 2nd Ed), 84(Problem 78 in 2nd Ed), 85(Problem 79 in 2nd Ed), 90(Problem 84 in 2nd Ed), 92(Problem 86 in 2nd Ed), 96(Problem 90 in 2nd Ed)	
	6	101(Problem 95 in 2nd Ed), 102(Problem 96 in 2nd Ed)	
5		1, 3, 5, 6, 10, 12, 13, 14, 21, 25	
6		2, 5, 6, 7, 8, 10, 11	
8	8.1-8.4	4 (ab)*, 7(a) (Problem 5 in 2nd Ed), 13 (Problem 11 in 2nd Ed), 14, 16a (Problem 14 in 2nd Ed), 18a(Problem 16 in 2nd Ed), 20(Problem 18 in 2nd Ed), 21a(Problem 19 in 2nd Ed), 23(MME only) (Problem 21 in 2nd Ed), 32(Problem 28 in 2nd Ed), 43(a-e, MME only)*, 50a(Problem 42 in 2nd Ed), 53a(Problem 45 in 2nd Ed)	
	8.5	4 (cd)*, 6 (ac)(Problem 4 in 2nd Ed), 7(bc), 8* (except c), 11 (Problem 9 in 2nd Ed), 12 (Problem 10 in 2nd Ed), 16bc, 18bc(Problem 16 in 2nd Ed), 21b(Problem 19 in 2nd Ed), 23(MLE only) (Problem 21 in 2nd Ed), 27(Problem 25 in 2nd Ed), 30*, 43(a-e, MLE only)*, 48*, 50(bc) (Problem 42 in 2nd Ed), 51, 53(b-d) (Problem 45 in 2nd Ed), 60(Problem 52 in 2nd Ed)	
	8.6	4 e*, 7d*, 62*, 63*	
	8.7-8.8	6b(Problem 4 in 2nd Ed), 16d, 18d, 21c (Problem 19 in 2nd Ed), 68, 71, 73	
9		3, 7, 9, 12, 13, 24(Problem 16 in 2nd Ed), 26(Problem 18 in 2nd Ed), 28(Problem 20 in 2nd Ed), 29(Problem 21 in 2nd Ed), 33(Problem 25 in 2nd Ed), 36(Problem 28 in 2nd Ed),41	

Figure 3.2: page 2

3.2 HW 1

3.2.1 1

HW1 Mathematics 502

By Nasser Abbasi

Problem 2, page 27

- **Question**
- **Answer**

a) This is a list of the sample space. Simply toss one die, then make a toss of the second die. The result is as shown:

```
s = {1, 2, 3, 4, 5, 6};
Table[{s[[i]], s[[j]]}, {i, 6}, {j, 6}];
space = Flatten[%, 1]
```

```
{{1, 1}, {1, 2}, {1, 3}, {1, 4}, {1, 5}, {1, 6}, {2, 1}, {2, 2}, {2, 3}, {2, 4}, {2, 5}, {2, 6},
 {3, 1}, {3, 2}, {3, 3}, {3, 4}, {3, 5}, {3, 6}, {4, 1}, {4, 2}, {4, 3}, {4, 4}, {4, 5}, {4, 6},
 {5, 1}, {5, 2}, {5, 3}, {5, 4}, {5, 5}, {5, 6}, {6, 1}, {6, 2}, {6, 3}, {6, 4}, {6, 5}, {6, 6}}
```

part b)
(1) This is event A. Look through each outcome in space and see if first+second die is less than or equal to 5

```
setA = Select[space, First[#] + Last[#] ≤ 5 &]
```

```
{{1, 1}, {1, 2}, {1, 3}, {1, 4}, {2, 1}, {2, 2}, {2, 3}, {3, 1}, {3, 2}, {4, 1}}
```

(2) This is event B, Look through each outcome in space and see if first die larger than second die

```
setB = Select[space, First[#] > Last[#] &]
```

```
{{2, 1}, {3, 1}, {3, 2}, {4, 1}, {4, 2}, {4, 3}, {5, 1},
 {5, 2}, {5, 3}, {5, 4}, {6, 1}, {6, 2}, {6, 3}, {6, 4}, {6, 5}}
```

(3) This is event C, Look through each outcome in space and see if first die is 4

```
setC = Select[space, First[#1] == 4 &]
```

```
{{4, 1}, {4, 2}, {4, 3}, {4, 4}, {4, 5}, {4, 6}}
```

Part c)
(1) This is $A \cap C$, which means event is in A and in C

```
setA ⋂ setC
```

```
{{4, 1}}
```

(2) This is $B \cup C$, which means event in B or in C or in both

```
setB ⋃ setC
```

```
{{2, 1}, {3, 1}, {3, 2}, {4, 1}, {4, 2}, {4, 3}, {4, 4}, {4, 5},
 {4, 6}, {5, 1}, {5, 2}, {5, 3}, {5, 4}, {6, 1}, {6, 2}, {6, 3}, {6, 4}, {6, 5}}
```

(3) This is $A \cap (B \cup C)$ which is A intersect B union C, i.e. event in A and also in B union C. First find $B \cup C$, which is event in B or C or both

```
setB ⋃ setC
```

```
{{2, 1}, {3, 1}, {3, 2}, {4, 1}, {4, 2}, {4, 3}, {4, 4}, {4, 5},
 {4, 6}, {5, 1}, {5, 2}, {5, 3}, {5, 4}, {6, 1}, {6, 2}, {6, 3}, {6, 4}, {6, 5}}
```

now find event in A or in the above or in both

```
setA ⋂ (setB ⋃ setC)
```

```
{{2, 1}, {3, 1}, {3, 2}, {4, 1}}
```

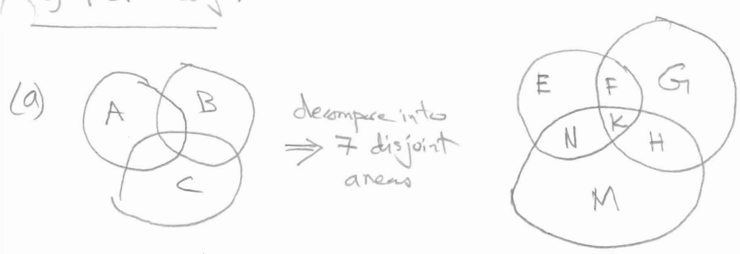
3.2.2 2

#6 Verify: $P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$

(a) using Venn diagram

(b) Formally.

(a)



So from Venn diagram, we see that we adding area A+B we are adding (F+K) area twice. so need to subtract this area once which is represented by $P(A)\cap P(B)$.
Similarly area A+C has added area (N+K) twice, so need to subtract this once, which is $P(A)\cap P(C)$.
Similarly, when area (K+H) is added twice between B, C. hence need to subtract (K+H) once, which is $P(C)\cap P(B)$.
But now we have added area (K) 3 times, and subtracted it 3 times also therefore, need to add it again once. but area (K) is represented by $P(A\cap B\cap C)$.

Now do the proof using formal argument →

(b) Since area E, F, G, K, N, H, M are disjoint, then from third axiom:

$$P(A \cup B \cup C) = P(E)+P(F)+P(G)+P(K)+P(N)+P(H)+P(M). \quad (1)$$

But $A = E+F+K+N$ and E, F, K, N are disjoint, so

$$P(A) = P(E)+P(F)+P(K)+P(N) \quad (2)$$

Similarly, $B = G+F+K+H$, hence

$$P(B) = P(G)+P(F)+P(K)+P(H) \quad (3)$$

Similarly $C = M+H+K+N$ hence

$$P(C) = P(M)+P(H)+P(K)+P(N) \quad (4)$$

add (2)+(3)+(4)

$$\boxed{P(A)+P(B)+P(C) = P(E)+P(G)+P(M)+2P(F)+2P(H)+2P(N)+3P(K)}$$

But on RHS we see that $P(A \cup B \cup C) = P(E)+P(F)+P(G)+P(K)+P(N)+P(H)+P(M)$. (5)

Hence (5) becomes

$$\boxed{P(A)+P(B)+P(C) = P(A \cup B \cup C)+P(F)+P(H)+P(N)+2P(K)} \quad (6)$$

Now we see that $P(F)+P(K) = P(A \cap B)$

and $P(H)+P(K) = P(B \cap C)$

and $P(N)+P(K) = P(A \cap C)$.

hence (6) becomes

$$\boxed{P(A)+P(B)+P(C) = P(A \cup B \cup C)+P(A \cap B)+P(B \cap C)+P(A \cup C)-P(K)} \rightarrow$$

But $P(K) = P(A \cap B \cap C)$.

Therefor we have

$$\boxed{P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cup C) + P(A \cap B \cap C)}$$

QED

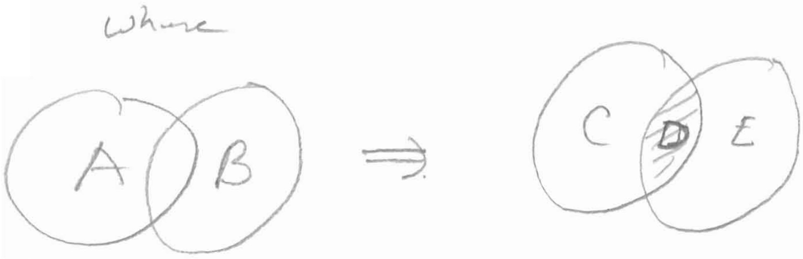
#7 prove Bonferroni inequality.

$$P(A\cap B) \geq P(A)+P(B)-1$$

answer we already know that

$$P(A\cup B) = P(A)+P(B) - P(D) \quad \text{—} (1)$$

where



i.e D is $A\cap B$.

this is from property D of probability measure so shown on page 5 of text Book.

Therefor we have from (1)

$$\boxed{P(A\cup B) = P(A)+P(B) - P(A\cap B)}$$

so $P(A\cap B) = P(A)+P(B) - P(A\cup B)$ —— (2)

now, $\boxed{P(A\cup B) \leq 1}$ let $P(A\cup B)=K$. so (2) becomes

$$P(A\cap B) = P(A)+P(B) - K$$

since $K\leq 1$ Then the above becomes

$$\boxed{P(A\cap B) \geq P(A)+P(B)-1}$$

QED

#9 Weather forecaster says prob. of rain on Saturday is 25% and prob. of rain on Sunday is 25%. Find prob of rain during weekend.

<u>Answer</u> Let $\tilde{S}$ = rainy Saturday.
S = not rainy Saturday
$\tilde{U}$ = rainy sunday.
U = not raing sunday.

hence Sample space is $\Omega = \{\tilde{S}\tilde{U}, \tilde{S}U, S\tilde{U}, SU\}$

i.e Ω

this is the Event E that represent rain occured during weekend.

So $\boxed{P(E) = P(\tilde{S}\tilde{U}) + P(\tilde{S}U) + P(S\tilde{U})}$ by definition. — (1)

but $P(\tilde{S}\tilde{U}) = P(\tilde{U}|\tilde{S})\ P(\tilde{S})$ by conditional prob.

but $P(\tilde{U}|\tilde{S}) = P(\tilde{U})$ since $\tilde{U}, \tilde{S}$ <u>are independent</u>

so $\boxed{P(\tilde{S}\tilde{U}) = P(\tilde{U})P(\tilde{S})}$

similarly $P(\tilde{S}U) = P(\tilde{S})\ P(U)$

and $P(S\tilde{U}) = P(S)\ P(\tilde{U})$

Hence (1) becomes

$$P(E) = P(\tilde{U})P(\tilde{S}) + P(\tilde{S})P(U) + P(S)\,P(\tilde{U})$$

$$= (.25)(.25) + (.25)(.75) + (.75)(.25)$$

$$= 0.4375$$

So $\boxed{43.75\%}$ chance it is Rainy weekend.

#11 if first 3 digits are 452. if all the sequences of the remaining 4 digits are equally likely. what is the probability that a randomly selected phone number contains 7 distinct digits?

$\{1,3,6,7,8,9,0\}$

Answer

4	5	2	·	·	·	·

Sample space Ω contains all possible outcomes of ways to select 4 digits. which is $\boxed{10^4}$ (since first 3 are fixed)

let A be the event of outcome which contains 7 digits that are all unique.

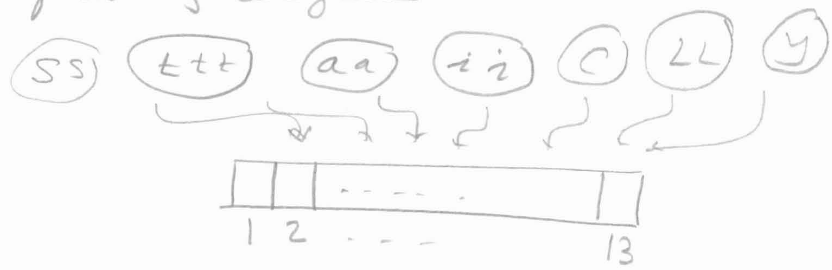
The first 3 digits are fixed. hence we have $1,3,6,7,8,9,0$ i.e 7 digits to select to fill the remaining 4 positions. We want to do this without replacement so as to have unique digits.

hence there are $7\times 6\times 5\times 4$ ways to do this.

$$\text{Hence } P(A) = \frac{7\times 6\times 5\times 4}{10^4} = 0.084$$
$$= \boxed{8.4\%}$$

we obtained from the letters of the word STATISTICALLY, using all letters?

Answer group letters that the same in separate groups. Hence we have the following diagram



we have a grouping of 7 groups.
lets starts with the ss group. there are $\binom{13}{2}$ ways to arrange these 2 letter, since ordering is not important.
after that lets consider ttt group. there are 11 slots left. hence there are $\binom{11}{3}$ ways to arrange these we continue this way to obtain:

$$\underset{'s'}{\binom{13}{2}}\underset{'t'}{\binom{11}{3}}\underset{'a'}{\binom{8}{2}}\underset{'i'}{\binom{6}{2}}\underset{'c'}{\binom{4}{1}}\underset{'L'}{\binom{3}{2}}\underset{'y'}{\binom{1}{1}}=\boxed{64{,}864{,}800}$$

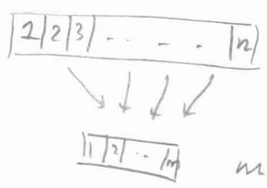
$\boxed{\text{PS.}}$ it does not matter which group we starting arranging i.e

$$\underset{'s'}{\binom{13}{2}}\underset{'t'}{\binom{11}{3}} \text{ is the same as } \underset{'t'}{\binom{13}{3}}\underset{'s'}{\binom{10}{2}}$$

#118

A Lot of n items contain k defectives, and m are selected randomly and inspected. How should the value of m be chosen so that the probability that at least one defective item turns up is .9 ? Apply your answer to (a) $n=1000$, $k=10$ and (b) $n=10{,}000$, $k=100$.

Answer



n items contains k defective.

m are slected.

(a) $n=1000$, $k=10$.

this mean if we pick one item, There is a 1% chance it is defective.

Let C_i be event of picking up an i^{th} item that is not defective.
So probability of picking up $\underline{m}$ items all not defective is

$$P[C_1 \cap C_2 \cap C_3 \cdots \cap C_m]$$

and so probability of at least one defective from m items is

$$\boxed{1 - P[C_1 \cap C_2 \cap \cdots \cap C_m]}$$

but $P[C_1 \cap C_2 \cdots \cap C_m] = P[C_2 \cap C_3 \cdots \cap C_m \mid C_1]\, P(C_1)$

Now once C_1 event occur, there will be $n-1$ items left, while k items remain defective so $\boxed{P(C_1) = \frac{n-k}{n}}$ while for $C_2 \cap C_3 \cdots \cap C_m$ now we will consider $n \to n-1$.

so $P(C_2 \cap C_3 \cdots \cap C_m) = P(C_3 \cap C_4 \cdots \cap C_m \mid C_2)\, P(C_2)$.

Now $\boxed{P(C_2) = \frac{n-1-k}{n-1}}$

we Continue this way. after m times, probability of picking m items all ok is

$$P(C_1)\,P(C_2)\cdots P(C_m) = \left(\frac{n-k}{n}\right)\left(\frac{n-1-k}{n-1}\right)\left(\frac{n-2-k}{n-2}\right)\cdots \left(\frac{n-(m-1)-k}{n-(m-1)}\right)$$

we are asked that defective item will appear has $p=0.9$.

So we need $1-\left(\frac{n-k}{n}\right)\left(\frac{n-1-k}{n-1}\right)\cdots\left(\frac{n-(m-1)-k}{n-(m-1)}\right)=0.9$

or $\boxed{\left(\frac{n-k}{n}\right)\left(\frac{n-1-k}{n-1}\right)\cdots\left(\frac{n-(m-1)-k}{n-(m-1)}\right)=0.1}$

Need to solve for $\underline{m}$. I can write a program and try for successive values of $\underline{m}$ until one is first found that satisfy the above.

doing the above program, this is the result

(a) $n=1000$, $k=10$ $\boxed{m=205}$

(b) $n=10{,}000$, $k=100$ $\boxed{m=227}$

35b

prove the following algebraically and by interpreting their meaning Combinatorially:

(b) $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$

Consider RHS: $\binom{n-1}{r-1} + \binom{n-1}{r} = \dfrac{(n-1)!}{(n-1-r+1)!\,(r-1)!} + \dfrac{(n-1)!}{(n-1-r)!\,r!}$

note that $(n-1)! = \dfrac{n!}{n}$ so above becomes

$$RHS = \frac{n!}{n\,(n-r)!\,(r-1)!} + \frac{n!}{n\,(n-1-r)!\,r!}$$

also $(r-1)! = \dfrac{r!}{r}$ and $(n-1-r)! = \dfrac{(n-r)!}{(n-r)}$ so above becomes

$$RHS = \frac{n!\,r}{n\,(n-r)!\,r!} + \frac{n!\,(n-r)}{n\,(n-r)!\,r!} = \frac{n!}{(n-r)!\,r!}\left[\frac{r}{n} + \frac{n-r}{n}\right]$$

$$= \frac{n!}{(n-r)!\,r!}\left[\frac{n}{n}\right] = \frac{n!}{(n-r)!\,r!} = \binom{n}{r} \quad \text{QED}$$

#42
How many ways can 11 boys on a soccer team be grouped into 4 Forwards, 3 midfielders, 3 defenders, 1 goalie?

this is a multinomial distribution.

$n_1 = 4, \ n_2 = 3, \ n_3 = 3, \ n_4 = 1$

so number of ways is $\binom{n}{n_1, n_2, n_3, n_4} = \frac{n!}{n_1!\, n_2!\, n_3!\, n_4!}$

$$= \frac{11!}{4!\ 3!\ 3!\ 1!} = \frac{11\times 10\times \cancel{9}^{3}\times \cancel{8}^{4}\times 7\times \cancel{6}^{2}\times 5\times \cancel{4}^{2}\times \cancel{3}\times \cancel{2}}{(\cancel{4}\times \cancel{3}\times \cancel{2})(\cancel{3}\times \cancel{2})(3\times 2)}$$

$= 11\times 10\times 3\times 4\times 7\times 5 = 46{,}200$ ways.

#45 show if the conditional probabilities exist

then $P(A_1 \cap A_2 \cap A_3 \cdots \cap A_n) =$

$P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cdots \cap A_{n-1})$

Answer

using $P(C \cap D) = P(C|D)\, P(D)$

(I will use $+$ for $\cap$)

$P(A_1 + A_2 + A_3 + \cdots + A_n) = P(A_n | A_1 + A_2 + \cdots + A_{n-1})\, P(A_1 + A_2 + \cdots A_{n-1})$

repeat the process on $P(A_1 + A_2 + \cdots + A_{n-1})$

$P(A_1 + A_2 + \quad + A_{n-1}) = P(A_{n-1} | A_1 + A_2 + \cdots + A_{n-1})\, P(A_1 + A_2 + \cdots + A_{n-2})$

Continue this way $\Rightarrow$

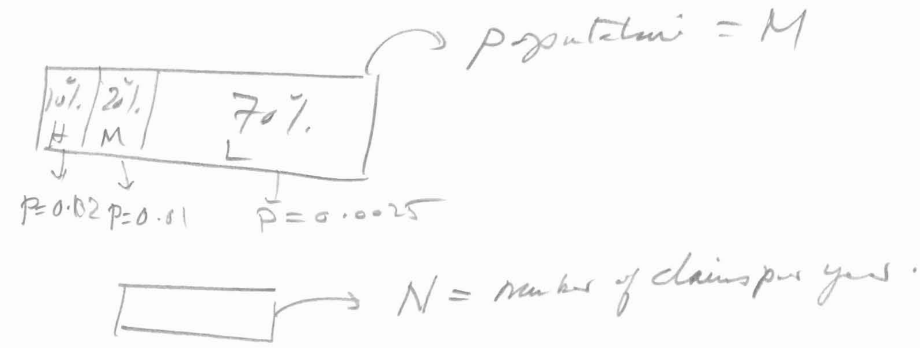
$P(A_1 + A_2 + \cdots + A_n) =$

$P(A_n | A_1 + \cdots + A_{n-1})\, P(A_{n-1} | A_1 + \cdots + A_{n-2}) \cdots P(A_2 | A_1)\, P(A_1)$ QED

53

Insurance Company has high-risk, medium Risk, Low Risk clients, who have respectively, probabilities 0.02, 0.01, 0.0025 of filing claims within a year. the proportions of the number of clients in the three Categories are 0.1, 0.2, 0.7 respectively. what proportion of the claims filed each year come from high Risk?

Answer



$$N = (10\%\ M)\, P(\text{High Risk}) + (20\%\ M)\, P(\text{Medium Risk}) + (70\%\ M)\, P(\text{Low Risk})$$

so $N = (0.1M)(0.02) + (0.2M)(0.01) + (0.7M)(0.0025)$.

but proportion of high Risk group filing a claim is

$$\frac{\text{Number of claims from high Risk}}{\text{Total number of claims}}$$

$$= \frac{(0.1M)(0.02)}{N} = \frac{(0.1)(0.02)}{(0.1)(0.02) + (0.2)(0.01) + (0.7)(0.0025)}$$

$= 0.3478$

hence $\boxed{34.78\%}$ of claims are from High Risk

#54 Consider Sequence of days and R_i event it rains on day i. Suppose $P(R_i|R_{i-1})=\alpha$ and $P(R_i^c|R_{i-1}^c)=\beta$

Suppose further that only today's weather is relevant to predicting tomorrow's.

a) if prob. it will rain today is p, what is prob. it will rain tomorrow?

b) what is prob. it will rain day after tommorow?

c) what is prob. it will rain n days from now? what happens as $n \to \infty$?

$\longrightarrow$

(a) $P(R_i) + P(R_i^c) = 1$

so $P(R_i) = 1 - P(R_i^c)$

$$P(R_i) = 1 - \left[P(R_i^c | R_{i-1}^c)\, P(R_{i-1}^c) \right]$$

but $P(R_i^c | R_{i-1}^c) = \beta$

and $P(R_{i-1}^c) = 1 - \alpha$.

so

$$P(R_i) = 1 - \left[\beta(1-\alpha) \right] = \boxed{1 - \beta + \beta\alpha}$$

#54

(a) Use Law of total probabilities. let $R_0 \equiv$ Rain today. $R_0^c \equiv$ No Rain today.

Then let $\Omega = \{ R_0, R_0^c \}$.

$$\text{so } P(R_1) = P(R_1|R_0)P(R_0) + P(R_1|R_0^c)\,P(R_0^c) \quad (1)$$

but $P(R_1|R_0) = \alpha$

$P(R_0) = p$

$P(R_0^c) = 1-p$

$P(R_1|R_0^c) = 1-\beta$ since $P(R_1^c|R_0^c) = \beta$

So (1) becomes

$$P(R_1) = \alpha p + (1-\beta)(1-p)$$

(c) Now apply the same principle again. Replace $R_0 \to R_1$, $R_1 \to R_2$, $\alpha \to P(R_1)$

so $\Omega = \{ R_1, R_1^c \}$

$$\text{so } P(R_2) = P(R_2|R_1)\,P(R_1) + P(R_2|R_1^c)\,P(R_1^c)$$

$$= \alpha\left[\alpha p + (1-\beta)(1-p)\right] + (1-\beta)\left[1 - (\alpha p + (1-\beta)(1-p))\right] \rightarrow$$

$$= \alpha^2 p + \alpha(1-\beta)(1-p) + (1-\beta) - (1-\beta)\left[\alpha p + (1-\beta)(1-p)\right]$$

$$= \alpha^2 p + \alpha(1-\beta)(1-p) + (1-\beta) - \left[\alpha p + (1-\beta)(1-p)\right] + \beta\left[\alpha p + (1-\beta)(1-p)\right]$$

$$= \alpha\left[\alpha p + (1-\beta)(1-p)\right] - \left[\alpha p + (1-\beta)(1-p)\right] + \beta\left[\alpha p + (1-\beta)(1-p)\right] + (1-\beta)$$

$$= (\alpha - 1 + \beta)\left[\alpha p + (1-\beta)(1-p)\right] + (1-\beta)$$

$$= (\alpha - 1 + \beta)\left[\alpha p + 1 - p - \beta + \beta p\right] + 1 - \beta$$

$$= \alpha^2 p + \alpha - \alpha p - \alpha\beta + \alpha\beta p - \alpha p - 1 + p + \beta - \beta p + \alpha\beta p + \beta - \beta p - \beta^2 + \beta^2 p + 1 - \beta$$

$$= \alpha^2 p + \alpha - 2\alpha p - \alpha\beta + 2\alpha\beta p + p - 2\beta p + \beta - \beta^2 + \beta^2 p$$

3.2.3 3

after n days

$$\Omega = \{ R_n, R_n^c \}$$

$$P(R_{n+1}) = P(R_{n+1} | R_n) P(R_n) + P(R_{n+1} | R_n^c) P(R_n^c)$$

we see that the Rule is $P(R_{i+1}) = \alpha P(R_i) + (1-\beta)(1-P(R_i))$

When $P(R_1) = \alpha p + (1-\beta)(1-p)$

$$P(R_2) = \alpha[\alpha p + (1-\beta)(1-p)] + (1-\beta)[1-(\alpha p + (1-\beta)(1-p))]$$

$$P(R_3) = \alpha\left[\alpha[\alpha p + (1-\beta)(1-p)] + (1-\beta)[1-(\alpha p + (1-\beta)(1-p))]\right]$$
$$+ (1-\beta)(1 - [\alpha(\alpha p + (1-\beta)(1-p)) + (1-\beta)(1-(\alpha p + (1-\beta)(1-p))])$$

$$\text{so } P(R_n) = \overbrace{\alpha[\alpha(\cdots}^{n} (\alpha p + (1-\beta)(1-p)] + [1-\beta](1-(\alpha p + (1-\beta)(1-p))$$
$$+ (1-\beta)[1 - (\overbrace{\alpha(\alpha(\cdots}^{n-1} (\alpha p + (1-\beta)(1-p) + (1-\beta)[1 - \cdots)]$$

So as $n \to \infty$

$$P(R_n) = 0 + 1 - \beta[\underbrace{\alpha(\alpha(\alpha \cdots)))}_{\to 0}]$$

$$= 0 + 1 - 0$$

$$= \underline{1}$$

3.2.4 4

#64

The probability axioms are:

1. $P(\Omega)=1$ where Ω is the sample space.
2. $A\subset\Omega \Rightarrow P(A)\geq 0$
3. if $A,B,C,\ldots$ are disjoint events in Ω, then

$$P(A\cup B\cup C\cup\cdots) = P(A)\ P(B)\ P(C)\cdots$$

So we need to show that $Q(A)\equiv P(A|B)$ satisfies the axioms.

$Q(A)=P(A|B)$

— to show $Q(\Omega)=1$

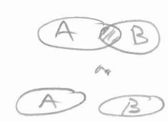
$$Q(\Omega)=P(\Omega|B) = \frac{P(\Omega\cap B)}{P(B)}$$

But $(\Omega\cap B)=B$ since $B\subseteq\Omega$

so $Q(\Omega)=\frac{P(B)}{P(B)}=\boxed{1}$

— to show $Q(A)\geq 0$. when $B>0$

$Q(A)\equiv P(A|B)$



if A,B disjoint then

$$Q(A)=\frac{P(A\cap B)}{P(B)}=\frac{P(\Phi)}{P(B)}=\frac{0}{P(B)}=\boxed{0}$$

if A,B not disjoint then

$Q(A)=\frac{P(A\cap B)}{P(B)}$ then $P(A\cap B)>0$ in this case.

$Q(A)=\frac{P(A\cap B)}{P(B)}>0$ so $\boxed{Q(A)\geq 0}$

#79

Many human diseases are genetically transmitted. here is an example model of disease. The genotype aa is diseased and dies before it mates. genotype Aa is carrier but is not diseased. AA is not carrier nor diseased.

(a) if 2 carriers mate, what is chance that their offspring are of each of the 3 genotypes?

(b) if male offspring of 2 carriers is not diseased, what is the probability that he is a carrier?

(c) suppose that the nondiseased offspring of part (b) mates with a member of the population for whom no family history is available and who is thus assumed to have propability p of being a carrier (p is very small number). what are the probabilities that their first offspring has genotype AA, Aa, aa?

(d) Suppose that first offspring of part (c) is Not diseased, what is the probability that the father is carrier?

	AA	Aa	aa
AA	{AA} $P(AA)=1$, $P(Aa)=0$ $P(aa)=0$	{AA, Aa} $P(AA)=1/2$, $P(Aa)=\frac{1}{2}$ $P(aa)=0$	{Aa} $P(AA)=0$, $P(Aa)=1$ $P(aa)=0$
Aa	{AA, Aa, AA, Aa} $P(AA)=\frac{1}{2}$ $P(Aa)=\frac{1}{2}$ $P(aa)=0$	{AA, Aa, Aa, aa} $P(AA)=\frac{1}{4}$ $P(Aa)=\frac{1}{2}$ $P(aa)=\frac{1}{4}$	{Aa, Aa, aa, aa} $P(AA)=0$, $P(Aa)=\frac{1}{2}$ $P(aa)=\frac{1}{2}$
aa	{Aa, Aa, Aa, Aa} $P(AA)=0$ $P(Aa)=1$ $P(aa)=0$	{Aa, Aa, aa, aa} $P(AA)=0$ $P(Aa)=\frac{1}{2}$ $P(aa)=\frac{1}{2}$	{aa} $P(AA)=0$ $P(Aa)=0$ $P(aa)=1$

b) offspring of 2 carriers as from Table:

$$\{Aa \underset{\text{mate}}{\longleftrightarrow} Aa) \Rightarrow \{AA, Aa, Aa, aa\}.$$

so $P(\text{Carrier} \mid \text{Not diseased}) = \dfrac{P(\text{Carrier} \cap \text{Not diseased})}{P(\text{Not diseased})}$

$$= \frac{\frac{2}{4} \to \{Aa, Aa\}}{\frac{3}{4} \to \{AA, Aa, Aa\}} = \boxed{\frac{2}{3}}$$

c) No diseased offspring of part (b) is $\{AA, Aa, Aa\}$

let m denote the mate from population for whom the probability of of being carrier is p.

let z denote the nondiseased offspring of part (b).

then we know that z can be only AA, Aa.

let X be the first offspring of m and z.

so
$$\begin{aligned}
P(X=AA) = {} & P(X=AA \mid m=AA \cap z=AA)\, P(m=AA \cap z=AA) \\
& + P(X=AA \mid m=AA \cap z=Aa)\, P(m=AA \cap z=Aa) \\
& + P(X=AA \mid m=Aa \cap z=AA)\, P(m=Aa \cap z=AA) \\
& + P(X=AA \mid m=Aa \cap z=Aa)\, P(m=Aa \cap z=Aa) \\
& + P(X=AA \mid m=aa \cap z=AA)\, P(m=aa \cap z=AA) \\
& + P(X=AA \mid m=aa \cap z=Aa)\, P(m=aa \cap z=Aa)
\end{aligned}$$

But $P(m=AA \cap z=AA) = P(m=AA)\, P(z=AA)$ since events are independent

but $P(m=AA) = 1-p$ since not carrier.

and $P(z=AA) = \frac{1}{3}$ from part (b). Since z can be $\{AA, Aa\}$ $\quad p=\frac{1}{3} \quad p=\frac{2}{3}$

and $P(m=Aa) = p$

and $P(m=AA \cap z=Aa) = P(m=AA)\, P(z=Aa) = (1-p)\frac{2}{3}$

and $P(m=Aa)\, P(z=AA) = p\,\frac{1}{3}$

Now $P(X=AA \mid m=AA \cap z=AA) = 1$ from Table.

$P(X=AA \mid m=AA \cap z=Aa) = \frac{1}{2}$

$P(X=AA \mid m=Aa \cap z=AA) = \frac{1}{2}$

$P(X=AA \mid m=Aa \cap z=Aa) = \frac{1}{4}$

$P(X=AA \mid m=aa \cap z=AA) = 0$

$P(X=AA \mid m=aa \cap z=Aa) = 0$

So
$$P(X=AA) = 1\,(1-p)\tfrac{1}{3} + \tfrac{1}{2}\,(1-p)\tfrac{2}{3} + \tfrac{1}{4}\,(p)\tfrac{1}{3} + \tfrac{1}{4}\,(p)\tfrac{2}{3} + 0 + 0$$

$$= \tfrac{1}{3}(1-p) + \tfrac{1}{3}(1-p) + \tfrac{1}{6}p + \tfrac{1}{6}p$$

$$= \tfrac{2}{3}(1-p) + \tfrac{1}{3}p$$

$$= \tfrac{2}{3} - \tfrac{2}{3}p + \tfrac{1}{3}p = \boxed{\tfrac{2}{3} - \tfrac{1}{3}p}$$

Now do same for $P(X=Aa)$

$P(X=Aa) = P(X=Aa \mid m=AA \cap z=AA)\ P(m=AA \cap z=AA)$

$+ \cdots$ the same as before.

only thing different is to find

$P(X=Aa \mid m=AA \cap z=AA) \longrightarrow 0$

and $P(X=Aa \mid m=AA \cap z=Aa) \longrightarrow \frac{1}{2}$

and $P(X=Aa \mid m=Aa \cap z=AA) \longrightarrow \frac{1}{2}$

and $P(X=Aa \mid m=Aa \cap z=Aa) \longrightarrow \frac{1}{2}$

and $P(X=Aa \mid m=aa \cap z=AA) \longrightarrow 0$ dies before it mates!

and $P(X=Aa \mid m=aa \cap z=Aa) \longrightarrow 0$ dies before it mates!

$$
\begin{aligned}
\text{so } P(X=Aa) &= 0\,(1-p)\tfrac{1}{3} \\
&+ \tfrac{1}{2}(1-p)\tfrac{2}{3} \\
&+ \tfrac{1}{2}(p)\tfrac{1}{3} \\
&+ \tfrac{1}{2}(p)\tfrac{2}{3} \\
&+ 0\,(1-p)\tfrac{1}{3} \\
&+ 0\,(1-p)\tfrac{2}{3}
\end{aligned}
$$

$$= \tfrac{1}{3}(1-p) + \tfrac{1}{6}p + \tfrac{1}{3}p$$

$$= \tfrac{1}{3} - \tfrac{1}{3}p + \tfrac{1}{6} + \tfrac{1}{3}p = \boxed{\tfrac{1}{3} + \tfrac{1}{6}p}$$

hence $P(X=aa) = 1 - \left[\left(\tfrac{1}{3} + \tfrac{1}{6}p\right) + \left(\tfrac{2}{3} - \tfrac{1}{3}p\right)\right]$

$$= 1 - \tfrac{1}{3} - \tfrac{1}{6}p - \tfrac{2}{3} + \tfrac{1}{3}p = \boxed{\tfrac{1}{6}p}$$

d) Need to find $P(\text{Father Carrier} \mid \text{first offspring Not diseased})$.

$$= \frac{P(\text{Father Carrier} \cap \text{first offspring Not diseased})}{P(\text{first offspring Not diseased})}$$

from part c, we found that first offspring has prob of $AA = \left(\frac{2}{3} - \frac{1}{3}p\right)$, $P(Aa) = \frac{1}{3} + \frac{1}{6}p$, $P(aa) = \frac{1}{6}p$.

so first offspring probability of Not diseased is

$$\left(\frac{2}{3} - \frac{1}{3}p\right) + \left(\frac{1}{3} + \frac{1}{6}p\right) = \boxed{1 - \frac{1}{2}p}$$

Prob Father is Carrier is $\frac{2}{3}$.

3.3 HW chapter 8

Maple worksheet

Mathematica notebook

3.4 Problem 84 chapter 8

Problem 90 chapter 4

by Nasser Abbasi 5:20 PM, oct 26, 2007

Problem:

Assuming that $X \sim N(0, \sigma^2)$ use the mgf to show that the odd moments are zero and the even moments are $\frac{(2n)!\ \sigma^{2n}}{2^n(n!)}$

Answer:

$$M_X(t) = e^{\frac{t^2\sigma^2}{2}}$$

First obtain a recursive formula for the moment generation function

$$\begin{aligned}
M_X^{(1)}(t) &= t\sigma^2 e^{\frac{t^2\sigma^2}{2}} \\
M_X^{(2)}(t) &= \sigma^2 e^{\frac{t^2\sigma^2}{2}} + \left(t\sigma^2\right)^2 e^{\frac{t^2\sigma^2}{2}} \\
&= \sigma^2 M_X(t) + t\sigma^2 M_X^{(1)}(t) \\
M_X^{(3)}(t) &= \sigma^2 M_X^{(1)}(t) + t\sigma^2 M_X^{(2)}(t) + \sigma^2 M_X^{(1)}(t) \\
&= 2\sigma^2 M_X^{(1)}(t) + t\sigma^2 M_X^{(2)}(t) \\
M_X^{(4)}(t) &= 2\sigma^2 M_X^{(2)}(t) + t\sigma^2 M_X^{(3)}(t) + \sigma^2 M_X^{(2)}(t) \\
&= 3\sigma^2 M_X^{(2)}(t) + t\sigma^2 M_X^{(3)}(t)
\end{aligned}$$

Hence

$$\boxed{M_X^{(r)}(t) = (r-1)\,\sigma^2 M_X^{(r-2)}(t) + t\sigma^2 M_X^{(r-1)}(t)}$$

Using the above, we generate odd and even moments.

$$\begin{aligned}
E\left(x^1\right) &= M_X^{(1)}(0) = 0 \\
E\left(x^2\right) &= M_X^{(2)}(0) = \sigma^2 M_X(0) = \sigma^2 \\
&\quad etc...
\end{aligned}$$

odd moments: Proof by induction., See class notes, Math 502 lecture 10/24/07

even moments:

proof by induction.

First show it is true for the base case $n = 1$.(base case)

From the above we see this is indeed the case because $E(x^2) = \sigma^2$, which is the same as saying $E(x^{2n}) = \frac{(2n)!\ \sigma^{2n}}{2^n(n!)}$ when $n = 1$

Now assume it is true for some $n \geq 1$, i.e. assume that

$$E\left(x^{2n}\right) = \frac{(2n)!\ \sigma^{2n}}{2^n\,(n!)} \tag{1}$$

Then we need to show that the relation is true for $n+1$ (the next even number), i.e. we need to show that

$$\boxed{E\left(x^{2(n+1)}\right) = \frac{(2(n+1))!\ \sigma^{2(n+1)}}{2^{(n+1)}((n+1)!)}} \tag{*}$$

Use the moment generation recursive formula to show the above. since from definition we know that

$$E\left(x^{2(n+1)}\right) = M_X^{2(n+1)}(0)$$

But we showed that

$$M_X^{(r)}(t) = (r-1)\,\sigma^2 M_X^{(r-2)}(t) + t\sigma^2 M_X^{(r-1)}(t) \tag{2}$$

Then replace r in (2) with $2(n+1)$ we obtain (and setting $t=0$)

$$\begin{aligned}
M_X^{(2(n+1))}(0) &= E\left(x^{2(n+1)}\right)\\
&= (2(n+1)-1)\,\sigma^2 M_X^{(2(n+1)-2)}(0)\\
&= (2n+2-1)\,\sigma^2 M_X^{(2n+2-2)}(0)\\
&= (2n+1)\,\sigma^2 M_X^{(2n)}(0)
\end{aligned}$$

But $M_X^{(2n)}(0)$ in the above is just $E(x^{2n})$ which we *assumed* in (1) to be $\frac{(2n)!\ \sigma^{2n}}{2^n(n!)}$, hence the above can be written as

$$E\left(x^{2(n+1)}\right) = (2n+1)\,\sigma^2\frac{(2n)!\ \sigma^{2n}}{2^n\,(n!)}$$

But $\sigma^2\sigma^{2n} = \sigma^{2(n+1)}$ and $(2n+1)\,(2n)! = (2n+1)!$ hence the above becomes

$$E\left(x^{2(n+1)}\right) = \frac{(2n+1)!\ \sigma^{2(n+1)}}{2^n\,(n!)} \tag{3}$$

But $(2n+1)! = \frac{(2n+2)!}{(2n+2)}$

Hence (3) becomes

$$\begin{aligned}
E\left(x^{2(n+1)}\right) &= \frac{(2n+2)!\ \sigma^{2(n+1)}}{(2n+2)\,2^n\,(n!)}\\
&= \frac{(2(n+1))!\ \sigma^{2(n+1)}}{2n\times 2^n\,(n!) + 2\times 2^n\,(n!)}\\
&= \frac{(2(n+1))!\ \sigma^{2(n+1)}}{2^{n+1}n\,(n!) + 2^{n+1}\,(n!)}\\
&= \frac{(2(n+1))!\ \sigma^{2(n+1)}}{2^{n+1}\,(n!\,(n+1))}
\end{aligned}$$

But $n!\,(n+1) = (n+1)!$ hence the above becomes

$$E\left(x^{2(n+1)}\right) = \frac{(2(n+1))!\,\sigma^{2(n+1)}}{2^{n+1}\,(n+1)!}$$

Compare to (*) we see it is the same. QED

3.5 Key solutions for HWs

Local contents

3.5.1 key solutions for chapter 1

(1)

Chapter 1 Home work :

Problem 2 :

a and b)

Possible outcomes of the 2nd die

Possible outcomes on the first die

	1	2	3	4	5	6
1				A	A	A
2	B		A	A	A	A
3	B	AB	(A)	A	A	A
4	ABC	ABC	ABC	AC	AC	AC
5	AB	AB	AB	AB	A	A
6	AB	AB	AB	AB	AB	A

Cell (i,j) denotes the event of rolling $i=1,\dots,6$ on the first die and $j=1,\dots,6$ on the 2nd die

C) (1) $A \cap C = C$

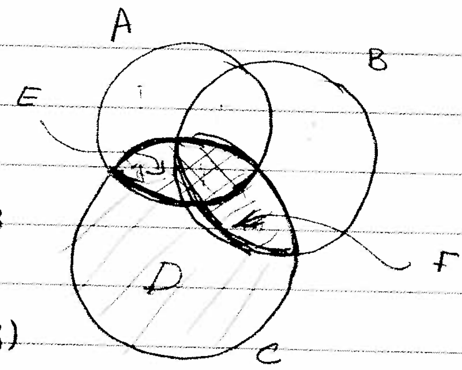
(2) $B \cup C$ = all the cells that either have B or C on them. for example $(4,2)$ or $(3,1)$ or $(4,6)$

(3) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$= (A \cap B) \cup C$

Thus, this would consist of all cells that either include both A and B or include C.

(2)

<u>Problem 6</u> :



$P(A \cup B \cup C) = P(A \cup B)$

$+ P(D)$ Axiom 3

$= P(A) + P(B) - P(A \cap B)$

$+ P(D)$ Addition rule $(*)$

$P(C) = P(D) + P(E \cup F)$ $E = A \cap C$

$P(E \cup F) = P(E) + P(F) - P(E \cap F)$ $F = B \cap C$

$\Rightarrow P(D) = P(C) - P(E \cup F)$

$= P(C) - P(E) - P(F) + P(E \cap F)$

$= P(C) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$ $(**)$

$*$ and $**$ give the desired result

3

Problem 7

Want to show

$$P(A\cap B) \geq P(A)+P(B)-1$$

$$P(A\cap B) = P(A)+P(B)-P(A\cup B)$$
$$\geq P(A)+P(B)-P(\Omega) \quad \text{since } \Omega \supset A\cup B$$
$$P(\Omega)\geq P(A\cup B)$$
$$= P(A)+P(B)-1 \quad \text{Axiom 1}$$

Problem 9: A = rains on Saturday

B = rains on Sunday

$A\cup B$ = rains on the weekend

$$P(A\cup B) = P(A)+P(B)-P(A\cap B)$$
$$= 0.25 + 0.25 - P(A\cap B)$$
$$\leq 0.50$$

No, it is less than or equal to 0.50, because $P(A\cap B) \geq 0$

4

Problem 11

452 $\square\,\square\,\square\,\square$

Total number of possible phone numbers $= 10^4$

Total number of distinct Phone numbers starting with 452

$$\boxed{7} \times \boxed{6} \times \boxed{5} \times \boxed{4} = 840$$

$$P = \frac{840}{10^4}$$

#16

$s \to 2$

$t \to 3$

$a \to 2$

$i \to 2$

$c \to 1$

$l \to 2$

$y = 1$

$$\frac{13!}{2!\,3!\,2!\,2!\,2!}$$

$$\underbrace{\sqcup \ \sqcup \ \sqcup \ \ldots \ \sqcup}_{13}$$

selecting a place for two S's

like multinomials

$$\underbrace{\binom{13}{2}}_{\substack{\text{group 1}\\ s}} \quad \underbrace{\binom{11}{3}}_{\substack{\text{group 2}\\ t}} \quad \underbrace{}_{\substack{\text{group 3}\\ a}} \quad \cdots \quad \underbrace{}_{\substack{\text{group 7}\\ l}}$$

5

Problem 18

n items k - defective

m selected and inspected :

E = at least one defective

Want to determine m such that

$P(E) = 0.90$

$P(E) = 1 - P(E^c)$

$= 1 - P(\text{No defective})$

$$= 1 - \frac{\binom{k}{0}\binom{n-k}{m}}{\binom{n}{m}} = 1 - \frac{(n-k)!\,(n-m)!}{n!\,(n-k-m)!}$$

m should be chosen so that

$$0.1 = \frac{(n-m)(n-m-1)\cdots(n-m-k+1)}{n\,(n-1)\cdots(n-k+1)}$$

$$= \prod_{i=0}^{k-1} \frac{n-m-i}{n-i}$$

~~$n = 1000$, $k = 10$ $\Rightarrow m = 100$~~

MATLAB Code:

```
n=1000;  k=10;      m = 205 (by trial)
prod(n-m-[0:k-1]./(n-[0:k-1]))
```

6

#42

$$\frac{11!}{4!\cdot 3!\cdot 3!} = 46,200$$

#35b

$$\binom{n-1}{r-1}+\binom{n-1}{r} = \frac{(n-1)!}{(r-1)!(n-r)!} + \frac{(n-1)!}{r!(n-r-1)!}$$

$$= \frac{r(n-1)! + (n-1)!(n-r)}{r!(n-r)!}$$

$$= \frac{(n-1)!(r+n-r)}{r!(n-r)!} = \frac{n!}{r!(n-r)!} = \binom{n}{r}$$

Explanation

$$\underbrace{\otimes}_{1}\ \underbrace{x\ \ x\ \ x\ \ ---\ \ x}_{n-1}$$

Choosing r objects from the n objects above can be accomplished in one of two ways

1) Choose $\otimes$, and r-1 from the remaining n-1 $\Rightarrow \binom{n-1}{r-1}$

or

2) Do not choose $\otimes$ and choose r from the remaining n-1.

7

#45 By induction:

true for $n=1$

Suppose it is true for $n=k-1$, and let $B=A_1\cap\cdots\cap A_{k-1}$

then

$$P(A_1\cap A_2\cdots\cap A_k)=P(B\cap A_k)$$
$$=P(A_k|B)\cdot P(B)$$

Now apply the induction hypothesis and the proof is complete.

#53 F = filing H = High risk M = medium risk L = low risk

$P(H)=0.1$, $P(M)=.20$ $P(L)=0.70$

$P(F|H)=.02$, $P(F|M)=.01$ $P(F|L)=.0025$

$$P(H|F)=\frac{P(F|H)\cdot P(H)}{P(F|H)\cdot P(H)+P(F|M)\cdot P(M)+P(F|L)\cdot P(L)}$$

$$=\frac{(.02)(.1)}{(.02)(.1)+(.2)(.01)+(.7)(.0025)}=.3478$$

#54

R_i = rains day i

$P(R_i|R_{i-1})=\alpha$ $P(R_i^c|R_{i-1}^c)=\beta$

$P(R_i|R_{i-1}\cap\cdots\cap R_0)=P(R_i|R_{i-1})$

a) $P(R_1)=P(R_1|R_0)\cdot P(R_0)+P(R_1|R_0^c)\cdot P(R_0^c)=\alpha p+(1-\beta)(1-p)$

$$=p(\alpha+\beta-1)+1-\beta$$

8

54a

$$P(R_1) = P(R_1|R_0)\cdot P(R_0) + P(R_1|R_0^c)\,P(R_0^c)$$

$$= \alpha p + (1-\beta)(1-p)$$

54b.

$$P(R_2) = P(R_2|R_1)\cdot P(R_1) + P(R_2|R_1^c)\,P(R_1^c)$$

$$= \alpha\, P(R_1) + (1-\beta)\,(1-P(R_1))$$

$$= P(R_1)\,[\alpha+\beta-1] + (1-\beta)$$

$$= [\alpha p + (1-\beta)(1-p)][\alpha+\beta-1] + (1-\beta)$$

$$= \alpha p(\alpha+\beta-1) + (1-\beta)[(1-p)(\alpha+\beta-1)+1]$$

9

#54C

$$P(R_n) = P(R_n|R_{n-1})\cdot P(R_{n-1}) + P(R_n|R_{n-1}^c)\cdot P(R_{n-1}^c)$$

$$= \alpha P(R_{n-1}) + (1-\beta)(1-P(R_{n-1}))$$

$$= P(R_{n-1})\left[\alpha+\beta-1\right] + 1-\beta$$

let $\gamma_n = P(R_n)$

Solve

$$\gamma_n = (\alpha+\beta-1)\gamma_{n-1} + 1-\beta \quad \text{with } \gamma_0 = p$$

$$\gamma_n = (\alpha+\beta-1)\gamma_{n-1} + (1-\beta)$$

$$= (\alpha+\beta-1)\left[(\alpha+\beta-1)\gamma_{n-2} + (1-\beta)\right] + 1-\beta$$

$$= (\alpha+\beta-1)^2\gamma_{n-2} + (1-\beta)(\alpha+\beta-1) + (1-\beta)$$

$$= (\alpha+\beta-1)^2\left[(\alpha+\beta-1)\gamma_{n-3} + (1-\beta)\right] + (1-\beta)(\alpha+\beta-1) + (1-\beta)$$

$$= (\alpha+\beta-1)^3\gamma_{n-3} + (1-\beta)(\alpha+\beta-1)^2 + (1-\beta)(\alpha+\beta-1) + (1-\beta)$$

$$= \cdots$$

$$= (\alpha+\beta-1)^n\gamma_0 + (1-\beta)\left[(\alpha+\beta-1)^{n-1} + (\alpha+\beta-1)^{n-2} + \cdots + 1\right]$$

10

54C Continued.

$$\gamma_n = (\alpha+\beta-1)^n p + (1-\beta)\frac{1-(\alpha+\beta-1)^n}{1-\alpha-\beta+1}$$

as $n \to \infty$

$$\gamma_n \to \frac{1-\beta}{2-\alpha-\beta}$$

11

#64

Axiom 1: $Q(\Omega) = P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$ ✓

Axiom 2: Suppose $A \subset \Omega$, then

$Q(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0$

Axiom 3: If A_1 and A_2 are two disjoint events

$$
\begin{aligned}
Q(A_1 \cup A_2) &= P(A_1 \cup A_2 | B) \\
&= \frac{P[(A_1 \cup A_2) \cap B]}{P(B)} \\
&= \frac{P[(A_1 \cap B) \cup (A_2 \cap B)]}{P(B)} \\
&= \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} \quad \text{By Axiom 3} \\
&= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} \\
&= P(A_1|B) + P(A_2|B) \\
&= Q(A_1) + Q(A_2)
\end{aligned}
$$

#79 (12)

aa ← diseased
Aa ← Carrier but not diseased
AA ← not Carrier and is not diseased

a) Aa ⟷ Aa F_A = father gives A
M_a = mother " a etc.

$P(AA) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$

$P(Aa) = P(F_A)P(M_a) + P(F_a)P(M_A) = \frac{1}{4}+\frac{1}{4} = \frac{1}{2}$

$P(aa) = P(F_a)P(M_a) = \frac{1}{4}$

	Father A	Father a
Mother A	AA	Aa
Mother a	Aa	aa

b) Male offspring is Aa ∪ AA
Carrier Aa

$$P(Aa \mid Aa \cup AA) = \frac{P[Aa \cap (Aa \cup AA)]}{P(Aa \cup AA)} = \frac{P(Aa)}{P(Aa \cup AA)} = \frac{1/2}{\frac{1}{4}+\frac{1}{2}} = 2/3$$

	A	a
A	AA	Aa
a	Aa	aa

c)

	offspring of Part b: Aa($\frac{2}{3}$)	AA($\frac{1}{3}$)
Aa(p)	E	F
AA(1-p)	G	H

$$P(AA) = P(AA|E)\cdot P(E) + P(AA|F)\cdot P(F) + P(AA|G)P(G) + P(AA|H)P(H)$$
$$= \left(\tfrac{1}{4}\right)\left(\tfrac{2}{3}\right)p + \left(\tfrac{1}{2}\right)\left(\tfrac{1}{3}\right)(p) + \left(\tfrac{1}{2}\right)\left(\tfrac{2}{3}\right)(1-p) + 1\left(\tfrac{1}{3}\right)(1-p)$$
$$= \frac{2}{3} - \frac{p}{3}$$

13

#79 c (cont)

$$P(Aa) = \left(\tfrac{2}{3}\right)(p)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)p\left(\tfrac{1}{2}\right) + \left(\tfrac{2}{3}\right)(1-p)\left(\tfrac{1}{2}\right) + \left(\tfrac{1}{3}\right)(1-p)(0)$$

$$= \frac{p}{3} + \frac{p}{6} + \frac{1-p}{3} = \frac{p}{6} + \frac{1}{3}$$

$$P(aa) = \tfrac{2}{3}p\left(\tfrac{1}{4}\right) = \frac{p}{6}$$

d)

$$P(F_{Aa} \mid AA \cup Aa) = \frac{P(AA \cup Aa \mid F_{Aa})\, P(F_{Aa})}{P(AA \cup Aa)}$$

$$= \frac{\left[P(AA \mid F_{Aa}) + P(Aa \mid F_{Aa})\right] P(F_{Aa})}{P(AA) + P(Aa)}$$

$$P(AA \mid F_{Aa}) = P(AA \mid F_{Aa} \cap M_{AA})\, P(M_{AA}) + P(AA \mid F_{Aa} \cap M_{Aa})\, P(M_{Aa}) + P(AA \mid F_{Aa} \cap M_{aa})\, P(M_{aa})$$

$$= \tfrac{1}{2}(1-p) + \tfrac{1}{4}(p) + 0$$

$$= \frac{1}{2} - \frac{1}{4}p$$

	father: $Aa\left(\tfrac{2}{3}\right)$	$AA\left(\tfrac{1}{3}\right)$
Mother: $Aa\ (p)$		
$AA\ (1-p)$		

$$P(Aa \mid F_{Aa}) = \tfrac{1}{2}(1-p) + \tfrac{1}{2}(p) + \tfrac{1}{2}(0)$$

$$= \frac{1}{2}$$

$$P(F_{Aa}) = \frac{2}{3} \qquad P(AA) + P(Aa) = \frac{2}{3} - \frac{p}{3} + \frac{p}{6} + \frac{1}{3} = 1 - \frac{p}{6}$$

14

#79d (Continued)

$$P(F_{Aa} \mid AA \cup Aa) = \frac{(1-\frac{1}{4}p)(\frac{2}{3})}{1-p/6}$$

$$\frac{P(AA \cup Aa \mid F_{Aa}) \cdot p(F_{Aa})}{p(AA \cup Aa)}$$

3.5.2 key solutions for chapter 2

45

Chapter 2

Problem 6.

$I_{A\cap B} = 1 \iff \omega \in A \cap B \iff \omega \in A$ and $\omega \in B$

$\iff I_A = 1$ and $I_B = 1$

$\iff I_A \cdot I_B = 1 \qquad (1)$

$I_{A\cap B} = 0 \iff \omega \notin A \cap B \iff \omega \in (A\cap B)^c$

$\iff \omega \in A^c$ or $\omega \in B^c$

$\iff I_A = 0$ or $I_B = 0$

$\iff I_A I_B = 0 \qquad (2)$

(1) and (2) $\Rightarrow I_{A\cap B} = I_A I_B$. ▨

$I_A I_B = 1 \iff I_A = 1$ and $I_B = 1 \iff \min(I_A, I_B) = 1 \quad (3)$

$I_A I_B = 0 \iff I_A = 0$ or $I_B = 0 \iff \min(I_A, I_B) = 0 \quad (4)$

(3) and (4) $\Rightarrow I_A I_B = \min(I_A, I_B)$ ▨

$I_{A\cup B} = 1 \iff \omega \in A$ or $\omega \in B$

$\iff I_A = 1$ or $I_B = 1$

$\iff \max(I_A, I_B) = 1 \qquad (5)$

$I_{A\cup B} = 0 \iff \omega \notin A\cup B \iff \omega \in A^c$ and $\omega \in B^c$

$\iff I_A = 0$ and $I_B = 0$

$\iff \max(I_A, I_B) = 0 \qquad (6)$

(5) and (6) $\Rightarrow I_{A\cup B} = \max(I_A, I_B)$.

16

#11

$$\frac{P(X=k)}{P(X=k-1)} = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}}$$

$$= \frac{n!}{k!\,(n-k)!} \cdot \frac{(k-1)!\,(n-k+1)!}{n!} \cdot \frac{p}{1-p}$$

$$= \left(\frac{n-k+1}{k}\right)\left(\frac{p}{1-p}\right)$$

$$P(X=k) \geq P(X=k-1) \iff \left(\frac{n-k+1}{k}\right)\left(\frac{p}{1-p}\right) \geq 1$$

$$\iff np - kp + p \geq k(1-p)$$

$$\iff np + p \geq k$$

The mode of the binomial is the greatest integer smaller than or equal to $np+p$.

17

#19

$$F(x) = P(X \le x)$$
$$= \sum_{i=1}^{x} (1-p)^{i-1} p = p \cdot \frac{1-(1-p)^i}{1-(1-p)} = 1-(1-p)^i$$

#21

$$P(X > n+k-1 \mid X > n-1) = \frac{P(X > n+k-1 \ \& \ X > n-1)}{P(X > n-1)}$$
$$= \frac{P(X > n+k-1)}{P(X > n-1)}$$
$$= \frac{1 - P(X \le n+k-1)}{1 - P(X \le n-1)} = \frac{1-1+(1-p)^{n+k-1}}{1-1+(1-p)^{n-1}}$$
$$= (1-p)^k$$
$$= P(X > k)$$

$$\underbrace{F\,F\,F \cdots F}_{n-1}\ \underbrace{F\,F \cdots F}_{k}$$

If we know that success has not occured up to and including the $n-1^{st}$ trial, then the probability that success will occure after $n+k-1^{st}$ trial is equivallent to success occuring after k trials since events are independent (see diagram above)

(18)

#27

prob of incidence $= \frac{1}{1000}$

$n = 100{,}000$ $\quad k = 0, 1, 2$

large n, small p $\Rightarrow$ Poisson approximation with $\lambda = np = 100{,}000 \cdot \left(\frac{1}{1000}\right) = 100$

$$P(X=k) \simeq \frac{e^{-100}\, 100^k}{k!} \qquad k = 0, 1, 2$$

#28

$$p_k = P(X=k) = \binom{n}{k} p^k q^{n-k}$$

$$= \frac{n!}{k!\,(n-k)!}\, p^k q^{n-k}$$

$$= \frac{n!\,(n-k+1)}{k\,(k-1)!\,(n-k+1)!}\, p^{k-1} \cdot p\; q^{n-k+1}\; q^{-1}$$

$$= \binom{n}{k-1} p^{k-1} q^{n-k+1}\, \frac{(n-k+1)p}{kq}$$

$$= P(X=k-1)\, \frac{(n-k+1)p}{kq}$$

$$= p_{k-1}\, \frac{(n-k+1)p}{kq}$$

and obviously $p_0 = P(X=0) = q^n$.

19

#28 (Continued)

$p_0 = (1-.0005)^{9000} = .0111$

$p_1 = \frac{(9000-1+1)(.0005)}{.9995}(.0111) = .04996$ etc.

Note that it would be difficult to compute these values using the direct formula.

#29

$p_0 = e^{-\lambda}$ obviously

$p_k = e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!}\cdot\frac{\lambda}{k} = \frac{\lambda}{k}p_{k-1}$

$p_0 = e^{-4.5}$ $p_1 = 4.5e^{-4.5}$ $p_2 = \frac{(4.5)^2}{2}e^{-4.5}$ etc.

#31 $\lambda = 2$ per hour

a. X = # of phone calls in a 10-min period

$X \sim P(\lambda = \frac{1}{3})$

$P(\text{phone rings}) = P(X \geq 1) = 1 - P(X=0)$

$= 1 - e^{-1/3} \cong 28\%$

20

#31 b

X = # of Phone calls in t-hour Period

$X \sim \mathcal{P}(2t)$

$P(X=0) = 0.5 \Rightarrow e^{-2t} = 0.5 \Rightarrow -2t = \ln(.5)$

$t = \frac{\ln(.5)}{-2} = 0.347$ hour $\cong 20.8$ min

21

#34

$$\int_{-1}^{1} (1+\alpha x)/2 \, dx = \frac{1}{2}\left(x+\alpha\frac{x^2}{2}\right)\Big|_{-1}^{1}$$
$$= \frac{1}{2}\left[1+\frac{\alpha}{2}+1-\frac{\alpha}{2}\right] = 1$$

$$F(x) = \int_{-1}^{x} \frac{1+\alpha t}{2}\, dt = \frac{1}{2}\left(t+\alpha\frac{t^2}{2}\right)\Big|_{-1}^{x}$$
$$= \frac{1}{2}\left[x+\alpha\frac{x^2}{2}+1-\frac{\alpha}{2}\right] \quad -1\le x\le 1$$

$$F(x) = \begin{cases} 0 & x\le -1 \\ \frac{1}{2}x+\frac{1}{2}+\frac{\alpha}{4}(x^2-1) & -1\le x\le 1 \\ 1 & x\ge 1 \end{cases}$$

find the inverse function

$$F(F^{-1}(x)) = x$$

$$F(F^{-1}(x)) = \frac{\alpha}{4}\left[F^{-1}(x)\right]^2 + \frac{1}{2}F^{-1}(x) + \frac{1}{2} - \frac{\alpha}{4} = x$$

$$F^{-1}(x) = \frac{-\frac{1}{2} \pm \sqrt{\frac{1}{4}-\alpha\left(\frac{1}{2}-\frac{\alpha}{4}-x\right)}}{\alpha/2}$$

$$= \frac{-1 \pm 2\sqrt{\frac{1}{4}-\alpha\left(\frac{1}{2}-\frac{\alpha}{4}-x\right)}}{\alpha}$$

$$x_p = F^{-1}(p) = \frac{-1+2\sqrt{1/4-\alpha\left(\frac{1}{2}-\frac{\alpha}{4}-p\right)}}{\alpha}$$

Note the root with - is not considered as that (?) F^{-1} will be out of range

Median $= F^{-1}(\frac{1}{2})$ $\quad Q_1 = F^{-1}(\frac{1}{4})$ $\quad Q_3 = F^{-1}(\frac{3}{4})$

22

#39

a. $\lim_{x \to \infty} F(x) = \lim_{x \to \infty} \left(\frac{1}{2} + \frac{1}{\pi} \tan^{-1}(x) \right)$

$= \frac{1}{2} + \frac{1}{\pi} \lim_{x \to \infty} \tan^{-1} x$

$= \frac{1}{2} + \frac{1}{\pi} \left(\frac{\pi}{2}\right) = 1$

$\lim_{x \to -\infty} F(x) = \frac{1}{2} + \frac{1}{\pi} \lim_{x \to -\infty} \tan^{-1}(x)$

$= \frac{1}{2} + \frac{1}{\pi} \left(\frac{-\pi}{2}\right) = 0$

$\frac{dF(x)}{dx} = \frac{1}{\pi(1+x^2)} > 0 \Rightarrow F(x)$ is non decreasing.

b. $f(x) = \frac{dF(x)}{dx} = \frac{1}{\pi(1+x^2)} \qquad -\infty < x < \infty$

c. $P(X > x) = 0.1 \iff P(X \le x) = 0.9$

$\Rightarrow F(x) = 0.9 \Rightarrow \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(x) = 0.9$

$\Rightarrow \tan^{-1}(x) = 0.4\pi \qquad x = \tan(.4\pi) = 3.08$

23

44

$$P(X=k) = P(T \le k+1) - P(T \le k)$$
$$= 1 - e^{-\lambda(k+1)} - 1 + e^{-\lambda k}$$
$$= e^{-\lambda k}(1 - e^{-\lambda})$$

Geometric dist^{n} with parameter $e^{-\lambda}$

49

a. $\Gamma(1) = \int_0^\infty u^0 e^{-u}\,du = \int_0^\infty e^{-u}\,du = 1$

b. $\Gamma(x+1) = \int_0^\infty u^x e^{-u}\,du$ $\qquad U = u^x \quad dV = e^{-u}\,du$

$\qquad dU = xU^{x-1} \quad V = -e^{-u}$

$$= -e^{-u} u^x \Big|_{u=0}^{\infty} + x\int_0^\infty e^{-u} u^{x-1}\,du$$

$$= x\,\Gamma(x-1)$$

c. Want to show $\Gamma(n) = (n-1)!$

by part (a), $\Gamma(1) = 1$

Suppose true for $n=k$ i.e. $\Gamma(k) = (k-1)!$

Then

$$\Gamma(k+1) \underset{\substack{\uparrow \\ \text{part } b}}{=} k\,\Gamma(k) = k(k-1)! = k! \quad \blacksquare$$

24

#49d

Statement is true for n=1 clearly.

Suppose the Statement is true for $n=2k-1$

i.e. (induction hypothesis):

$$\Gamma\left(\frac{2k-1}{2}\right) = \frac{\sqrt{\pi}\,(2k-2)!}{2^{(2k-2)!}\left(\frac{2k-2}{2}\right)!}.$$

Want to show

$$\Gamma\left(\frac{2k+1}{2}\right) = \frac{\sqrt{\pi}\,(2k)!}{2^{2k}\,(k!)}$$

Proof:

$$\Gamma\left(\frac{2k+1}{2}\right) = \Gamma\left(\frac{2k-1}{2}+1\right) = \frac{2k-1}{2}\,\Gamma\left(\frac{2k-1}{2}\right)$$

$$=\left(\frac{2k-1}{2}\right)\frac{\sqrt{\pi}\,(2k-2)!}{2^{(2k-2)}\left(\frac{2k-2}{2}\right)!} \quad \text{by hyp.}$$

$$= \frac{\sqrt{\pi}\,(2k-1)!}{2^{2k-1}\,(k-1)!}\cdot\frac{k}{k}$$

$$= \frac{\sqrt{\pi}\,k\,(2k-1)!}{k!\,2^{2k-1}}\cdot\frac{2}{2} = \frac{\sqrt{\pi}\,(2k)!}{k!\,2^{2k}}$$

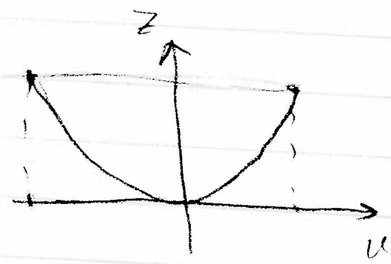
Q.E.D.

26

#54

$$F_U(u) = \begin{cases} \frac{1}{2}(u+1) & -1 \le u \le 1 \\ 0 & u \le -1 \\ 1 & u \ge 1 \end{cases}$$

Let $Z = U^2$



$$\begin{aligned} F_Z(z) &= P(Z \le z) \\ &= P(U^2 \le z) \\ &= P(-\sqrt{z} \le U \le \sqrt{z}) \\ &= P(U \le \sqrt{z}) - P(U \le -\sqrt{z}) \\ &= \frac{1}{2}(\sqrt{z}+1) - \frac{1}{2}(-\sqrt{z}+1) \\ &= \sqrt{z} \end{aligned}$$

$f_Z(z) = \frac{1}{2} z^{-1/2} \qquad 0 \le z \le 1$

27

#60

$$F_Y(y) = P(Y \le y)$$
$$= P(e^Z \le y)$$
$$= P(Z \le \log y)$$
$$= \Phi(\log y)$$

$$f_Y(y) = \frac{1}{y}\phi(\log y) \qquad 0 \le y \le \infty$$
$$= \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{1}{2}(\log y - \mu)^2/\sigma^2}$$

#62. Let $g(x) = ax+b$, then

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

$$g(g^{-1}(x)) = x \Rightarrow a g^{-1}(x) + b = x$$
$$\Rightarrow g^{-1}(x) = \frac{x-b}{a}$$

$$\frac{d}{dy} g^{-1}(x) = \frac{1}{a}$$

$\therefore$

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{x-b}{a}\right)$$

28

\# 66

$$F(x) = \int_1^x \alpha t^{-\alpha-1}\,dt$$

$$= \alpha \, t^{-\alpha}/-\alpha \Big|_1^x = -t^{-\alpha}\Big|_1^x = 1 - x^{-\alpha}$$

$$F(F^{-1}(x)) = 1 - [F^{-1}(x)]^{-\alpha} = x$$

$$[F^{-1}(x)]^{-\alpha} = 1 - x \quad \Rightarrow F^{-1}(x) = (1-x)^{-1/\alpha}$$

Use $(1-U)^{-1/\alpha}$ $U \sim$ unif(0,1)

\#67

a. $$f(x) = e^{-(x/\alpha)^{\beta}} \cdot -(-\beta)(x/\alpha)^{\beta-1}\left(\frac{1}{\alpha}\right)$$

$$= \beta/\alpha \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^{\beta}}$$

b. $$P(X \le x) = P((W/\alpha)^{\beta} \le x)$$

$$= P\left(\beta \log\left(\frac{W}{\alpha}\right) \le \log(x)\right)$$

$$= P\left(\log\left(\frac{W}{\alpha}\right) \le \frac{1}{\beta} \log x\right)$$

$$= P\left(\frac{W}{\alpha} \le (e^{\log x})^{1/\beta}\right)$$

$$= P(W \le \alpha x^{1/\beta})$$

$$= 1 - e^{-\left[\frac{\alpha x^{1/\beta}}{\alpha}\right]^{\beta}}$$

$$= 1 - e^{-x}$$

CDF of exponential with $\lambda = 1$

* c. $F^{-1}(U) = [-\log(1-U)]^{1/\beta}$

Problem 67

a) $F(x) = 1 - e^{-(x/\alpha)^\beta}$

$f(x) = \frac{d}{dx} F(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta}$

$= \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(x/\alpha)^\beta}$

b)
$F_X(x) = P(X \le x)$

$= P[(W/\alpha)^\beta \le x]$

$= P[\frac{W}{\alpha} \le x^{1/\beta}]$

$= P[W \le \alpha x^{1/\beta}]$

$= F_W(\alpha x^{1/\beta})$

$= 1 - e^{-[\frac{\alpha x^{1/\beta}}{\alpha}]^\beta}$

$= 1 - e^{-x}$ which is cdf of exponential.

c) Generate $U \sim unif(0,1)$ is generated the value

$F^{-1}(U) = -\alpha [\log(1-U)]^{1/\beta}$

is computed

#70

$$P(V \le v) = P(U^{-\alpha} \le v)$$

$$= P\left(\frac{1}{U^{\alpha}} \le v\right)$$

$$= P\left(U^{\alpha} \ge \frac{1}{v}\right)$$

$$= 1 - P(U \le v^{-1/\alpha})$$

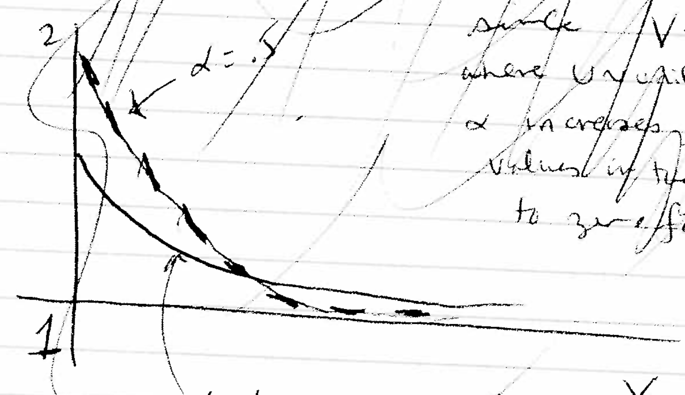
$$= 1 - F_U(v^{-1/\alpha})$$

$$= 1 - v^{-1/\alpha} \qquad v \ge 1$$

$$f_V(v) = \frac{\partial}{\partial v}(1 - v^{-1/\alpha}) = \frac{1}{\alpha} v^{-(1+\frac{1}{\alpha})} \qquad v \ge 1$$

~~As α increases, the tail goes to zero ~~faster~~ slower~~

~~This makes sense since $V = U^{-\alpha}$ where $U \sim \text{Unif}(0,1)$. As α increases, the values in the tails go to zero faster.~~



~~It makes sense since the graph of $f(x) = \frac{1}{x^{\alpha}}$ goes to zero slower for larger values of α when $x \ge 1$.~~

70 Continued

As α increases, the tails go to zero slower. Does this make sense?

Note

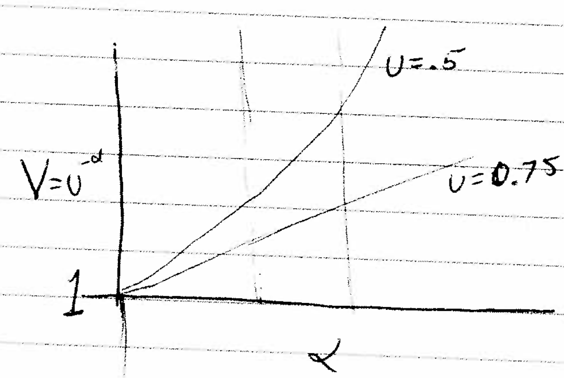
$$V = U^{-\alpha} \qquad \alpha > 0 \qquad 0 \le U \le 1$$

As a function of α, how does V change?

$$\frac{dV}{d\alpha} = -\log(U)\, U^{-\alpha} \geq 0 \quad \text{since } U \in [0,1]$$

$\therefore$ $V = U^{-\alpha}$ is an increasing function of α.

that is for larger values of α, values of $U^{-\alpha}$ are larger, and thus the tail of the density goes to zero slower for larger values of α



29

#71 $U \sim \mathcal{U}(0,1)$

Let $X \sim F$

$Y = k \quad \text{if } F(k-1) \le U \le F(k)$

$$P(Y=k) = P(F(k-1) \le U \le F(k))$$

$$= P(U \le F(k)) - P(U \le F(k-1))$$

$$= F(k) - F(k-1)$$

$$= P(X \le k) - P(X \le k-1)$$

$$= P(X = k)$$

$Y \sim X$ $\boxtimes$

For the Geometric r.v.

$$F(x) = \sum_{k=1}^{x} (1-p)^{k-1} p$$

$$= p \sum_{k=0}^{x-1} (1-p)^{n} = p \cdot \frac{1-(1-p)^x}{1-(1-p)} = 1-(1-p)^x$$

Generate a $u = U(0,1)$

$u = F(k) \Leftrightarrow u = 1-(1-p)^k \Leftrightarrow (1-p)^k = 1-u$

$\Leftrightarrow k \log(1-p) = \log(1-u) \Leftrightarrow k = \frac{\log(1-u)}{\log(1-p)}$.

$u = F(k-1) \Leftrightarrow u = 1-(1-p)^{k-1} \Leftrightarrow k = \frac{\log(1-u)}{\log(1-p)} + 1$

Choose an integer k in the interval $\left[\frac{\log(1-u)}{\log(1-p)}, \frac{\log(1-u)}{\log(1-p)} + 1\right]$

3.5.3 key solutions for chapter 3

6

#68

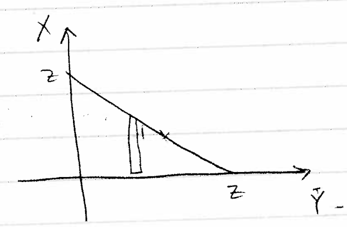
Let T_i be the amount of time it takes cashier i to finish. The time that the next person in line gets served is

$$\min(T_i) \sim \exp(n\lambda)$$ ← see problem 65

So the 2nd person in line will have to wait the sum of two independent exponential dist^ns with parameter $(n\lambda)$

So let $X \sim \exp(n\lambda)$ $Y \sim \exp(n\lambda)$

then $Z = X + Y$



$$f_Z(z) = \int_0^z (n\lambda)^2 e^{-n\lambda x} e^{-n\lambda(z-x)}\, dx$$

$$= (n\lambda)^2 z e^{-\lambda z} \sim \Gamma(2, n\lambda)$$

60

#70 Continued

Let $Y = X_{(1)}$ and $Z = X_{(5)}$

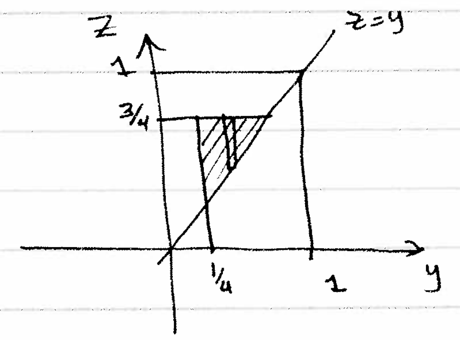
We first find the joint density of Y, Z.

Refering to the notes

$$f_{YZ}(y,z) = \binom{5}{0,1,3,1,0} (z-y)^3$$

$$= 20(z-y)^3 \qquad 1 > z > y > 0$$

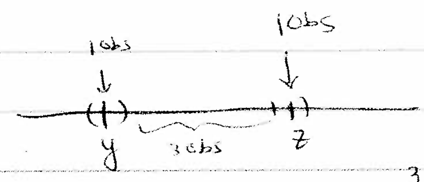
$$P(Z \le 3/4, Y \ge \tfrac{1}{4})$$



$$= \int_{1/4}^{3/4} \int_{y}^{3/4} 20(z-y)^3 \, dz \, dy$$

$$= \frac{1}{32}$$

Note:



$$f_{YZ}(y,z) \simeq f(y)dy \, f(z)dz \, [F(z) - F(y)]^3 \binom{5}{1,3,1}$$

For the Uniform, we get

$$f_{YZ}(y,z) = \frac{5!}{3!}(z-y)^3$$

59

§3.7

#65 Let $T_1, \ldots, T_n$ be lifetime of components $1, 2, \ldots, n$. $T_i \sim exp(\lambda_i)$

Since components are connected in series, the system will fail if any one of the components fail. Then

T = The lifetime of the system = $\min\limits_{i=1,\ldots,n} T_i$

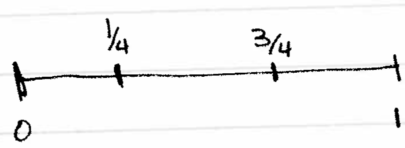
$$P(T \leq t) = P(\min_{i=1,\ldots,n} T_i \leq t) = 1 - P(\min_{i=1,\ldots,n} T_i \geq t)$$

$$= 1 - P(T_1 \geq t, T_2 \geq t, \ldots, T_n \geq t)$$

$$= 1 - P(T_1 \geq t) \ldots P(T_n \geq t)$$

$$= 1 - e^{\lambda_1 t} e^{\lambda_2 t} \cdots e^{\lambda_n t}$$

$$= 1 - e^{(\sum_{i=1}^{n} \lambda_i) t} \Rightarrow T \sim exp(\sum_{i=1}^{n} \lambda_i)$$

#70 

The question to find

$$P(X_{(5)} \leq 3/4, X_{(1)} \geq \tfrac{1}{4})$$

58

#64

$U = X + Y \qquad X \perp\!\!\!\perp Y \qquad X \sim \exp(\lambda)$

$V = X/Y \qquad\qquad\qquad\qquad Y \sim \exp(\lambda)$

$$f_{XY}(x,y) = \begin{cases} \lambda^2 e^{-\lambda(x+y)} & x>0 \quad y>0 \\ 0 & \text{o.w.} \end{cases}$$

$$J(x,y) = \begin{vmatrix} 1 & 1 \\ \frac{1}{y} & \frac{-x}{y^2} \end{vmatrix} = -\frac{x}{y^2} - \frac{1}{y} = \frac{-x-y}{y^2}$$

$$|J(x,y)| = \frac{|x+y|}{y^2} = \frac{x+y}{y^2}$$

$$X = \frac{UV}{1+V} \qquad Y = \frac{U}{1+V}$$

$$f_{UV}(u,v) = \lambda^2 e^{-\lambda\left[\frac{uv}{1+v} + \frac{u}{1+v}\right]} \cdot \frac{[u/(1+v)]^2}{\frac{uv+u}{1+v}}$$

$$= \lambda^2 e^{-\lambda u} \cdot \frac{u}{(1+v)^2} \qquad u \geq 0 \quad v \geq 0$$

Note that the density can be factored into two functions $g(u)$ and $h(v)$. Therefore U and V are independent.

57

§3.6 #58 Note: a_1 and a_2 are required to be nonzero.

Let $\phi_{X_1 X_2}$ denote the density of X_1, X_2.

The Jacobian of the transformation is given by

$$J(x_1, x_2) = \begin{vmatrix} a_1 & 0 \\ 0 & a_2 \end{vmatrix} = a_1 a_2 \neq 0 \qquad x_1 = \frac{Y_1 - b_1}{a_1} \quad x_2 = \frac{Y_2 - b_2}{a_2}$$

$$f_{Y_1, Y_2}(y_1, y_2) = \phi_{X_1 X_2}\left(\frac{Y_1 - b_1}{a_1}, \frac{Y_2 - b_2}{a_2}\right) / a_1 a_2$$

It can be shown that f_{Y_1, Y_2} is the bivariate normal density with parameters

$$\mu_{Y_1} = a_1 \mu_{X_1} + b_1 \qquad \sigma_{Y_1} = a_1 \sigma_{X_1}$$

$$\mu_{Y_2} = a_2 \mu_{X_2} + b_2 \qquad \sigma_{Y_2} = a_2 \sigma_{X_2}$$

and ρ remains the same.

56

#5 (cont.)

c) The density is not uniform over the disk. The following density is uniform over the disk:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi} & -\sqrt{1-x^2} \le y \le \sqrt{1-x^2}, \quad -1 \le x \le 1 \\ 0 & \text{o.w.} \end{cases}$$

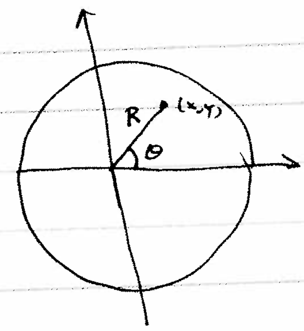
55

§3.6 #55

$R \sim Unif\,[0,1]$

$\Theta \sim Unif\,[0,2\pi)$

$R \perp\!\!\!\perp \Theta$



a)

$X = R\cos\Theta \qquad R = \sqrt{X^2+Y^2}$

$Y = R\sin\Theta \qquad \Theta = \tan^{-1}\left(\frac{Y}{X}\right)$

$$J(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$$

$\theta = \tan^{-1}\left(\frac{Y}{X}\right) \qquad r = X^2+Y^2$

$$f_{R,\Theta}(r,\theta) = \begin{cases} \frac{1}{2\pi} & 0 \le r \le 1 \quad 0 \le \theta \le 2\pi \\ 0 & o.w. \end{cases}$$

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2\pi\,(x^2+y^2)^{1/2}} & 0 \le X^2+Y^2 \le 1 \\ 0 & o.w. \end{cases}$$

b) $$f_X(x) = \frac{1}{2\pi}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2+y^2)^{-1/2}\,dy = \frac{1}{2\pi}\log\frac{1+\sqrt{1-x^2}}{1-\sqrt{1-x^2}}$$

$-1 \le X \le 1$

$$f_Y(y) = \frac{1}{2\pi}\log\frac{1+\sqrt{1-y^2}}{1-\sqrt{1-y^2}} \qquad -1 \le y \le 1$$

§3.6

#2

(54)

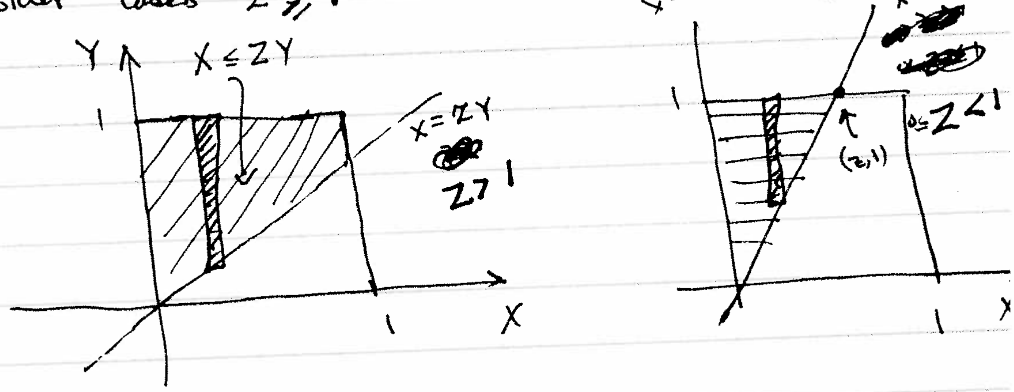
Let $X \sim U(0,1)$ indep $Y \sim U(0,1)$

Let $Z = \frac{X}{Y}$ $\qquad Z \geq 0$

$$F_Z(z) = P(Z \le z) = P\left(\frac{X}{Y} \le z\right)$$

$$= P(X \le zY)$$

Consider Cases $Z \geq 1$ and $0 \le Z \le 1$



Case $Z \geq 1$

$$F_Z(z) = \int_0^1 \int_{\frac{x}{z}}^1 dy\, dx = \int_0^1 \left(1 - \frac{x}{z}\right) dx = x - \frac{x^2}{2z}\Big|_0^1 = 1 - \frac{1}{2z}$$

Case $0 \le z \le 1$

$$F_Z(z) = \int_0^z \int_{\frac{x}{z}}^1 dy\, dx = x - \frac{x^2}{2z}\Big|_0^z = z - \frac{z}{2}$$

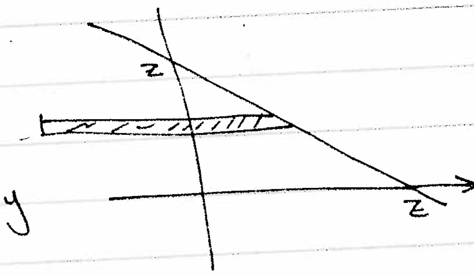
$$f_Z(z) = \begin{cases} \frac{1}{2} z^{-2} & z \geq 1 \\ \frac{1}{2} & 0 \le z \le 1 \end{cases}$$

53

§3.6

#47 $X \sim N(0,1)$ indep $Y \sim N(0,1)$

Determine the dist^{n} $Z = X + Y$



$$F_Z(z) = P(X+Y \le z)$$

$$= \iint_{\{(x,y):\, x+y \le z\}} f_X(x)\, f_Y(y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}\, dx\, dy$$

Let $u = x + y$, $du = +dx$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} \frac{1}{2\pi} e^{-\frac{1}{2}\left[u^2 - 2uy + 2y^2\right]}\, du\, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} \frac{1}{2\pi} e^{-\frac{1}{2}\left[(\sqrt{2}y - \frac{1}{\sqrt{2}}u)^2 + \frac{1}{2}u^2\right]}\, du\, dy$$

See ~~[illegible]~~ page 82 of text.

$$= \frac{1}{2\pi} \int_{-\infty}^{z} e^{-\frac{1}{4}u^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\sqrt{2}y - \frac{1}{\sqrt{2}}u)^2}\, dy\, du$$

$$= \frac{1}{2\pi} \int_{-\infty}^{z} e^{-\frac{1}{4}u^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left[2(y - \frac{1}{2}u)^2\right]}\, dy\, du$$

Compare to $N(\frac{1}{2}u, \frac{1}{2})$

$$= \frac{1}{2\pi} \int_{-\infty}^{z} e^{-\frac{1}{4}u^2} \left(\sqrt{2\pi}\, \frac{1}{\sqrt{2}}\right) du = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{z} e^{-\frac{1}{4}u^2}\, du$$

$Z \sim N(0, 2)$ with $f_Z(z) = \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}z^2}$ $\quad -\infty < z < \infty$

52

§ 3.6

#4

The joint pmf of N_1 and N_2 is (because of indep.)

$$p(n_1, n_2) = \frac{e^{-(\lambda_1+\lambda_2)} \lambda_1^{n_1} \lambda_2^{n_2}}{n_1! \; n_2!}$$

$$P(N=k) = P(N_1+N_2=k)$$

$$= \sum_{\{(n_1,n_2):\, n_1+n_2=k\}} p(n_1, n_2)$$

$$= \sum_{n_1=0}^{k} p(n_1, k-n_1)$$

$$= \sum_{n_1=0}^{k} \frac{e^{-(\lambda_1+\lambda_2)} \lambda_1^{n_1} \lambda_2^{k-n_1}}{n_1! \; (k-n_1)!}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_{n_1=0}^{k} \frac{k!}{n_1! \, (k-n_1)!} \lambda_1^{n_1} \lambda_2^{k-n_1}$$

$$= e^{-(\lambda_1+\lambda_2)} \cdot \frac{1}{k!} (\lambda_1+\lambda_2)^k$$

$$= p_N(k)$$

where $N \sim$ Poiss$(\lambda_1+\lambda_2)$

51

§3.6 #~~3~~2a

independent

$T \sim \exp(\lambda) \qquad W \perp\!\!\!\perp T$

$P(W=1) = P(W=-1) = \frac{1}{2}$

$X = WT$

Consider two cases:

$x \leq 0$

$$F_X(x) = P(X \leq x)$$
$$= P(WT \leq x,\ W=-1) + \underbrace{P(WT \leq x,\ W=1)}_{=0}$$
$$= P(WT \leq x \mid W=-1)\, P(W=-1)$$
$$= \frac{1}{2} P(T \geq -x) = \frac{1}{2}\left[1-(1-e^{+\lambda x})\right] = \frac{1}{2} e^{+\lambda x}$$

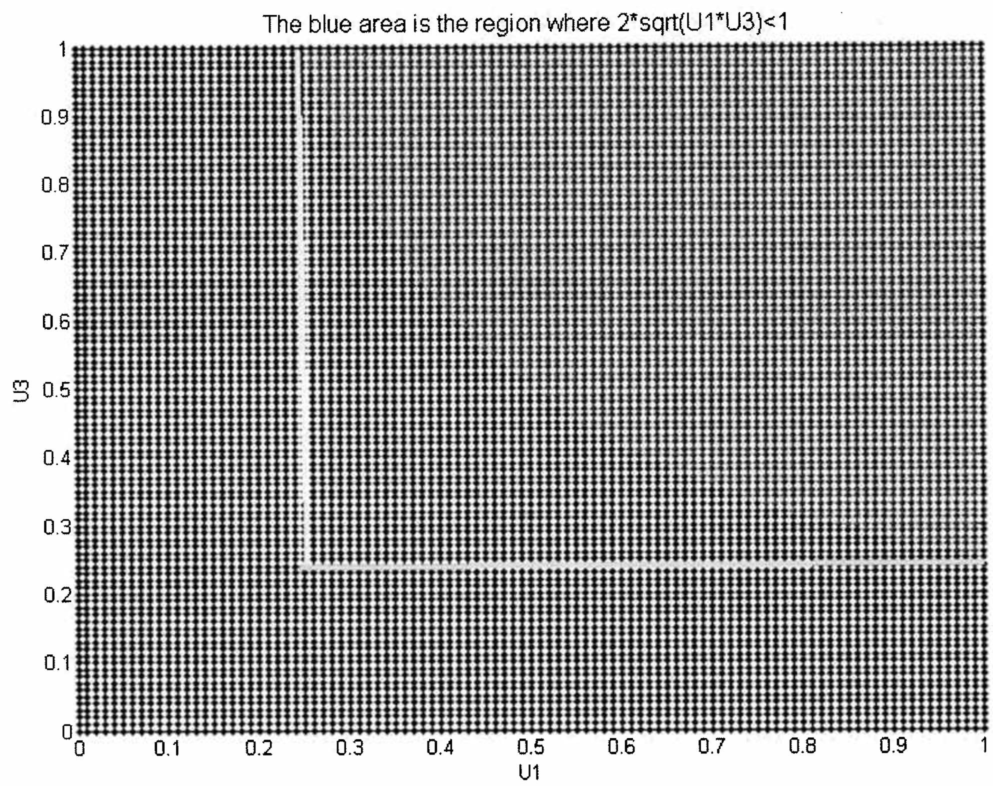
$$f_X(x) = \frac{+\lambda}{2} e^{+\lambda x} \qquad x<0 \qquad (*)$$

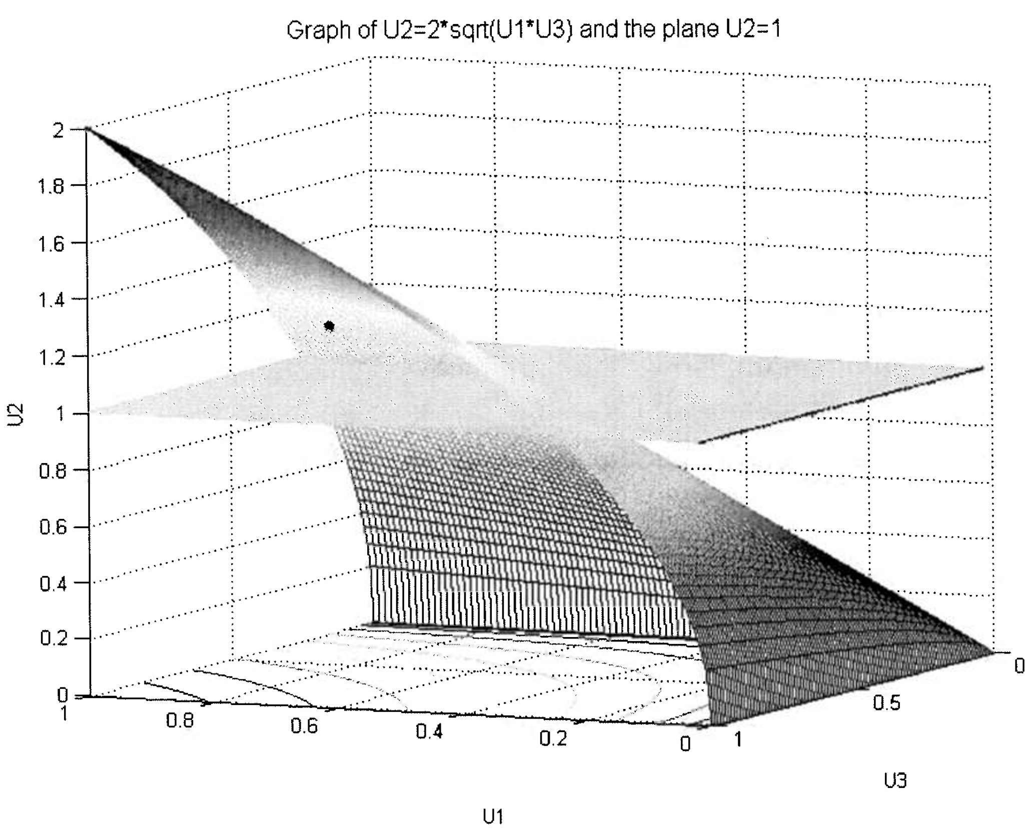
$x \geq 0$

$$F_X(x) = P(X \leq x)$$
$$= P(WT \leq x,\ W=1) + \underbrace{P(WT \leq x,\ W=-1)}_{=0}$$
$$= P(WT \leq x \mid W=1)\ P(W=1)$$
$$= \frac{1}{2} P(T \leq x) = \frac{1}{2}\left(1-e^{-\lambda x}\right)$$

$$f_X(x) = \frac{\lambda}{2} e^{-\lambda x} \qquad x>0 \qquad (**)$$

$(*)$ and $(**)$ $\Rightarrow$ $f_X(x) = \frac{\lambda}{2} e^{-\lambda |x|}$


The blue area is the region where 2*sqrt(U1*U3)<1
U3
U1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


Graph of U2=2*sqrt(U1*U3) and the plane U2=1
U2
U1
U3
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0.5

§3.6

#11 Roots of the quadratic equation are real iff $U_2^2 - 4U_1U_3 \geq 0$. Thus we need the following probability:

$$P[U_2^2 - 4U_1U_3 \geq 0] = 1 - P[U_2^2 \leq 4U_1U_3]$$

$$= 1 - P[-2\sqrt{U_1U_3} \leq U_2 \leq 2\sqrt{U_1U_3}]$$

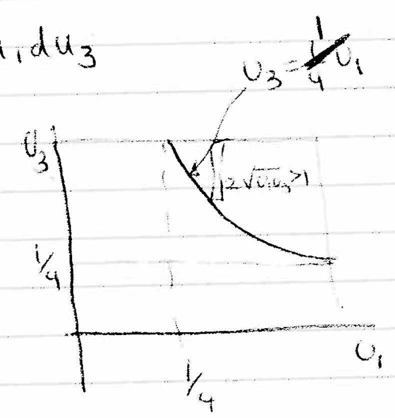
$$= 1 - P[0 \leq U_2 \leq \min(1, 2\sqrt{U_1U_3})]$$

Below $0 \leq U_1 \leq 1$ and $0 \leq U_2 \leq 1$ is assumed.

$$P[0 \leq U_2 \leq \min(1, 2\sqrt{U_1U_3})] = \iint_{2\sqrt{U_1U_3} \leq 1} \int_0^{2\sqrt{u_1u_3}} du_2\,du_1\,du_3$$

$$+ \iint_{2\sqrt{U_1U_3} > 1} \int_0^1 du_2\,du_1\,du_3$$ see the 3-dimentional graph

$$= \iint_{2\sqrt{U_1U_3} \leq 1} 2\sqrt{U_1U_3}\,du_1\,du_3 + \iint_{2\sqrt{U_1U_3} > 1} du_1\,du_3$$



$$\iint_{2\sqrt{U_1U_3} > 1} du_1\,du_3 = \int_{1/4}^1 \int_{1/4u_1}^1 du_3\,du_1$$

$$= \frac{3}{4} - \frac{1}{2}\log(2)$$

$$\iint_{2\sqrt{U_1U_3} \leq 1} 2\sqrt{U_1U_3}\,dU_1\,dU_3 = \int_0^{.25}\int_0^1 2\sqrt{U_1U_3}\,dU_3\,dU_1 + \int_{.25}^1 \int_0^{1/4U_1} 2\sqrt{U_1U_3}\,dU_3\,dU_1 = \frac{1}{9} + \frac{1}{3}\log(2)$$

$\therefore P[U_2^2 - 4U_1U_3 \geq 0] = 1 - \frac{3}{4} + \frac{1}{2}\log(2) - \frac{1}{9} - \frac{1}{3}\log(2) \cong .2456$

Problem 40 Contined

Now if we generate

a. $Z \sim \text{unif}(1, \ldots, 100)$

b. accept z with probability $p_{Y|X}(44|z)$

which is

$$p_{Y|X}(44|z) = \begin{cases} \frac{1}{z} & 44 \leq z \leq 100 \\ 0 & o.w. \end{cases}$$

we see that the generated z's are between 44 to 100 only, and the proportion is controlled by $\frac{1}{z}$

Which means that you would get the same proportion as $X|Y=44$.

Problem 40 Continued.

$X \sim \text{Unif}(1, 2, \ldots, 100)$

$Y | X x \sim \text{Unif}(1, 2, \ldots, x)$

$$P(X=x \mid Y=44) = \frac{P(Y=44 \mid X=x) \cdot P(X=x)}{P(Y=44)}$$

$$= \frac{P(Y=44 \mid X=x) \cdot P(X=x)}{\sum_{x=1}^{100} P(Y=44 \mid X=x) \cdot P(X=x)}$$

$$= \frac{P(Y=44 \mid X=x) \cdot P(X=x)}{\sum_{X=44}^{100} P(Y=44 \mid X=x) \cdot P(X=x)}$$

$$= \begin{cases} \dfrac{(1/x)\left(\frac{1}{100}\right)}{\sum_{X=44}^{100} \left[\frac{1}{100x}\right]} & 44 \le X \le 100 \\ 0 & \text{o.w.} \end{cases}$$

$$= \begin{cases} \dfrac{1}{x \sum_{X=44}^{100} \frac{1}{x}} & 44 \le x \le 100 \\ 0 & \text{o.w.} \end{cases}$$

We know that $X \geq 44$ and with the probability decreasing proportional to $\frac{1}{x}$.

Problem 40

To clarify let's reword the algorithm:

a. Generate $Z \sim p_X(z)$.
b. Accept z with probability $p_{Y|X}(y|z)$ for a given value of y.

Let $\mathbb{Z}$ be the random variable that is generated by going through steps a and b.

Then

$$p_Z(z) = \mathbb{P}(Z=z) = p_X(z) \cdot p_{Y|X}(y|z)$$

$$= P(X=z) \cdot \frac{P(Y=y, X=z)}{P(X=z)}$$

$$= P(X=z \mid Y=y) \cdot P(Y=y)$$

Note that $P(Y=y)$ is just a constant value for a given y. thus the proportion of generated numbers are controlled by the conditional $\mathbb{P}(X=z \mid Y=y)$ or $p_{X|Y}$

To give a simple example, let
$X \sim$ Bernoulli $(\frac{1}{2})$
$Y \sim$ Binomial $(2, \frac{1}{2})$ $\quad X \perp\!\!\!\perp Y$

Consider $y=2$.
Generate $Z \sim$ Bernoulli$(\frac{1}{2})$ i.e. $Z = \begin{cases} 1 & \text{w/p } \frac{1}{2} \\ 0 & \text{w/p } \frac{1}{2} \end{cases}$

$P(Y=2 \mid Z=1) = P(Y=2)$

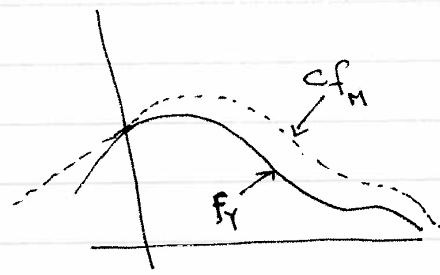
Now
$$\left.\begin{aligned} P(Z=0) &= \left(\tfrac{1}{2}\right)\left(\tfrac{1}{4}\right) = \tfrac{1}{8} \\ P(Z=1) &= \tfrac{1}{2}\left(\tfrac{1}{4}\right) = \tfrac{1}{8} \end{aligned}\right\}$$

Note that at the end $P(Z=0) = P(Z=1) = \frac{1}{2}$ since no other value is generated

49

#38 Since we generate independent random variates, acceptance or rejection of a value at any given iteration is independent of any other iteration. The probability that we accept is

$$\cancel{P(accept) = \frac{\int f_M \, random}{c \int f_M \, random \, dy}}$$



$$P(\text{accept}) = \frac{\int f_Y(t)\,dt}{c \int f_M(t)\,dt}$$

$$= \frac{1}{c}$$

where $c = \sup_y \frac{f_Y(y)}{f_M(y)}$

The closer f_Y to f_M, the closer c will be to one, and therefore the rate of acceptance goes up.

(48)

39 (Continued)

So the algorithm will be as follows

Step 1: Generate U and V indep Unif(0,1) and let $m = -\frac{1}{2}\log[1-(1-e^{-4})V]$.

Step 2: Let $c = \frac{15(1-e^{-4})}{8}$. If

$$U < \frac{1}{c} \cdot \frac{\frac{15}{16}(m-1)^2(2-m)^2}{\left(\frac{2}{1-e^{-4}}\right)e^{-2m}}$$

$$= \frac{1}{4}\,\frac{(m-1)^2(2-m)^2}{e^{-2m}}$$

Then set $Y = m$ (or $X = m+1$ for the shift)

o.w. reject and go to step 1.

The probability of acceptance, in general is

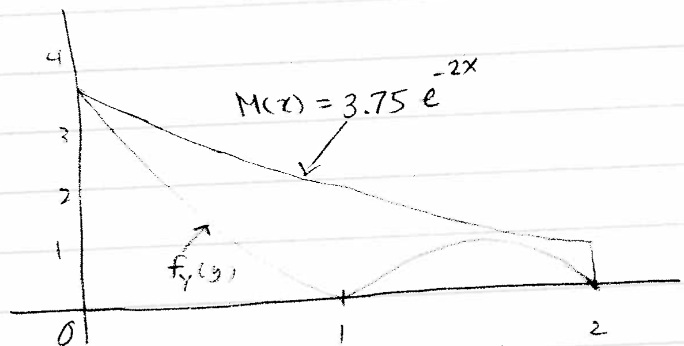
$$\frac{\int_0^2 \left(\frac{15}{16}\right)(y-1)^2(2-y)^2\,dy}{\int_0^2 3.75\,e^{-2x}\,dx} = \frac{1}{1.847} \cong 0.54$$

Note this is $\frac{1}{c}$

Note that If we would have chosen Uniform in place of m(x), the ratio would have been $\frac{1}{7.5} \cong 0.13$

47

(#37 Continued) The following figure shows the graph of $f_Y(y)$ and $c\,m(x)$, where $c = \frac{3.75(1-e^{-4})}{2}$.



Note that generating random values from $m(x)$ is simple, because its cumulative distribution function is

$$G(x) = \frac{1-e^{-2x}}{1-e^{-4}}$$

and

$$G^{-1}(x) = -\frac{1}{2}\log[1-(1-e^{-4})x]$$

Thus we can genenerate a $U \sim unif(0,1)$ and use

$$-\frac{1}{2}\log[1-(1-e^{-4})U].$$

Now

$$\sup_{0 \le y \le 2} \frac{\frac{15}{16}(y-1)^2(2-y)^2}{\frac{2}{1-e^{-4}}e^{-2y}}$$

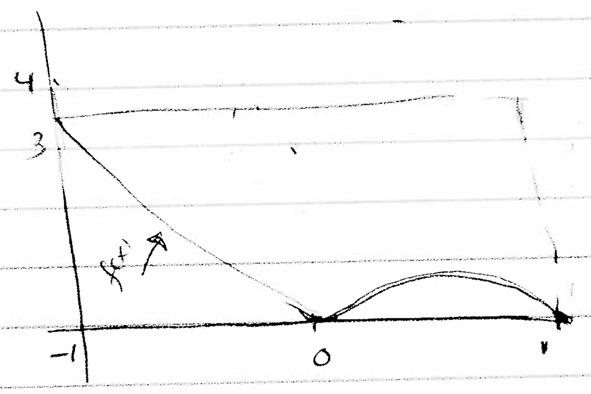
this maximum occurs at $y=0$ with value

$$c = \frac{15}{4}\Big/\frac{2}{1-e^{-4}} = \frac{15(1-e^{-4})}{8}$$

46

§3.5 #37 first note that the given $f(x)$ is not a density. We multiply it by a constant to make it a density. The result is

$$f(x) = \frac{15}{16} x^2(1-x)^2 \qquad -1 \le x \le 1$$

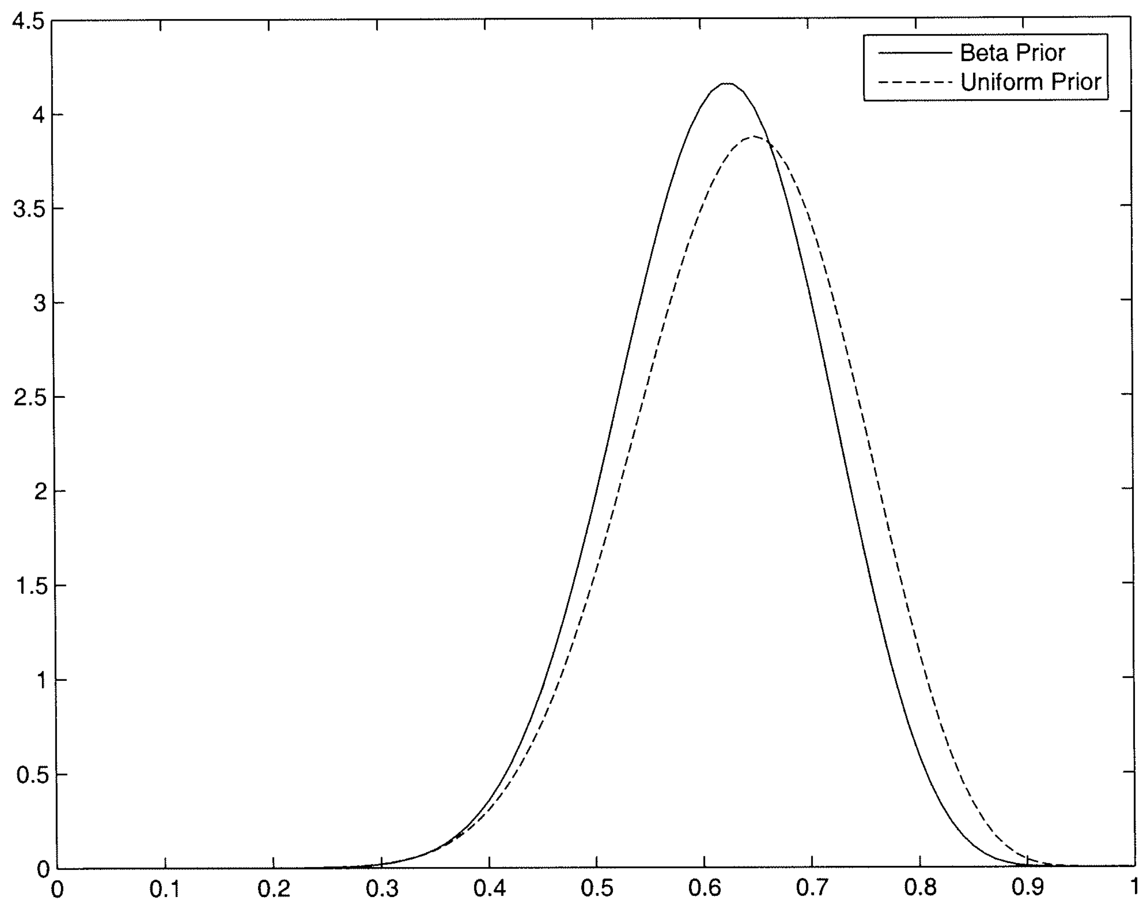


The maximum of $f(x)$ occures at $x=-1$ and it is $\frac{15}{4} = 3.75$. One can use $M \sim \text{Unif}(-1,1)$ and with $c = 3.75/2$. However, this will be very inefficient. Noting that the density is decreasing like an exponential function, I will try a density of the form $m(x) = c\, e^{-\lambda x}$ $0 \le x \le 2$ (I shift from $(-1,1)$ to $(0,2)$ for convenience). If we shift the density of X we get $Y = X+1$ with

$$f(y) = \frac{15}{16}(y-1)^2(2-y)^2 \qquad 0 \le y \le 2$$

After examining a few candidates for λ, $\lambda = 2$ seem to work well. So I choose the density

$$m(x) = \frac{2}{1-e^{-4}}\, e^{-2x} \qquad 0 < x < 2$$


Beta Prior
Uniform Prior
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

#34

$\Theta \sim Beta(a=3, b=3)$

$X \mid \Theta=\theta \sim Binomial(n, \theta)$

$$f_{\Theta|X}(\theta|x) = \frac{f_{X|\Theta}(x|\theta) f_{\Theta}(\theta)}{\int_0^1 f_{X|\Theta}(x|\theta) f_{\Theta}(\theta)\, d\theta}$$

$$= \frac{\binom{n}{x} \theta^{x}(1-\theta)^{n-x} \cdot 30 \cdot \theta^{2}(1-\theta)^{2}}{30\binom{n}{x} \int_0^1 \theta^{x+2}(1-\theta)^{n-x+2}\, d\theta}$$

$$= \frac{\theta^{x+2}(1-\theta)^{n-x+2}}{\int_0^1 \theta^{x+2}(1-\theta)^{n-x+2}\, d\theta}$$

$\Theta \mid X \sim Beta(a=x+3,\ b=n-x+3)$

The plot for $n=20$ and $x=13$ is shown on the next page along with the posterior based on Uniform. It can be seen that the posterior based on the beta is concentrated more towards 0.5, as compared to that corresponding to the uniform. Thus our prior has made a difference

#33a

$$f_\Theta(\theta) = \begin{cases} 1 & 0 \le \theta \le 1 \\ 0 & o.w. \end{cases}$$

N = number of times until a head comes up.

$N \sim \text{geometric}(\theta)$

$$f_{N|\theta}(n|\theta) = (1-\theta)^{n-1}\theta \qquad n = 1, 2, \ldots$$

$$f_{\theta|N}(\theta|n) = \frac{f_{N|\theta}(n|\theta)\, f_\Theta(\theta)}{\int_0^1 f_{N|\theta}(n|\theta)\, f_\Theta(\theta)\, d\theta}$$

$$= \frac{(1-\theta)^{n-1}\theta}{\int_0^1 (1-\theta)^{n-1}\theta\, d\theta}$$

$$= \frac{(1-\theta)^{n-1}\theta}{\Gamma(2)\Gamma(n)/\Gamma(n+2)}$$

$$= n(n+1)(1-\theta)^{n-1}\theta \qquad 0 \le \theta \le 1$$

$\theta|N \sim \text{Beta}(a=2, b=n)$

#32

Max $\theta^{x}(1-\theta)^{n-x}$

Let $f(\theta|x) = \theta^{x}(1-\theta)^{n-x}$

Consider $\ell(\theta|x) = \cancel{x}\log f(\theta|x) = x\log\theta + (n-x)\log(1-\theta)$

$$\frac{\partial}{\partial\theta}\ell(\theta|x) = \frac{x}{\theta} - \frac{n-x}{1-\theta} = 0$$

$$\frac{x - x\theta - n\theta + x\theta}{\theta(1-\theta)} = 0$$

$$\Rightarrow \boxed{\theta = \frac{x}{n}}$$

This makes sense, because the estimate of θ is the proportion of heads obtained.

~~#33 $\theta \sim unif(0,1)$~~

~~$(N|\theta) \sim Geom(\theta)$~~

~~$f(\theta|N) =$~~

Chapter 3 # 25

Let $Z \sim \text{Bernoulli}(\frac{1}{2})$

then

$$Y = X \quad \text{if } Z = 0$$
$$Y = -X \quad \text{if } Z = 1$$

$$\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(Y \le y \mid Z=0)\,P(Z=0) + P(Y \le y \mid P(Z=1) \\
&= P(X \le y)\,\tfrac{1}{2} + P(-X \le y)\,\tfrac{1}{2} \\
&= \tfrac{1}{2}\left[F_X(y) + 1 - F_X(-y)\right]
\end{aligned}$$

$$f_Y(y) = \frac{\partial}{\partial y} F_Y(y) = \frac{1}{2}\left[f_X(y) + f_X(-y)\right]$$

Now clearly

$$f_Y(y) = f_Y(-y).$$

45

§3.5 #24 $P \sim \text{unif}(0,1)$

$X \mid P=p \sim \text{Bernouli}(p)$

Want to find Conditional dist$^{\underline{n}}$ of $P|X$

$$f_{P|X}(p|x) = \frac{f_{P,X}(P=p, X=x)}{f_X(x)}$$

$$= \frac{f_{X|P}(X=x \mid P=p)\, f_P(p)}{\int_0^1 f_X(x|p)\, f_p(p)\, dp}$$

$$= \frac{p^x (1-p)^{1-x}}{\int_0^1 p^x (1-p)^{1-x}\, dp} \qquad x=0,1$$

If $x=0$ $\int_0^1 (1-p)\, dp = \frac{(1-p)^2}{2}\Big|_0^1 = \frac{1}{2}$

If $x=1$ $\int_0^1 p\, dp = \frac{p^2}{2}\Big|_0^1 = \frac{1}{2}$

$\therefore$ $f_{p|x}(p|x) = 2p^x(1-p)^{1-x}$ $\quad x=0,1$

$0 \le p \le 1$

$$f_{p|x}(p|x) = \begin{cases} 2p & \text{if } x=0 \\ 2(1-p) & \text{if } x=1 \end{cases}$$

§3.5 #23

$X|N \sim \text{Binomial}(N, p)$

$N \sim \text{Binomial}(m, r)$

$$p_X(x) = \sum_{n=x}^{m} p_{X|N}(x|n)\, p_N(n) \qquad x = 0, 1, \ldots, m$$

$$= \sum_{n=x}^{m} \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{n} r^n (1-r)^{m-n}$$

$$= \sum_{n=x}^{m} \frac{n!}{x!(n-x)!} \cdot \frac{m!}{n!(m-n)!} p^x r^n (1-p)^{n-x} (1-r)^{m-n}$$

$$= \frac{m!}{x!} \left(\frac{p}{1-p}\right)^x (1-r)^m \sum_{n=x}^{m} \frac{1}{(n-x)!(m-n)!} \left[\frac{r(1-p)}{1-r}\right]^n$$

Let $q = n - x$

$$= \frac{m!}{\underbrace{(m-x)!\,x!}} \left(\frac{p}{1-p}\right)^x (1-r)^m \sum_{q=0}^{m-x} \frac{(m-x)!}{q!\,(m-q-x)!} \left[\frac{r(1-p)}{1-r}\right]^{q+x}$$

$$= \binom{m}{x} \left(\frac{p}{1-p}\right)^x (1-r)^m \sum_{q=0}^{m-x} \binom{m-x}{q} \left(\frac{r(1-p)}{1-r}\right)^{q+x}$$

$$= \binom{m}{x} \left(\frac{p}{1-p}\right)^x (1-r)^m \left(\frac{r(1-p)}{1-r}\right)^m \sum_{q=0}^{m-x} \binom{m-x}{q} \left(\frac{1-r}{r(1-p)}\right)^{m-q}$$

$$= \binom{m}{x} p^x (1-p)^{m-x} r^m \left[1 + \frac{1-r}{r(1-p)}\right]^{m-x}$$

$$= \binom{m}{x} p^x (1-p)^{m-x} r^m \left(\frac{1-rp}{r(1-p)}\right)^{m-x}$$

$$= \binom{m}{x} (pr)^x (1-rp)^{m-x}$$

$X \sim \text{Binomial}(m, rp)$

44

§3.5 #22

Let $X_0 = N(t_0, t_1) \sim \mathcal{P}(\lambda(t_1 - t_0))$

$X_1 = N(t_1, t_2) \sim \mathcal{P}(\lambda(t_2 - t_1))$

$X_2 = N(t_0, t_2) \sim \mathcal{P}(\lambda(t_2 - t_0))$

We seek the Conditional distribution of $X_0 | X_2 = n$

$$P(X_0 = k \mid X_2 = n) = \frac{P(X_0 = k,\ X_2 = n)}{P(X_2 = n)}$$

$$= \frac{P(X_2 = n \mid X_0 = k)\, P(X_0 = k)}{P(X_2 = n)}$$

$$= \frac{P(X_1 = n - k)\, P(X_0 = k)}{P(X_2 = n)}$$

$$= \frac{\dfrac{e^{-\lambda(t_2 - t_1)} [\lambda(t_2 - t_1)]^{n-k}}{(n-k)!} \; \dfrac{e^{-\lambda(t_1 - t_0)} [\lambda(t_1 - t_0)]^{k}}{k!}}{e^{-\lambda(t_2 - t_0)} [\lambda(t_2 - t_0)]^{n} / n!}$$

$$= \binom{n}{k} \left(\frac{t_1 - t_0}{t_2 - t_0}\right)^{k} \left(\frac{t_2 - t_1}{t_2 - t_0}\right)^{n-k} \sim \text{Binomial}\left(n, \frac{t_1 - t_0}{t_2 - t_0}\right.$$

43

§ 3.5

\# 20 $X_1 \sim Unif(0,1)$

$X_2 | X_1 \sim Unif[0, X_1]$

$$f_{X_1,X_2}(x_1,x_2) = f_{X_2|X_1}(x_2|x_1)\, f_{X_1}(x_1)$$

$$= \frac{1}{x_1} \qquad 0 \le x_2 \le x_1 \qquad x_1 \in [0,1]$$

\# 21 X = Concentration of the Chemical $\sim f_X(x)$

Let $Z = \begin{cases} 1 & \text{if the Chemical is detected} \\ 0 & \text{otherwise} \end{cases}$

$$P(Z=1 | X=x) = R(x)$$
$$P(Z=0 | X=x) = 1 - R(x)$$

Y = Concentration of the chemical given that it is detected.

Distribution of Y is $X | Z=1$

$$f_Y(y) = f_{X|Z}(y|z=1) = \frac{f_{X,Z}(y,1)}{f_Z(1)} = \frac{f_{Z|X}(z=1|y) f_X(y)}{\int_0^\infty f_{Z|X}(1|x)\, f_X(x)\,dx}$$

$$= \frac{R(y)\, f_X(y)}{\int_0^\infty R(x) f_X(x)\,dx}$$

§3.5

42

12c

$$f_{Y|X}(y|x) = \frac{1/8\,(x^2-y^2)\,e^{-x}}{\frac{1}{6}e^{-x}\,x^3} = \frac{3}{4}\,\frac{x^2-y^2}{x^3}$$

$-x \le y \le x$

$x \geq 0$

if $y>0$

$$f_{X|Y}(x|y) = \frac{1/8\,(x^2-y^2)\,e^{-x}}{1/4\,e^{y}\,(1+y)} = \frac{1}{2}\,\frac{(x^2-y^2)e^{-x}}{(1+y)e^{-y}}$$

$x>y \quad , x\geq 0$

~~if $y>0$ then $x>y$~~

~~if $y<0$ then $x>y$~~

≥ 0

If $y<0$

$$f_{X|Y}(x|y) = \frac{\frac{1}{8}(x^2-y^2)e^{-x}}{\frac{1}{4}e^{y}(1-y)} = \frac{1}{2}\,\frac{(x^2-y^2)e^{-x}}{e^{y}(1-y)}$$

$y \geq -x \qquad x>0$

41

§3.5

8c.

$$f_{Y|X}(y|x) = \frac{(6/7)(x+y)^2}{\frac{1}{7}(6x^2+6x+2)} = \frac{6(x+y)^2}{6x^2+6x+2}$$

$0 \le y \le 1$

$0 \le x \le 1$

$$f_{X|Y}(x|y) = \frac{(6/7)(x+y)^2}{\frac{1}{7}(6y^2+6y+2)} = \frac{6(x+y)^2}{6y^2+6y+2}$$

$0 \le y \le 1$

$0 \le x \le 1$

9b.

$$f_{Y|X}(y|x) = \frac{3/4}{3/4(1-x^2)} = \frac{1}{1-x^2} \qquad 0 \le y \le 1-x^2$$

$$f_{X|Y}(x|y) = \frac{3/4}{3/2\sqrt{1-y}} = \frac{1}{2\sqrt{1-y}} \qquad -\sqrt{1-y} \le x \le \sqrt{1-y}$$

40

§3.4

#40

$X \backslash Y$	0	1	2
0	0	1	2
1	1	2	3
2	2	3	4

← $X+Y$

$$P(X+Y=0) = P(X=0 \cap Y=0) = \frac{1}{9}$$

$$P(X+Y=1) = P(X=0 \cap Y=1) + P(X=1, Y=0) = \frac{2}{9}$$

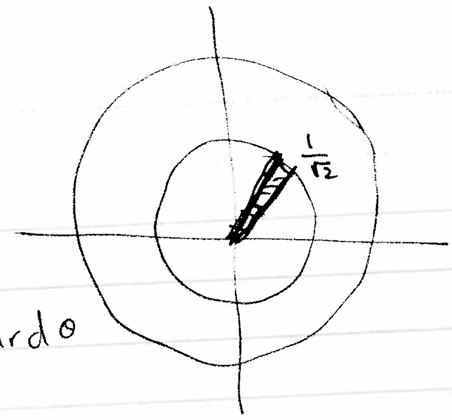
$$P(X+Y=2) = \frac{3}{9}$$

$$P(X+Y=3) = \frac{2}{9}$$

$$P(X+Y=4) = \frac{1}{9}$$

39

#15c (Continued)



c. $P(X^2+Y^2 \le \frac{1}{2})$

$$= \frac{3}{2\pi} \int_0^{2\pi} \int_0^{\frac{1}{\sqrt{2}}} \sqrt{1-r^2}\, r\, dr\, d\theta$$

$$= 0.6464$$

d. $f_X(x) = \frac{3}{2\pi} \int_{-\sqrt{u}}^{\sqrt{u}} \sqrt{u-y^2}\, dy$ where $u = 1-x^2$

$$= \frac{3}{2\pi} \left(\frac{1}{2} u\pi\right)$$

$$= \frac{3}{4} u = \frac{3}{4} (1-x^2)$$

$$f_Y(y) = \frac{3}{4}(1-y^2)$$

15e)

$$f_{X|Y}(x|y) = \frac{\frac{3}{2\pi}\sqrt{1-(x^2+y^2)}}{\frac{3}{4}(1-y^2)}$$

$$f_{Y|X}(y|x) = \frac{\frac{3}{2\pi}\sqrt{1-(x^2+y^2)}}{\frac{3}{4}(1-x^2)}$$

38

§3.5

#14 b

$$f_{X|Y}(x|y) = \frac{x e^{-x(y+1)}}{1/(y+1)^2} = (y+1)^2 x e^{-x(y+1)}$$

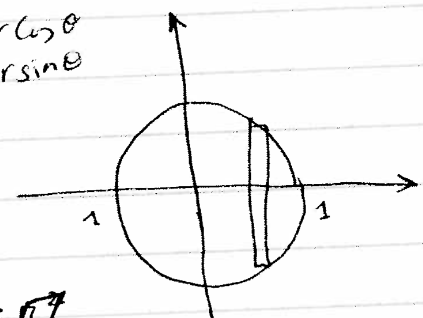
$$0 \le y \le \infty$$

$$f_{Y|X}(y|x) = \frac{x e^{-x(y+1)}}{e^{-x}} = x e^{-xy} \qquad y>0,\ x>0$$

#15

$x = r\cos\theta$
$y = r\sin\theta$

a) $$\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-(x^2+y^2)}\, dy\, dx$$



$$= \int_{-\pi}^{\pi}\int_{0}^{1} \sqrt{1-r^2}\, r\, dr\, d\theta = \frac{2}{3}\pi$$

$$\Rightarrow \quad C = \frac{3}{2\pi}$$

b) see next page

37

#19 (Cont.)

$$= \int_0^\infty \alpha\beta\left(\frac{1}{\beta}\right)\left[1-e^{-\beta t_1}\right]e^{-\alpha t_1}\,dt_1$$

$$= \alpha\left[\int_0^\infty e^{-\alpha t_1}\,dt_1 - \int_0^\infty e^{-(\alpha+\beta)t_1}\,dt_1\right]$$

$$= \alpha\left[\frac{-1}{\alpha}e^{-\alpha t_1}\Big|_0^\infty + \frac{1}{\alpha+\beta}e^{-(\alpha+\beta)t_1}\Big|_0^\infty\right]$$

$$= \alpha\left(1/\alpha - \frac{1}{\alpha+\beta}\right) = \frac{\beta}{\alpha+\beta}$$

b) Note that if $X \sim \exp(\lambda)$ then $Y = aX$ $(a>0)$ is distributed as $\exp\left(\frac{\lambda}{a}\right)$.

because

$$F_Y(y) = P(Y \le y) = P(aX \le y) = P\left(X \le \frac{y}{a}\right)$$
$$= 1 - e^{-\frac{\lambda}{a}y}$$

$\therefore$ Let $Z = 2T_2 \sim \exp\left(\beta/2\right)$

$$P(T_1 > Z) = \frac{\beta/2}{\alpha+\beta/2} = \frac{\beta}{2\alpha+\beta}$$

36

#14a

$$f_X(x) = \int_0^\infty x e^{-x(y+1)}\, dy$$

$$= e^{-x} \qquad x \geq 0$$

$$f_Y(y) = \int_0^\infty x e^{-x(y+1)}\, dx \qquad u = x \quad dv = e^{-x(y+1)}$$

$$du = dx \quad v = \frac{-1}{y+1} e^{-x(y+1)}$$

$$= \frac{-x}{y+1} e^{-x(y+1)}\Big|_0^\infty - \int_0^\infty \frac{-1}{y+1} e^{-x(y+1)}\, dx$$

$$= \frac{1}{y+1} \int_0^\infty e^{-x(y+1)}\, dx = \left(\frac{1}{y+1}\right) \frac{-1}{y+1} e^{-x(y+1)}\Big|_0^\infty$$

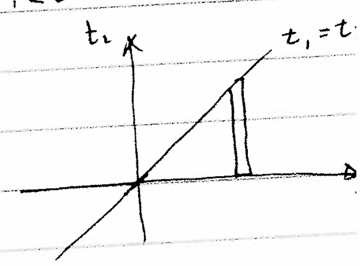
$$= \frac{1}{(y+1)^2} \qquad y \geq 0$$

Since $f(x,y) \neq f_X(x)\, f_Y(y)$, X and Y are not independent.

#19

$T_1 \sim exp(\alpha)$ $\qquad T_2 \sim exp(\beta)$

$$f(t_1, t_2) = \alpha\beta e^{-\alpha t_1 - \beta t_2} \qquad 0 < t_1 < \infty \quad 0 < t_2 < \infty$$

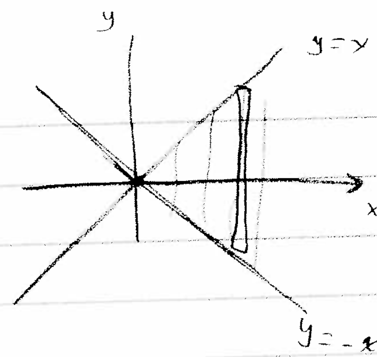


a)

$$P(T_1 > T_2) = \int_0^\infty \int_0^{t_1} \alpha\beta e^{-\alpha t_1 - \beta t_2}\, dt_2\, dt_1$$

$$= \int_0^\infty \alpha\beta e^{-\alpha t_1} \frac{-1}{\beta} e^{-\beta t_2}\Big|_0^{t_1} dt_1$$

34



#12

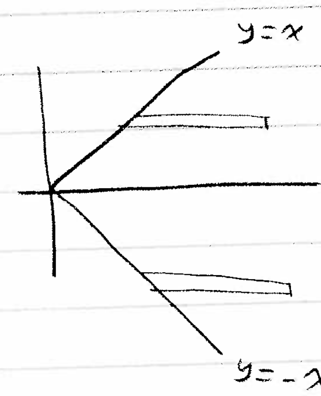
a) $c = 1 \Big/ \int_0^{\infty}\int_{-x}^{x} (x^2-y^2)e^{-x}\,dy\,dx$

$= 1/8$

b) $f_X(x) = \int_{-x}^{x} \tfrac{1}{8}(x^2-y^2)e^{-x}\,dy$

$= \frac{1}{6} e^{-x} x^3 \qquad 0 \le x \le \infty$

If $y>0$



$f_Y(y) = \int_y^{\infty} \tfrac{1}{8}(x^2-y^2)e^{-x}\,dx \qquad y>0$

$= \frac{1}{4} e^{-y}(y+1) \qquad y>0$

If $y<0$

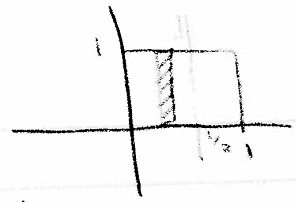
$f_Y(y) = \int_{-y}^{\infty} \frac{1}{8}(x^2-y^2)e^{-x}\,dx$

$= \frac{1}{4} e^{y}(1-y) \qquad y<0$

33

#8 (cont)

a (iii)



$$P(X < \tfrac{1}{2}) = \int_0^{1/2}\int_0^1 \tfrac{6}{7}(x+y)^2\,dy\,dx$$

$$= 2/7$$

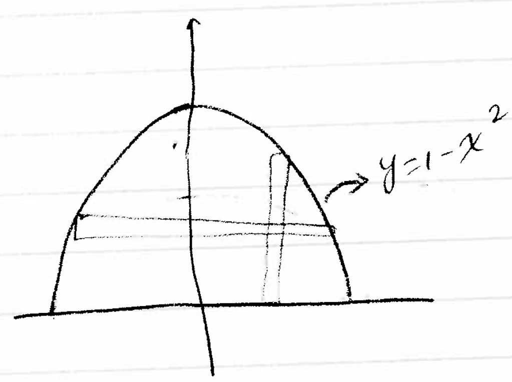
8b

$$f_X(x) = \int_0^1 \tfrac{6}{7}(x+y)^2\,dy =$$

$$= \tfrac{1}{7}(6x^2 + 6x + 2) \qquad 0 \le x \le 1$$

$$f_Y(y) = \tfrac{1}{7}(2 + 6y + 6y^2) \qquad 0 \le y \le 1$$

#9



$$f_{XY}(x,y) = \begin{cases} c & 0 \le y \le 1-x^2 \\ & -1 \le x \le 1 \\ 0 & \text{o.w.} \end{cases}$$

$$\Rightarrow c = 1 / \left\{ \int_{-1}^{1}\int_0^{1-x^2} dy\,dx \right. \quad \Rightarrow c = 3/4$$

$$f_X(x) = \int_0^{1-x^2} 3/4\,dy = 3/4\,(1-x^2) \qquad -1 \le x \le 1$$

$$f_Y(y) = \int_{-\sqrt{1-y}}^{\sqrt{1-y}} 3/4\,dy = 3/2\sqrt{1-y} \qquad 0 \le y \le 1$$

32

Problem 6 (Cont.)

$$f_X(x) = \int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} \frac{1}{\pi ab}\, dy$$

$$= \frac{1}{\pi ab}\left\{2\frac{b}{a}\sqrt{a^2-x^2}\right\}$$

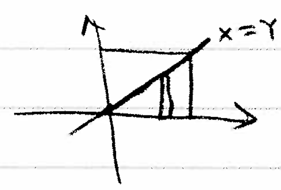
$$= \frac{2}{\pi a^2}\sqrt{a^2-x^2} \qquad -a \le x \le a$$

$f_Y(y)$ is obtained similarly

Problem 8

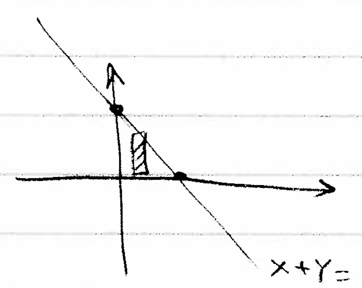
a (i)

$$P(X>Y) = \int_0^1 \int_0^x \frac{6}{7}(x+y)^2\, dy\, dx$$



$$= \frac{1}{2}$$

a (ii)



$$P(X+Y \le 1)$$

$$= \int_0^1 \int_0^{1-x} \frac{6}{7}(x+y)^2\, dy\, dx$$

$$= 3/14$$

(31)

Problem 3:

$n = 10$ $\quad p_i$ = probability that the i^{th} player wins

$p_i = \frac{1}{3}$ $\quad i = 1, 2, 3$

Let N_i = number of games won by the i^{th} player

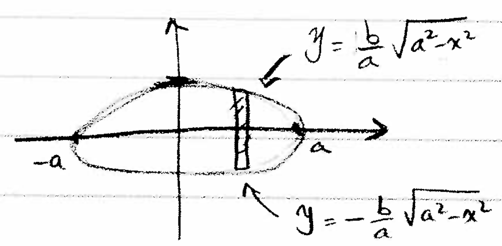
$$P(N_1 = n_1,\ N_2 = n_2,\ N_3 = n_3) = \binom{10}{n_1, n_2, n_3} \left(\frac{1}{3}\right)^{n_1} \left(\frac{1}{3}\right)^{n_2} \left(\frac{1}{3}\right)^{n_3}$$

with $n_1 + n_2 + n_3 = 10$

Problem 6:

w log assume $a, b > 0$.

$$f_{X,Y}(x,y) = \begin{cases} c & -\frac{b}{a}\sqrt{a^2-x^2} \le y \le \frac{b}{a}\sqrt{a^2-x^2} \\ & -a \le x \le a \\ 0 & \text{o.w.} \end{cases}$$



where c is a constant. we need to find c first

c satisfies

$$\int_{-a}^{a} \int_{-\frac{b}{a}\sqrt{a^2-x^2}}^{\frac{b}{a}\sqrt{a^2-x^2}} c \, dy \, dx = 1 \quad \Rightarrow$$

$$\int_{-a}^{a} 2c \frac{b}{a} \sqrt{a^2-x^2} \, dx = 1 \quad \Rightarrow$$

$$\frac{bc}{a} \left\{ x\sqrt{a^2-x^2} + a^2 \operatorname{atan}\left(\frac{x}{\sqrt{a^2-x^2}}\right) \right\} \Big|_{-a}^{a} = \frac{bc}{a} \left\{ \pi a^2 \right\}$$

$\Rightarrow abc\pi = 1 \Rightarrow c = \frac{1}{\pi ab}$

30

Chapter 3

Problem 1:

a. $p_X(x) = \sum_{y=1}^{4} p_{X,Y}(x,y)$

x	1	2	3	4
$P_X(x)$	0.19	.32	0.31	0.18

$p_Y(y) = \sum_{x=1}^{4} p_{X,Y}(x,y)$

y	1	2	3	4
$P_Y(y)$	0.19	0.32	0.31	0.18

b. $P_{X|Y=1}(x) = \frac{P(X=x, Y=1)}{P(Y=1)}$

x	1	2	3	4
$P_{X\|Y=1}$	$\frac{0.1}{0.19}$	$\frac{0.05}{.32}$	$\frac{0.02}{.31}$	$\frac{.02}{.18}$

$P_{Y|X=1}(y) = \frac{P(Y=y, X=1)}{P(X=1)}$

y	1	2	3	4
$P_{Y\|X=1}(y)$	$\frac{0.1}{0.19}$	$\frac{.05}{.32}$	$\frac{.02}{.31}$	$\frac{.02}{.18}$

3.5.4 key solutions for chapter 4

1

§4.1 & 4.2

#1

$$\int_{-\infty}^{\infty} |x| f(x)\,dx \le M \int_{-\infty}^{\infty} f(x)\,dx \le M$$

$\therefore$ $E(x)$ exists.

#2

x	1	2	...	n
$p(x)$	$\frac{1}{n}$	$\frac{1}{n}$		$\frac{1}{n}$

$$E(x) = \sum_{k=1}^{n} k\left(\tfrac{1}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} k = \frac{1}{n}\,\frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$Var(x) = E(x^2) - (E(x))^2$$

$$E(x^2) = \sum_{k=1}^{n} k^2\left(\tfrac{1}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{\cancel{n}}\,\frac{\cancel{n}(2n+1)(n+1)}{6} = \frac{(2n+1)(n+1)}{6}$$

$$Var(x) = \frac{(2n+1)(n+1)}{6} - \frac{(n+1)^2}{4}$$

#6 $f(x) = 2x \quad 0 \le x \le 1$

a) $E(x) = \int_0^1 2x^2\,dx = 2\frac{x^3}{3}\Big|_0^1 = \frac{2}{3}$

b) $F_Y(y) = P(Y \le y) = P(x^2 \le y) = P(0 \le x \le \sqrt{y}) = \int_0^{\sqrt{y}} 2x\,dx$

$= x^2\Big|_0^{\sqrt{y}} = y$

$f_Y(y) = 1 \quad 0 \le y \le 1$

$E(Y) = \int_0^1 1\cdot y\,dy = \frac{y^2}{2}\Big|_0^1 = \frac{1}{2}$

②

c) $E(X^2) = \int_0^1 x^2 (2x)\,dx = 2\int_0^1 x^3 dx = 2\left.\frac{x^4}{4}\right|_0^1 = \frac{1}{2}$

d) $Var(X) = \int_0^1 (x-\frac{2}{3})^2 (2x)\,dx$

$= \int_0^1 (x^2 - \frac{4}{3}x + \frac{4}{9})(2x)dx$

$= \int_0^1 (2x^3 - \frac{8}{3}x^2 + \frac{8}{9}x)\,dx$

$= \left.\frac{2x^4}{4} - \frac{8}{3}\frac{x^3}{3} + \frac{8}{9}\frac{x^2}{2}\right|_0^1$

$= \frac{1}{2} - \frac{8}{9} + \frac{4}{9} = \frac{7}{18}$

or

$Var(X) = E(X^2) - (E(X))^2$

$= \frac{1}{2} - (\frac{2}{3})^2 = \frac{7}{18}$

#8 $\sum_{k=1}^{\infty} P(X \geq k) = P(X\geq 1) + P(X\geq 2) + P(X\geq 3) + \ldots$

$P(X\geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + \ldots$

$P(X\geq 2) = P(X=2) + P(X=3) + P(X=4) + \ldots$

$P(X\geq 3) = P(X=3) + P(X=4) + \ldots$

$P(X\geq 4) = P(X=4) + \ldots$

$\sum_{k=1}^{\infty} P(X\geq k) = P(X=1) + 2P(X=2) + 3P(X=3) + 4P(X=4) + \ldots$

$= \sum_{x=1}^{\infty} x\, P(X=x) = E(X)$

4.1 & 4.2

3

#8 Cont

$X \sim Geom(p)$

$P(X=k) = (1-p)^{k-1} p \qquad k=1,2,\ldots$

$P(X \le k) = p\sum_{j=1}^{k} (1-p)^{k-1} = p \cdot \frac{1-(1-p)^k}{p} = 1-(1-p)^k$

$P(X \ge k) = 1 - P(X \le k-1) = (1-p)^{k-1}$

$E(X) = \sum_{k=1}^{\infty} (1-p)^{k-1} = \frac{1}{p}$

#10

Let X = # of items searched to reach the desired item. Let $\otimes$ be the desired item, then possible arrangements are:

$\otimes \; \times \; \times \; - \; - \; \times \qquad X=1$

$\times \; \otimes \; \times \; - \; - \; \times \qquad X=2$

$\times \; \times \; \otimes \; - \; - \; \times \qquad X=3$

$\vdots$

$\times \; \times \; \times \; - \; - \; \otimes \qquad X=n$

$P(X=i) = \frac{1}{n}$

$E(X) = \sum_{k=1}^{n} k \frac{1}{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$

4.1 & 4.2

4

#12

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

$$= \int_{-\infty}^{\infty} (x-\xi) f(x)\, dx + \int_{-\infty}^{\infty} \xi f(x)\, dx$$

$$= \int_{-\infty}^{\infty} (x-\xi) f(x)\, dx + \xi$$

But $\int_{-\infty}^{\infty} (x-\xi) f(x)dx = \int_{-\infty}^{\infty} u\, f(u+\xi)\, du = 0 \quad u = x-\xi$

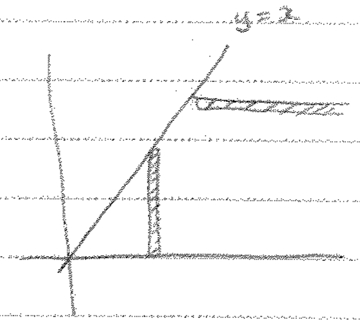
since $g(u) = u\, f(u+\xi)$ is an odd function

Note $g(-u) = -u\, f(-u+\xi) = -u f(u+\xi) = -g(u)$

$\therefore\ E(X) = \xi$

#13

$$E(X) = \int_0^{\infty} x\, f(x)\, dx$$



$$= \int_0^{\infty} \int_0^{x} 1\, dy\, f(x)\, dx$$

$$= \int_0^{\infty} \int_0^{x} f(x)\, dy\, dx$$

$$= \int_0^{\infty} \int_y^{\infty} f(x)\, dx\, dy$$

$$= \int_0^{\infty} (1-F(y))\, dy \quad ▨$$

$X \sim \exp(\lambda) \qquad F(x) = 1 - e^{-\lambda x}$

$$E(X) = \int_0^{\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}.$$

4.1 & 4.2

5

#16

$$E(Z) = E\left(\frac{Z-\mu}{\sigma}\right) = \frac{1}{\sigma} E(Z-\mu)$$

$$= \frac{1}{\sigma}\{E(Z) - E(\mu)\} = \frac{1}{\sigma}\{\mu-\mu\} = 0$$

$$Var(Z) = Var\left(\frac{Z-\mu}{\sigma}\right) = \frac{1}{\sigma^2} Var(Z-\mu)$$

$$= \frac{1}{\sigma^2}\{Var(Z)\} = \frac{1}{\sigma^2}\cdot\sigma^2 = 1$$

#18 $F_{U_{(n)}}(u) = u^n \qquad u \in [0,1]$

$F_{U_{(1)}}(v) = 1 - [1-v]^n \qquad v \in [0,1]$

Using the result of problem 13:

$$E(U_{(n)}) = \int_0^1 (1-u^n)\,du = u - \frac{u^{n+1}}{n+1}\Big|_0^1 = 1 - \frac{1}{n+1}$$

$$E(U_{(1)}) = \int_0^1 [1-v]^n\,dv = -\frac{(1-v)^{n+1}}{n+1}\Big|_0^1 = \frac{1}{n+1}$$

$$E\{U_{(n)} - U_{(1)}\} = E(U_{(n)}) - E(U_{(1)}) = 1 - \frac{2}{n+1} = \frac{n-1}{n+1}$$

(6)

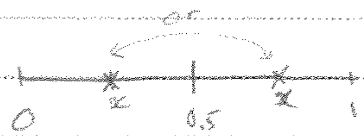
§4.[redacted] B²
#20

Let R = ration of larger piece to the smaller piece

and let

X = the place where the stick is cut

$X \sim Unif(0,1)$



$$E(R) = E(R \mid X \le 0.5)\, P(X \le 0.5) + E(R \mid X \ge 0.5)\, P(X > 0.5)$$

$$= \frac{1}{2} E\left(\frac{1-X}{X} \,\middle|\, X \le 0.5\right) + \frac{1}{2} E\left(\frac{X}{1-X} \,\middle|\, X \ge 0.5\right)$$

$$= \frac{1}{2} \left\{ \int_0^{0.5} 2\, \frac{1-x}{x}\, dx + \int_{0.5}^{1} 2\, \frac{x}{1-x}\, dx \right\}$$

note that this integral is not finite, so the ~~expectation does not exist~~

Note that $X \mid X \le 0.5 \sim Unif(0, 0.5)$

However let

T = Ratio of smaller to larger piece. Then

$$E(T) = E(T \mid X \le 0.5)\, P(X \le 0.5) + E(T \mid X \ge 0.5)\, P(X \ge 0.5)$$

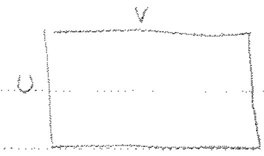
$$= \frac{1}{2} \left\{ E\left(\frac{X}{1-X} \,\middle|\, X \le 0.5\right) + E\left(\frac{1-X}{X} \,\middle|\, X \ge 0.5\right) \right\}$$

$$= \frac{1}{2} \left\{ \int_0^{0.5} \frac{2\cancel{X}}{1-X}\, dx + \int_{0.5}^{1} 2\left(\frac{1-X}{X}\right) dx \right\} = 2\log(2) - 1$$

(7)

§4.1

#22



Area = UV $U \sim \text{Unif}(0,1)$
$V \sim \text{Unif}(0,1)$
$U \perp\!\!\!\perp V$

$$E(\text{Area}) = \int_0^1 \int_0^1 uv \, du \, dv$$

$$= \int_0^1 v \left. \frac{u^2}{2} \right|_0^1 dv = \frac{1}{2} \int_0^1 v \, dv = \frac{1}{4}$$

#25

$$E(R^2) = E(X_1^2 + X_2^2)$$

$$= E(X_1^2) + E(X_2^2)$$

$$E(X_1^2) = \int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha+1} e^{-\lambda x} \, dx$$

Note that using the density of $\Gamma(\alpha+2, \lambda)$

$$\int_0^\infty x^{\alpha+1} e^{-\lambda x} \, dx = \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}}$$

$$\therefore \ E(X_1^2) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}} = \frac{(\alpha+1)\alpha}{\lambda^2}$$

$$E(R^2) = \frac{2(\alpha+1)\alpha}{\lambda^2}$$

8

§4.1

#26

$$E(X_1+X_2+\cdots+X_r) = \sum_{j=1}^{r} \frac{n}{n-j+1}$$

$$= \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{n-r+1}$$

$$= n \sum_{j=1}^{r} \frac{1}{n-j+1}$$

#30 $$E\left(\frac{1}{X+1}\right) = \sum_{x=0}^{\infty} \frac{1}{x+1} \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \frac{1}{\lambda} \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!}$$

$$= \frac{1}{\lambda} \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \frac{1}{\lambda}\left[\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!} - e^{-\lambda}\right]$$

$$= \frac{1}{\lambda}\left[1 - e^{-\lambda}\right]$$

#32 $$E\left(\frac{1}{X}\right) = \int_0^{\infty} \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \underbrace{x^{\alpha-2} e^{-\lambda x}}_{\text{Compare to } \Gamma(\alpha-1,\lambda)} dx$$

$$= \frac{\lambda^{\alpha}}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha-1)}{\lambda^{\alpha-1}} = \frac{\lambda}{\alpha-1}.$$

7

§4.1, 4.2

#35

Let X_1 = # of trials until 1^{st} success

X_2 = # of trials after the first success to the 2^{nd} success

X_i = " " " " $i-1$ " " " i^{th} "

X_r = " " " " $r-1$ " " " r^{th} "

$\underbrace{FFF\ldots FS}_{X_1}\ \underbrace{FF\ldots F\ldots FS}_{X_2}\ \ldots\ldots S\ \underbrace{FF\ldots FS}_{X_r}$

Then $X = X_1 + X_2 + \ldots + X_r \sim$ negative binomial (p, r)

Note that $X_i \sim$ Geom(p) and $E(X_i) = \frac{1}{p}$

$$E(X) = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}.$$

§4.2

10

#2 $X \sim exp(\frac{1}{\sigma})$ since $Var(X) = \sigma^2 = \frac{1}{\lambda^2}$

$E(X) = \frac{1}{\lambda} = \sigma$

$P(|X - E(X)| > k\sigma) = P(|X - \sigma| > k\sigma)$

$= 1 - P(|X - \sigma| \le k\sigma)$

$= 1 - P(-k\sigma < (X - \sigma) < k\sigma)$

$= 1 - P((1-k)\sigma < X < (1+k)\sigma)$

$= 1 - P(0 < X < (1+k)\sigma)$ for $k \ge 1$

$= 1 - F_X((1+k)\sigma)$

$= 1 - (1 - e^{-(1+k)\sigma/\sigma})$

$= e^{-(1+k)}$

k	2	3	4
$e^{-(1+k)}$	.0498	.0183	.0067
$1/k^2$	0.25	.111	.0625

The bounds are very large relative to the exact value.

11

§4.1 §4.2

#49 $E(X) = E(Y) = \mu \qquad \sigma_X \neq \sigma_Y$

$$Z = \alpha X + (1-\alpha) Y \qquad 0 \le \alpha \le 1$$

a. $E(Z) = \alpha E(X) + (1-\alpha) E(Y)$
$= \alpha \mu + (1-\alpha)\mu = \mu$

b. $Var(Z) = \alpha^2 \sigma_X^2 + (1-\alpha)^2 \sigma_Y^2$

$$\frac{\partial}{\partial \alpha} Var(Z) = 2\alpha \sigma_X^2 - 2(1-\alpha)\sigma_Y^2 = 0$$

$$\alpha(2\sigma_X^2 + 2\sigma_Y^2) = 2\sigma_Y^2$$

$$\alpha = \frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}$$

Note:
If $\sigma_X = \sigma_Y$ then $\alpha = \frac{1}{2}$ is optimal.

If $\sigma_Y > \sigma_X$ more weight should be given to X than Y and vise versa

c. Now $Var\left(\frac{X+Y}{2}\right) = \frac{1}{4}(\sigma_X^2 + \sigma_Y^2)$

Thus $\frac{X+Y}{2}$ is only advantagous if either
(1) $\sigma_X^2 + \sigma_Y^2 \le 4\sigma_X^2$ or $\sigma_X^2 + \sigma_Y^2 \le 4\sigma_Y^2$

§4.2

12

#50 $E(X_i)=\mu \quad Var(X_i)=\sigma^2$

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}(n\mu) = \mu$$

$$Var(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

#57 $Var(XY) = E(X^2Y^2) - (EXY)^2$

$$= (EX^2)(EY^2) - (EX)^2 E(Y)^2$$

$$= (\sigma_X^2+\mu_X^2)(\sigma_Y^2+\mu_Y^2) - \mu_X^2\mu_Y^2$$

13

§4.3

#43

$$\mathrm{Var}(X-Y) = E(X-Y)^2 - [E(X-Y)]^2$$
$$= E(X^2 - 2XY + Y^2) - [E(X) - E(Y)]^2$$
$$= \underbrace{E(X^2)}_{①} - \underbrace{2E(XY)}_{③} + \underbrace{E(Y^2)}_{②} - \underbrace{(E(X))^2}_{①} + \underbrace{2E(X)E(Y)}_{③} - \underbrace{(E(Y))^2}_{②}$$
$$= \underbrace{\mathrm{Var}(X)}_{①} + \underbrace{\mathrm{Var}(Y)}_{②} - \underbrace{2\,\mathrm{Cov}(X,Y)}_{③}$$

#45 Define

$$I_{ij} = \begin{cases} 1 & \text{if the } i^{th} \text{ trial results in the } j^{th} \text{ outcome} \quad i=1,\dots,n \quad j=1,\dots,r \\ 0 & \text{o.w.} \end{cases}$$

$i=1,\dots,n$ and $I_{il} \perp\!\!\!\perp I_{jk}$ for $i \neq j$

$\therefore\ \mathrm{Cov}(I_{il}, I_{jk}) = 0$ for $i \neq j$.

$$N_i = \sum_{l=1}^{n} I_{li} \qquad N_j = \sum_{l=1}^{n} I_{lj}$$

$$\mathrm{Cov}(N_i, N_j) = \mathrm{Cov}\left(\sum_{l=1}^{n} I_{li}, \sum_{k=1}^{n} I_{kj}\right)$$
$$= \sum_{l=1}^{n}\sum_{k=1}^{n} \mathrm{Cov}(I_{li}, I_{kj}) = \sum_{k=1}^{n} \mathrm{Cov}(I_{ki}, I_{kj})$$

$P(I_{ki} I_{kj} = 1) = 0$, $P(I_{ki} I_{kj} = 0) = 1$, $P(I_{ki} = 1) = p_i$

$$\mathrm{Cov}(I_{ki}, I_{kj}) = E(I_{ki} I_{kj}) - E(I_{ki})\,E(I_{kj})$$
$$= 0 - p_i p_j \implies \mathrm{Cov}(N_i, N_j) = -n p_i p_j$$

§4.3

14

45 Continued

$N_i \sim \text{Binomial}(n, p_i) \Rightarrow Var(N_i) = n p_i (1-p_i)$

$$\rho = \frac{-n p_i p_j}{\sqrt{n^2 p_i p_j (1-p_i)(1-p_j)}} = \frac{-\sqrt{p_i p_j}}{\sqrt{(1-p_i)(1-p_j)}}$$

46

$Cov(U, V) = bd \; Cov(X, Y)$

$Var(U) = b^2 \; Var(X)$

$Var(V) = d^2 \; Var(Y)$

$$\rho_{UV} = \frac{Cov(U, V)}{\sqrt{Var(U)\,Var(V)}} = \frac{bd \; Cov(X, Y)}{\sqrt{b^2 d^2 \; Var(X)\,Var(Y)}}$$

$$= \frac{bd}{|bd|} \frac{Cov(X, Y)}{\sqrt{Var(X)\,Var(Y)}} = \frac{bd}{|bd|} \rho_{XY}$$

$$= \text{sign}(bd) \, \rho_{XY}$$

$\Rightarrow |\rho_{UV}| = |\rho_{XY}|$

(15)

#60 $P(Y \geq y) = P(Y \leq -y)$

$$S = \begin{cases} 1 & \text{w prob } \frac{1}{2} \\ -1 & \text{w prob } \frac{1}{2} \end{cases}$$

$X = SY \qquad S \perp Y$

$Cov(X, Y) = E(XY) - E(X)\,E(Y)$

$= E(XY) \qquad$ since $E(Y) = 0$

$= E(SY^2)$

$\overset{\text{by independence}}{=} E(S)\,E(Y^2) = 0 \qquad$ since $E(S) = 0$

To show that X and Y are not necessarily independent it is sufficient to take an example

Let

X	-1	0	1
$p(x)$	$1/3$	$1/3$	$1/3$

Y \ S	1	-1
-1	-1	1
0	0	0
1	1	-1

$\underbrace{\qquad\qquad}_{X}$

$P(X=0 \mid Y=1) = 0$

But

$P(X=0) \neq 0$

So X and Y are not independent

162

§4.4

#63

$$f(x,y) = \begin{cases} \frac{6}{7}(x+y)^2 & 0 \le x \le 1 \quad 0 \le y \le 1 \\ \end{cases}$$

Recall: $f_X(x) = \frac{1}{7}(6x^2+6x+2) \quad 0 \le x \le 1$

$f_Y(y) = \frac{1}{7}(6y^2+6y+2) \quad 0 \le y \le 1$

$$f_{Y|X}(y|x) = \frac{(6/7)(x+y)^2}{\frac{1}{7}(6x^2+6x+2)} = \frac{6(x+y)^2}{6x^2+6x+2} \quad 0 \le y \le 1$$

a) $E(X) = \int_0^1 \frac{x}{7}(6x^2+6x+2)\,dx = 9/14 \Rightarrow E(Y) = 9/14$

$E(X^2) = \int_0^1 \frac{x^2}{7}(6x^2+6x+2)\,dx = \frac{101}{210} \Rightarrow E(Y^2) = \frac{101}{210}$

$E(XY) = \int_0^1 \int_0^1 XY\left(\frac{6}{7}\right)(x+y)^2\,dx\,dy = \frac{17}{42}$

$Cov(X,Y) = \frac{17}{42} - \frac{9}{14} \cdot \frac{9}{14} = -.0085$

$Var(Y) = Var(X) = \frac{101}{210} - \left(\frac{9}{14}\right)^2 = .0677$

$$\rho = \frac{-.0085}{.0677} = -.1256$$

b)

$$E(Y|X) = \int_0^1 y \frac{6(x+y)^2}{6x^2+6x+2}\,dy = \frac{(6x^2+8x+3)}{4(3x^2+3x+1)}$$

§4.4

#66

$X =$ waiting time

$$Y = \begin{cases} 1 & \text{if fast elevator is selected} \\ 0 & \text{if slow " " "} \end{cases}$$

$$E(X) = E\{E(X|Y)\}$$

$$= E(X|Y=0)\,P(Y=0) + E(X|Y=1)\,P(Y=1)$$

$$= 3\left(\tfrac{1}{3}\right) + 1\left(\tfrac{2}{3}\right) = 5/3$$

#67 Let $X =$ base $\sim$ Unif(0,1)

$Y =$ height

$Y|X \sim$ unif$(0, X)$

$$\text{Area} = \frac{XY}{2}$$

$$E\left(\frac{XY}{2}\right) = E\left\{E\left(\frac{XY}{2}\Big|X\right)\right\}$$

$$= E\{X\ E(Y|X)\}$$

$$= E\left\{\frac{X}{2}\cdot\frac{X}{2}\right\} = \frac{1}{4}EX^2$$

$$EX^2 = \int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$$

$E(\text{Area}) = \frac{1}{6}$.

§4.4

#67 cont.

Let C be the circumference.

$$C = 2(X+Y)$$

$$E(C) = E\{E(C|X)\}$$
$$= 2E\{E(X+Y|X)\}$$
$$= 2E\{E(X|X)\} + E\{E(Y|X)\}$$
$$= 2\left(E(X) + E\left(\tfrac{X}{2}\right)\right)$$
$$= 2\left(\tfrac{1}{2} + \tfrac{1}{4}\right) = 1.5$$

#68

$$Var(Y) = Var(E(Y|X)) + E(Var(Y|X))$$

Since $Var(E(Y|X)) \geq 0$ the result is obvious.

#70 $$E(X|Y=y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$$
$$= \int_{-\infty}^{\infty} x f_X(x)\,dx = E(X)$$

#73

$N \sim \text{Binomial}(n, p)$

$Y = \#$ of heads obtained the second time

$Y|N \sim \text{Binomial}(N, p)$

$$E(Y+N) = E\{E(Y+N|N)\}$$

$$= E\{E(Y|N)\} + E(N)$$

$$= E\{Np\} + E(N)$$

$$= np^2 + np$$

Alter native method:

let $X_i = \begin{cases} 1 & \text{if heads is obtained the } 2^{nd} \text{ round} \\ 0 & \text{o.w.} \end{cases}$

$$E\left(N + \sum_{i=1}^{N} X_i\right) = E(N) + E\left(E\left(\sum_{i=1}^{N} X_i \,|\, N\right)\right)$$

$$= E(N) + E\left\{\sum_{i=1}^{N} E(X_i|N)\right\}$$

$$= np + E\left\{\sum_{i=1}^{N} p\right\}$$

$$= np + E(Np) = np + np^2$$

§4.4

#26 $T \sim \exp(\lambda)$

$U|T \sim \text{uniform}(0,T)$

$$E(U) = E\{E(U|T)\}$$

$$= E\left\{\frac{T}{2}\right\} = \frac{1}{2\lambda}$$

$$Var(U) = Var\{E(U|T)\} + E\{var(U|T)\}$$

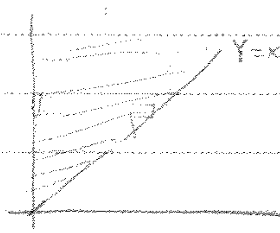
$$= Var\left(\frac{T}{2}\right) + E\left(\frac{T^2}{12}\right)$$

$$= \frac{1}{4} Var(T) + \frac{1}{12}\{var(T) + (E(T))^2\}$$

$$= \frac{1}{4\lambda^2} + \frac{1}{12}\left\{\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\right\}$$

$$= \frac{5}{12\lambda^2}$$

#77



a) $Cov(X,Y) = E(XY) - E(X)E(Y)$

$f_X(x) = \int_x^\infty e^{-y}\,dy = -e^{-y}\Big|_x^\infty = e^{-x}$ $x>0$

$f_Y(y) = \int_0^y e^{-y}\,dx = ye^{-y}$

$X \sim exp(1)$ $Y \sim \Gamma(\alpha=2, \lambda=1)$

$E(X)=1$ $E(Y) = \alpha/\lambda = 2$

$E(XY) = \int_0^\infty \int_0^y xy\, e^{-y}\, dx\, dy$

$= \int_0^\infty y e^{-y} \frac{x^2}{2}\Big|_0^y dy$

$= \frac{1}{2}\int_0^\infty y^3 e^{-y}\,dy$ Compare to $\Gamma(\alpha=4, \lambda=1$

$= \frac{1}{2} \cdot \frac{\Gamma(4)}{1} = 3$

$Cov(X,Y) = 3-2 = 1$

7b (cont)

$$f_{X|Y}(x|y) = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y} \qquad 0 < x < y$$

$$E(X|Y=y) = \int_0^y x/y \, dx = \frac{1}{y} \left. \frac{x^2}{2} \right|_0^y = \frac{y}{2}$$

$$f_{Y|X}(y|x) = \frac{e^{-y}}{e^{-x}} \qquad x < y < \infty$$

$$E(Y|X=x) = \int_x^\infty y \, e^{-y}/e^{-x} \, dy$$

$$= e^{x} \int_x^\infty y e^{-y} \, dy$$

$$= e^{x} \, e^{-x}(x+1) = x+1$$

7c)

$$E(X|Y) = \frac{Y}{2}$$

Let $Z = \frac{Y}{2}$ $\qquad F_Z(z) = P\left(\frac{Y}{2} \le z\right)$

$$= P(Y \le 2z)$$

$$= F_Y(2z)$$

$$f_Z(z) = 2 f_Y(2z) = 2\{2z e^{-2z}\} = 4z e^{-2z} \qquad z \ge 0$$

77c cont

$E(Y|X) = x+1$

$X \sim \exp(1)$ $\quad f_X(x) = e^{-x}$

$Z = x+1$

$f_Z(z) = e^{z-1}$ $\quad z \geq 1$

§4.5

#83 $X_i \sim \text{Binomial}(n_i, p)$

$$Z = \sum_{i=1}^{n} X_i$$

$$M_Z(t) = E\left(e^{\sum X_i t}\right)$$

$$= \prod_{i=1}^{n} E e^{X_i t}$$

$$= \prod_{i=1}^{n} M_{X_i}(t)$$

$$= \prod_{i=1}^{n} (pe^t + 1 - p)^{n_i}$$

$$= (pe^t + 1 - p)^{\sum n_i}$$

$Z \sim \text{Binomial}(\sum n_i, p)$

#84

$$M_Z(t) = \prod_{i=1}^{n} (p_i e^t + 1 - p_i)^{n_i}$$

And there is no way to have this in the form of a binomial mgf.

#85

$$M_x(t) = E(e^{tx}) = \sum_{x=1}^{\infty} e^{tx} (1-p)^{x-1} p$$

$$= pe^t \sum_{x=1}^{\infty} \left((1-p)e^t\right)^{x-1}$$

$$= \frac{pe^t}{1-(1-p)e^t} \sum_{x=1}^{\infty} [(1-p)e^t]^{x-1} (1-(1-p)e^t)$$

$$= \frac{pe^t}{1-(1-p)e^t}$$

$$M_x'(t) = \frac{pe^t}{(1-e^t+pe^t)^2}$$

$$M_x'(0) = \frac{p}{p^2} = \frac{1}{p} = E(X)$$

$$M_x''(t) = \frac{-pe^t(pe^t - e^t - 1)}{(1-e^t+pe^t)^3}$$

$$M_x''(0) = \frac{p(2-p)}{p^3} = \frac{2-p}{p^2}$$

$$Var(X) = \frac{(2-p)}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

#90

$$M_X(t) = \exp\left\{\frac{\sigma^2 t^2}{2}\right\}$$

$$M_X'(t) = \cancel{t}\,t\sigma^2 M_X(t)$$

$$M_X''(t) = \sigma^2\left\{M_X(t) + tM_X'(t)\right\}$$

$$M_X'''(t) = \sigma^2\left\{2M_X'(t) + tM_X''(t)\right\}$$

$$M^{IV}(t) = \sigma^2\left\{3M_X''(t) + tM_X'''(t)\right\}$$

In general it can be shown that

$$M_X^{(r)}(t) = \sigma^2\left\{(r-1)M_X^{(r-2)}(t) + tM^{(r-1)}(t)\right\}$$

If r is odd, $r-2$ is odd and hence $M_X^{(r-2)}(0) = 0$ so it is easy to see that

$$M_X^{(r)}(0) = 0 \text{ for } r \text{ odd.}$$

For even r, we prove by induction.

Assume $\mu_{2n} = \{(2n)!\,\sigma^{2n}\}/\{2^n(n!)\}$

$$\mu_{2n+2} = M_X^{(2n+2)}(0) = \sigma^2\left\{(2n+1)M_X^{(2n)}(0) + 0\,M^{(2n+1)}(t)\right\}$$

$$= \frac{\sigma^2(2n+1)(2n)!\,\sigma^{2n}}{2^n n!} = \frac{(2n+2)!\,\sigma^{2n+2}}{(2n+2)2^n n!} = \frac{(2n+2)!\,\sigma^{2n+2}}{2^{n+1}(n+1)!}$$

#92 $\Theta \sim \Gamma(\lambda, \alpha)$ α integer

$X|\Theta \sim \text{Poisson}(\Theta)$

let $Z = \alpha + X$

$$M_Z(t) = e^{\alpha t} M_X(t)$$

$$= e^{\alpha t} E(e^{tX})$$

$$= e^{\alpha t} E\{\underbrace{E(e^{tX} \mid \Theta)}\}$$

$$= e^{\alpha t} E\{e^{\Theta(e^t-1)}\}$$ Since $X|\Theta$ is Poisson

$$= e^{\alpha t} \int_0^\infty e^{\Theta(e^t-1)} \frac{\lambda^\alpha \Theta^{\alpha-1} e^{-\lambda\Theta}}{\Gamma(\alpha)} d\Theta$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{\alpha t} \int_0^\infty \Theta^{\alpha-1} e^{\Theta(e^t-\lambda-1)} d\Theta$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{\alpha t} \frac{\Gamma(\alpha)}{(1+\lambda-e^t)^\alpha} = \frac{\lambda^\alpha e^{\alpha t}}{(1+\lambda-e^t)^\alpha}$$

$$= \left(\frac{\lambda e^t}{1+\lambda-e^t}\right)^\alpha$$

$$M_Y(t) = \left(\frac{\frac{\lambda}{\lambda+1} e^t}{1-\frac{1}{1+\lambda}e^t}\right)^\alpha$$ $Y \sim$ negative Binomial

#92 cont

Let $X_1, \ldots, X_n \sim \text{geom}(p)$ iid

$$X = \sum_{i=1}^{r} X_i \quad \text{Negative binomial } (r, p)$$

$$M_X(t) = \prod_{i=1}^{r} M_{X_i}(t)$$

$$= \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$$

$Z \sim$ negative binomial with

$p = \frac{\lambda}{1+\lambda}$ and $r = \alpha$

#96

$$M_{XY}(s,t) = E\left(e^{tY+sX}\right)$$

$$\frac{\partial}{\partial t}\left(\frac{\partial}{\partial s} M_{XY}(s,t)\right) = E\left(XY e^{tY+sX}\right)$$

$$E(XY) = \frac{\partial}{\partial t}\left(\frac{\partial}{\partial s} M_{XY}(0,0)\right)$$

#101 $X \sim \text{Poisson}(\lambda)$

$Y = \sqrt{X}$

$g(x) = \sqrt{x}$ $\quad g'(x) = \frac{1}{2} x^{-1/2}$ $\quad g''(x) = -\frac{1}{4} x^{-3/2}$

$$E(Y) \approx g(\mu_X) + \frac{1}{2}\sigma_X^2 g''(\mu_X)$$

$$= \sqrt{\lambda} + \frac{\lambda}{2}\left(-\frac{1}{4}\lambda^{-3/2}\right)$$

$$= \sqrt{\lambda} - \frac{1}{8\sqrt{\lambda}}$$

$$Var(Y) \approx \sigma_X^2 \left[g'(\mu_X)\right]^2$$

$$= \lambda \left[\frac{1}{2}\lambda^{-1/2}\right]^2 = \frac{1}{2}$$

#102

$$E(\Theta) = \tan^{-1}(y_0/x_0)$$

$$Var(\Theta) = \sigma^2/(x_0^2 + y_0^2)$$

3.5.5 key solutions for chapter 5

Chapter 5

#1 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \Rightarrow$ ~~$Var(\bar{X}_n) = \frac{1}{n^2} Var(X_i) = \frac{\sigma^2}{n}$~~

$Var(\bar{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2$

By Chebyshev

$$P(|\bar{X}_n - \mu| > \varepsilon) \le \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sum_{i=1}^{n}\sigma_i^2 / n^2}{\varepsilon^2} \to 0 \text{ as } n \to \infty$$

#3 $N \sim \text{Poisson}(10{,}000) \approx N(10{,}000,\ 10{,}000)$

$$P(N > 10{,}200) \cong P\left(\frac{N - 10{,}000}{\sqrt{10{,}000}} > \frac{10{,}200 - 10{,}000}{100}\right)$$

$$= P(Z > 2) \cong .028$$

#5 $n \to \infty \quad p \to 0 \quad np \to \lambda$

$X \sim \text{Binomial}(n, p) \to \text{Poisson}(\lambda)$

$M_X(t) = (pe^t + 1 - p)^n$

$\log M_X(t) = n \log(1 + p(e^t - 1))$

$= n\{p(e^t - 1) + O(p^2(e^t - 1))\}$ by Taylor expansion

$= np(e^t - 1) + \underbrace{n\,O(p^2(e^t - 1))}_{\to 0}$

$\to \lambda(e^t - 1)$ Note that if $Y \sim \text{Poisson}(\lambda)$

then $M_Y(t) = \exp\{\lambda(e^t - 1)\}$

#6 $X \sim \Gamma(\alpha, \lambda)$

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}$$

Let $Z = \dfrac{X - \alpha/\lambda}{\sqrt{\alpha}/\lambda} = \dfrac{\lambda X - \alpha}{\sqrt{\alpha}}$

$$M_Z(t) = E\left\{ e^{\left(\frac{\lambda}{\sqrt{\alpha}} X - \sqrt{\alpha}\right)t} \right\}$$

$$= e^{-\sqrt{\alpha}\, t}\, M_X\left(\frac{\lambda}{\sqrt{\alpha}} t\right)$$

$$= e^{-\sqrt{\alpha}\, t} \left(\frac{\lambda}{\lambda - \frac{\lambda}{\sqrt{\alpha}} t}\right)^{\alpha}$$

$$\log M_Z(t) = -\sqrt{\alpha}\, t + \left[\log(\lambda) - \log\left(\lambda - \frac{\lambda}{\sqrt{\alpha}} t\right)\right]\alpha$$

$$= -\sqrt{\alpha}\, t + \left[\log\lambda - \log\lambda - \log\left(1 - \frac{1}{\sqrt{\alpha}} t\right)\right]\alpha$$

$$= -\left\{\sqrt{\alpha}\, t + \alpha \log\left(1 - \frac{1}{\sqrt{\alpha}} t\right)\right\}$$

$$= -\left\{\cancel{\sqrt{\alpha}\, t} + \alpha\left(\cancel{-\frac{1}{\sqrt{\alpha}} t} - \frac{1}{2\alpha} t^2 - \frac{1}{3}\left(\frac{t}{\sqrt{\alpha}}\right)^3 - \cdots\right)\right\}$$

$$\longrightarrow t^2/2 \quad \text{as} \quad \alpha \to \infty \qquad \blacksquare$$

#10

$X =$ # of sixes that appear $\sim$ Binomial $(100, \frac{1}{6})$

X is approximatated by $\tilde{X} \sim N(\frac{100}{6}, \frac{500}{36})$

$P(15 \le X \le 20) \doteq P(15 \le \tilde{X} \le 20)$

$= P\left(\frac{15 - 100/6}{\sqrt{500/36}} \le Z \le \frac{20 - 100/6}{\sqrt{500/36}}\right)$

$= P(-.447 \le Z \le .894) = .487$

(True value is 0.46)

~~$X_i \sim$ Binomi~~

Let $X_i = 1, 2, 3, 4, 5, 6$ w/ Prob $\frac{1}{6}$

$P\left(\sum_{i=1}^{100} X_i < 300\right) = P\left(\frac{1}{100} \sum_{i=1}^{100} X_i \le 3\right)$

$\bar{X} = \frac{1}{100} \sum_{i=1}^{100} X_i \sim N\left(3.5, \frac{\sqrt{105/36}}{\sqrt{100}}\right) = N(3.5, (.1707)^2)$

note $Var(X_i) = E(X_i^2) - (E(X_i))^2$

$= (1+4+9+16+25+36)/_6 - (3.5)^2$

$= \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{105}{36}$

$P(\bar{X} \le 3) = P\left(Z \le \frac{3 - 3.5}{.1707}\right) = P(Z \le -2.929$

$= .002$

#12 Let x_i^* be the true value $i=1,\dots,100$
and x_i ~~...~~ be x_i^* with the round off. So

$$x_i^* = x_i + \varepsilon_i \qquad \varepsilon_i \sim \text{unif}[-\tfrac{1}{2},\tfrac{1}{2}]$$

a. $E(\varepsilon_i)=0 \quad Var(\varepsilon_i)=\frac{1}{12}$

$$P\left(\sum_{i=1}^{100}\varepsilon_i>1\right) = P\left(\bar{\varepsilon}>\frac{1}{100}\right) = P\left(Z>\frac{.01-0}{\sqrt{1200}}\right)$$

$$= P(Z>.0003) \overset{\sim}{=} .50$$

$$\bar{\varepsilon} \sim N\left(0,\frac{1}{1200}\right)$$

b. $P\left(\bar{\varepsilon}>\frac{5}{100}\right) = P\left(Z>\frac{.05}{\sqrt{1200}}\right) = P(Z>.0014) = .499$

#13 $X_i = \begin{cases} 50 & \text{if he steps north at } i^{th} \text{ minute} \\ -50 & \text{" " " south " "} \end{cases}$

L = location after 1 hour $= \sum_{i=1}^{60} X_i$

$E(X_i)=0 \quad Var(X_i) = 5000$

$L \sim \cancel{N(0, (5000)\ldots)} \; N(0,(5000)60)$

#14 $E(X_i) = 50(2/3) - 50(\frac{1}{3}) = \frac{50}{3}$

$Var(X_i) = 5000$

$L \sim N\left(60\left(\frac{50}{36}\right), 5000(60)\right)$

#21

a) Let $Z_i = \dfrac{f(X_i)}{g(X_i)}$

$$E(Z_i) = \int_a^b \frac{f(x)}{g(x)}\, g(x)\, dx = \int_a^b f(x)\, dx$$

$$\Rightarrow E\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right) = \int_a^b f(x)\, dx = I(f)$$

b) $\mathrm{Var}(Z_i) = E(Z_i^2) - \left(E(Z_i)\right)^2$

$$= \int_a^b \frac{f^2(x)}{(g(x))^2}\, g(x)\, dx - \left(I(f)\right)^2$$

$$= \int_a^b f^2(x)/g(x)\, dx - \left(I(f)\right)^2$$

$$\therefore \mathrm{Var}(\hat{I}(f)) = \frac{1}{n}\left[\int_a^b \frac{f^2(x)}{g(x)}\, dx - \left(I(f)\right)^2\right]$$

Let $[a,b] = [-1, 1]$. Take $g(x) = \frac{3}{2}x^2 \quad -1 \le x \le 1$

And take $f(x) = 1$. This gives an example with infinite variance.

Take $g(x) \sim \mathrm{Unif}(0,1)$ and $f(x) = 1$. This is a finite example.

#21c

That it is the same as Exmpl A is obvious.

If $g(x)=1$, then

$$Var(\hat{I}(f)) = \frac{1}{n}\left[\int_0^1 f^2(x)\,dx - [I(f)]^2\right]$$

Choosing a $g(x) \neq 1$ would be more efficient if

$$\int_0^1 f^2(x)\,dx > \int_0^1 \frac{f^2(x)}{g(x)}\,dx$$

25

Chapter 6

#2.

$F = \frac{U/m}{V/n} = \frac{n}{m}\left(\frac{U}{V}\right) \quad W = V \qquad U = \frac{m}{n} FW$

$\frac{\partial F}{\partial U} = \frac{n}{mV} \qquad \frac{\partial F}{\partial V} = -\frac{n}{m}\frac{U}{V^2}$

$\frac{\partial W}{\partial U} = 0 \qquad \frac{\partial W}{\partial V} = 1$

$|J|^{-1} = \frac{mV}{n}$

$F_{UV}(u,v) = \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} v^{\frac{n}{2}-1} u^{\frac{m}{2}-1} e^{-v/2} e^{-u/2}$

$F_{F,W}(f,w) = \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} \cdot w^{\frac{n}{2}-1}\left(\frac{m}{n} f w\right)^{\frac{m}{2}-1} e^{-w/2} e^{-\frac{mfw}{2n}}$

$\int_0^\infty F_{F,W}(f,w)\,dw = \frac{1}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}\left(\frac{m}{n}\right)^{\frac{m}{2}-1} f^{\frac{m}{2}-1}$

Work out the remaining Algebra

$\int_0^\infty w^{\frac{m+n}{2}-2} e^{-w\left[\frac{1}{2}+\frac{mf}{2n}\right]} dw$

$\alpha = \frac{m+n}{2} + 1 \qquad \lambda = \frac{1}{2} + \frac{mf}{2n}$

$= \frac{\cancel{(m/n)^{\frac{m}{2}-1} f^{\frac{m}{2}-1}}}{\cancel{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}} \cancel{\left(\frac{1}{2}+\frac{mf}{2n}\right)^{-\frac{m+n}{2}}} \Gamma\left(\frac{m+n}{2}+1\right)$

$= \frac{\cancel{\left(\frac{m}{n}\right)^{\frac{m}{2}-1} f^{\frac{m}{2}-1}}}{2^{\frac{m+n}{2}}\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} \left(\frac{1}{2}\right)^{-1-\frac{m+n}{2}} \left(\frac{m}{n}\right)^{-1} \left(1 + \frac{mf}{n}\right)^{-\frac{m+n}{2}} \left(\frac{1}{2}\frac{mf}{2n}\right)^{-1} \left(\frac{m+n}{2}\right)\Gamma\left(\frac{m+n}{2}\right)$

work the remaining detail

#8 Let $X \sim exp(1)$
$Y \sim exp(1)$

Consider $U = \frac{X}{2}$

$M_U(t) = M_X(\frac{t}{2}) = \left(\frac{1}{1-2t}\right)$

Note that if $W \sim \chi^2(2)$
then $M_W(t) = \left(\frac{\frac{1}{2}}{\frac{1}{2} - t}\right) = \frac{1}{1-2t}$

So $U \sim \chi^2(2)$

$\Rightarrow$ $X/_2 \sim \chi^2(2)$, $Y/_2 \sim \chi^2(2)$

$$\frac{[X/2]/2}{[Y/2]/2} \sim F(2,2)$$

$\Rightarrow \frac{X}{Y} \sim F(2,2)$

3.5.6 key solutions for chapter 8

Chapter 8

#4 $X \sim bin(n, p)$

a) $f_X(x|p) = \binom{n}{x} p^x (1-p)^{n-x}$

$$l(p) = \log f_X(x|\theta) = \log \binom{n}{x} + x \log p + (n-x) \log(1-p)$$

$$\frac{\partial l(p)}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p} = \frac{x(1-p) - p(n-x)}{(1-p)p}$$

$$= \frac{x - \cancel{px} - pn + \cancel{px}}{(1-p)p} = 0$$

$$\Rightarrow x = pn \Rightarrow \boxed{\hat{p} = \frac{x}{n}}$$

b)

$$Var(\hat{p}) = Var\left(\frac{X}{n}\right) = \frac{1}{n^2} Var(X) = \frac{1}{n^2} np(1-p)$$

$$= \frac{p(1-p)}{n} \quad (*)$$

Cramer-Rao lower bound $= \dfrac{1}{I(p)}$

$$I(p) = E\left[\frac{\partial}{\partial p} \log f(x|\theta)\right]^2$$

$$= E\left\{\frac{x - pn}{p(1-p)}\right\}^2 = \frac{E(X-np)^2}{p^2(1-p)^2} = \frac{Var(X)}{p^2(1-p)^2}$$

$$= \frac{np(1-p)}{p^2(1-p)^2} = \frac{n}{p(1-p)} \Rightarrow \frac{1}{I(p)} = \frac{p(1-p)}{n} \quad (**)$$

$$\boxed{* = **}$$

#5

a) $E(X) = \frac{1}{p} \Rightarrow \bar{X} = \frac{1}{p} \Rightarrow \hat{p} = \frac{1}{\bar{X}}$

b) $\log(f(x|p)) = \log(p) + (x-1)\log(1-p)$

Given a sample $X_1, \ldots, X_n$

$$\ell(p) = \sum_{i=1}^{n} \{\log p + (x_i - 1)\log(1-p)\}$$

$$= n(\log p - \log(1-p)) + \log(1-p)\sum_{i=1}^{n} x_i$$

$$= n\{\log p - \log(1-p) + \bar{x}\log(1-p)\}$$

$$= n\{\log p + (\bar{x}-1)\log(1-p)\}$$

$$\frac{\partial \ell}{\partial p} = n\left\{\frac{1}{p} - \frac{\bar{x}-1}{1-p}\right\} = 0$$

$$\Rightarrow \hat{p} = 1/\bar{X}$$

c) $Var(\hat{p}) \to \frac{1}{nI(p)}$

$$\frac{\partial^2 \ell}{\partial p^2} = n\left\{\frac{-1}{p^2} + \frac{\bar{X}-1}{(1-p)^2}\right\} \Rightarrow -E\left(\frac{\partial^2 \ell}{\partial p^2}\right) = +\frac{n}{p^2} + \frac{n}{(1-p)^2} - \frac{nE(\bar{X})}{(1-p)^2}$$

$$= n/p^2(1-p) \qquad \text{Note } E\bar{X} = \frac{1}{p}$$

$$\Rightarrow Var(\hat{p}) \to \frac{p^2(1-p)}{n}$$

#9

In the method of Bootstrap, we generate at each step 23 Poisson ($\lambda = 24.9$) and compute the mean of the generated sample. The question is what is the variance of these means.

Since $Var(\bar{X}) = \frac{1}{n} var(X_i) = \frac{\lambda}{n}$

we have variance of the generated $\bar{x}$ should be $\frac{24.9}{23}$ or in general $\frac{\hat{\lambda}}{n}$

$\therefore S_{\hat{\lambda}} = \sqrt{\frac{\hat{\lambda}}{n}}$.

#10

MM estimates: $\hat{\lambda} = \frac{\bar{X}}{\hat{\sigma}^2}$ $\hat{\alpha} = \frac{\bar{X}^2}{\hat{\sigma}^2}$

ML estimates: $\hat{\lambda} = \frac{\hat{\alpha}}{\bar{X}}$

$$n \log \hat{\alpha} - n \log(\bar{x}) + \sum \log x_i - n \frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} = 0$$

It's difficult to assess true variances. MM estimates are easier to calculate, but on the other hand ML estimates are most efficient among unbiased estimates. So asymptotically ML estimates maybe better.

#11 $f(x|\alpha) = \frac{1+\alpha x}{2}$ $-1 \le x \le 1$ $-1 \le \alpha \le 1$

a)

$E(\hat{\alpha}) = E(3\bar{X}) = 3E(\bar{X})$

$= 3\ E(X) = 3\int_{-1}^{1} \frac{x+\alpha x^2}{2} dx$

$= 3 \left\{ \frac{1}{2} \left[\frac{x^2}{2} + \alpha \frac{x^3}{3} \Big|_{-1}^{1} \right] \right.$

$= \alpha$

b) $Var(\hat{\alpha}) = 9\,Var(\bar{X}) = \frac{9}{n} Var(X)$

$E(X^2) = \frac{1}{2}\int_{-1}^{1} (x^2 + \alpha x^3)\, dx$

$= \frac{1}{2} \left\{ \frac{x^3}{3} + \alpha \frac{x^4}{4} \Big|_{-1}^{1} \right\}$

$= \frac{1}{2} \left\{ \frac{2}{3} \right\} = \frac{1}{3}$

$Var(\hat{\alpha}) = \frac{9}{n} \left\{ \frac{1}{3} - \left(\frac{\alpha}{3}\right)^2 \right\}$

$= \frac{1}{n}(3 - \alpha^2)$

#c) $\hat{\alpha} \overset{app}{\sim} N\left(\alpha, \frac{1}{n}(3-\alpha^2)\right)$

$\hat{\alpha} \overset{app}{\sim} N\left(0, \frac{3}{25}\right)$

$P(|\hat{\alpha}| > .5) = 1 - P(|\hat{\alpha}| \le .5)$

$= 1 - P(-.5 \le \hat{\alpha} \le .5) = 1 - P\left(\frac{-.5-0}{\sqrt{3/25}} \le Z \le \frac{+.5}{\sqrt{3/25}}\right)$

#14 $E(X) = 0$ since f is odd

$$Var(X) = E(X^2) = \int_{-\infty}^{0} \frac{x^2}{2\sigma} e^{x/\sigma}\,dx + \int_0^{\infty} \frac{x^2}{2\sigma} e^{-x/\sigma}\,dx$$

$$= 2\int_0^{\infty} \frac{x^2}{2\sigma} e^{-x/\sigma}\,dx$$

$$= 2\sigma^2$$

$$2\sigma^2 = \frac{1}{n}\sum X_i^2$$

$$\Rightarrow \sigma^2 = \frac{1}{2n}\sum X_i^2$$

$$\Rightarrow \hat{\sigma} = \sqrt{\frac{1}{2n}\sum X_i^2}$$

b) $\log f(x|\sigma) = -\log 2 - \log\sigma - \frac{|x|}{\sigma}$

$$\ell(\sigma) = n\{-\log 2 - \log\sigma\} - \frac{1}{\sigma}\sum_{i=1}^{n}|x_i|$$

$$\ell'(\sigma) = \frac{-n}{\sigma} + \frac{\sum|x_i|}{\sigma^2} \Rightarrow \hat{\sigma} = \frac{1}{n}\sum|x_i|$$

c) $\cancel{E\left(\frac{\partial}{\partial\sigma}\log f(x|\sigma)\right)^2 = E\left\{-\frac{1}{\sigma} + \frac{|x|}{\sigma^2}\right\}^2}$

$$\cancel{= E\left\{\frac{1}{\sigma^2} + \frac{1}{\sigma^4}E(X^2)\right\} = -\frac{1}{\sigma} + \frac{1}{\sigma^2}(2\sigma^2)}$$

$$\cancel{= 2 - \frac{1}{\sigma}}$$

#14C $\frac{\partial}{\partial\sigma}\log f(x|\sigma) = -\frac{1}{\sigma} + \frac{|x|}{\sigma^2}$

$\frac{\partial^2}{\partial\sigma^2}\log f(x|\sigma) = \frac{1}{\sigma^2} - 2\frac{|x|}{\sigma^3}$

$I(\sigma) = -E\left\{\frac{1}{\sigma^2} - 2\frac{|x|}{\sigma^3}\right\}$

$E(|x|) = \frac{1}{2\sigma}\int_{-\infty}^{\infty}|x|e^{-\frac{|x|}{\sigma}}dx$

$= \frac{1}{2\sigma}\left\{\int_{-\infty}^{0}-x e^{x/\sigma}dx + \int_0^\infty x e^{-x/\sigma}dx\right\}$

$= \frac{1}{\sigma}\int_0^\infty x e^{-x/\sigma}dx = \sigma$

$\Rightarrow I(\sigma) = -E\left\{\frac{1}{\sigma^2} - \frac{2}{\sigma^2}\right\} = \frac{1}{\sigma^2}$

$\therefore$

$\text{avar}(\hat{\sigma}) = \frac{1}{nI(\theta)} = \frac{\sigma^2}{n}$

#16 a) $E(X^2) = Var(X) + (E(X))^2$

$$E(X^2) = \frac{2}{9(3\alpha+1)} + \frac{1}{9}$$

$$\frac{2}{9(3\alpha+1)} + \frac{1}{9} = m_2 \quad \text{where } m_2 = \frac{1}{n}\sum X_i^2$$

Solve for α

$$\hat{\alpha} = \frac{3m_2 - 1}{1 - 9m_2}$$

b) $\log f(x|\alpha) = \log \Gamma(3\alpha) - \log \Gamma(\alpha) - \log \Gamma(2\alpha) + (\alpha - 1)\log x_i + (2\alpha - 1)\log(1-x)$

$$l(\alpha) = \sum_{i=1}^{n} \{ \log \Gamma(3\alpha) + \log \Gamma(\alpha) - \log \Gamma(2\alpha) + (\alpha-1)\log x_i + (2\alpha-1)\log(1-x_i) \}$$

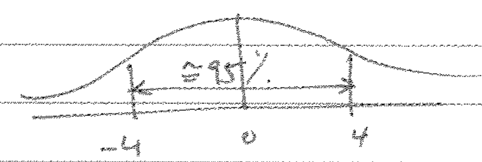
$$\frac{\partial l(\alpha)}{\partial \alpha} = \sum_{i=1}^{n} \left\{ \frac{3\Gamma'(3\alpha)}{\Gamma(3\alpha)} + \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \frac{2\Gamma(2\alpha)}{\Gamma'(2\alpha)} + \log x_i + 2\log(1-x_i) \right\} = $$

c) $$\frac{\partial^2 l(\alpha)}{\partial \alpha^2} = n \left\{ \frac{9\Gamma''(3\alpha)\Gamma(3\alpha) - 3[\Gamma'(3\alpha)]^2}{(\Gamma(3\alpha))^2} + \frac{\Gamma''(\alpha)\Gamma(\alpha) - (\Gamma'(\alpha))^2}{(\Gamma(\alpha))^2} - \frac{4(\Gamma'(2\alpha))^2 - 4\Gamma(2\alpha)\Gamma''(2\alpha)}{(\Gamma'(2\alpha))^2} \right\}$$

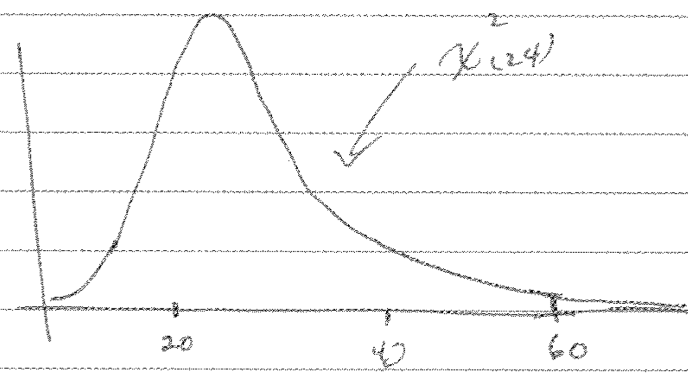
Note That $I(\alpha) = -E\left(\frac{\partial^2 l(\alpha)}{\partial \alpha}\right) = -\frac{\partial^2 l(\alpha)}{\partial \alpha^2}$ in this case

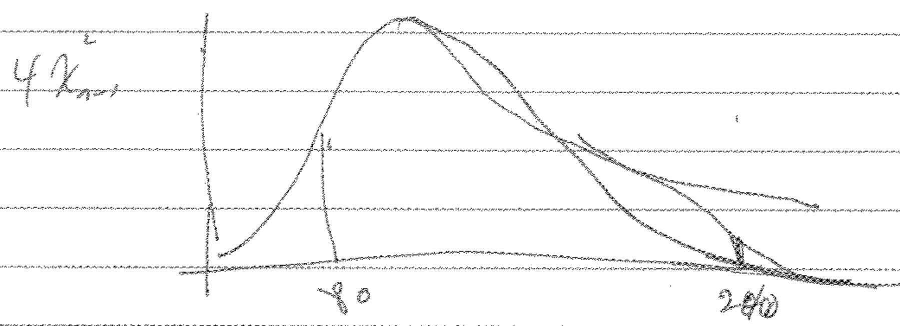
Thus $AVar \cong \dfrac{1}{nI(\theta)}$

#18 $\bar{X} \sim N(0, \frac{100}{25})$



$\hat{\sigma}^2 = \frac{\sigma^2}{n}\chi^2_{n-1} = 4\chi^2_{(n-1)}$





#1.9

a) $E(X)=\int_\theta^\infty x e^{-(x-\theta)}\,dx$

MM:

$$= e^{\theta}\int_\theta^\infty x e^{-x}\,dx = e^{\theta}\, e^{-\theta}(\theta+1) = \theta+1$$

$$\theta+1 = \frac{1}{n}\sum x_i$$

$$\hat{\theta} = \bar{X}-1$$

b) $f(x|\theta) = \begin{cases} e^{-x+\theta} & x \geq \theta \\ 0 & o.w. \end{cases}$

$$L(\theta) = \begin{cases} e^{n\theta - \sum_{i=1}^n x_i} & x_1 \geq \theta \quad x_2 \geq \theta, \ldots x_n \geq \theta \\ 0 & o.w. \end{cases}$$

$$\sum x_i \geq n\theta \Rightarrow n\theta - \sum_{i=1}^n x_i < 0$$

The maximum is attained if $n\theta - \sum_{i=1}^n x_i$ is largest

~~This happens when $\theta = (x_1, \ldots, x_n)$~~ However, note that if $\theta > x_i$ Then the likelihood becomes zero. So the maximum is attained at a value of θ such that $\theta = \min(X_1, \ldots, X_n)$

#21

Method of moments estimate $= 2\bar{X} + 1$

$= 2(888) + 1$ see note

MLE $= \max_{1 \le i \le n}(X_i) = 888$ see note

Note $n = 1$ in this case

#25

a) Let $T = \min(T_1, T_2, T_3, T_4, T_5)$

We have observed $T_{(1)} = 100$

likelihood is

$$L(\tau) = \frac{5}{\tau} e^{-5T/\tau}$$ see top of Page 101 of the text

b) $l(\tau) = \log 5 - \log \tau - 5T/\tau$

$l'(\tau) = -\frac{1}{\tau} + \frac{5T}{\tau^2} \Rightarrow \hat{\tau} = 5T$

$\therefore \tau = 500$

c) $P(\hat{\tau} \le x) = P(5T \le x)$

$\Rightarrow P(T \le \frac{x}{5}) = \int_0^{\frac{x}{5}} \frac{5}{\tau} e^{-5T/\tau} \, dT$

$= -e^{-5T/\tau} \Big|_0^{x/5} = 1 - e^{-x/\tau} \Rightarrow \hat{\tau} \sim \exp(\frac{1}{\tau})$

d) $\text{var}(\hat{\tau}) = \hat{\tau}^2 \Rightarrow \text{s.e.}(\hat{\tau}) = \hat{\tau}$.

#28 $\hat{\mu} = 3.61$ $\hat{\sigma}^2 = 3.2$

b) $(2.8, 4.4)$ $(2.6, 4.6)$ $(2.2, 5.0)$

c) $(2.1, 7.1)$ $(1.9, 8.2)$ $(1.6, 11.1)$

d)

Take 95% CI: $z_{\alpha/2}\,\hat{\sigma}/\sqrt{n} = 1.96(3.2)^{1/2}/\sqrt{n} = 1/2$

$\sqrt{n} \cong 2(1.96)(3.2)^{1/2}$

$n \cong 16(3.2) \cong 52$

#38 Use data disk for data.

#42 a) $\hat{\theta} = \bar{X}\sqrt{2/n}$

b) $\ell(\theta) = \sum_{i=1}^{n}\left[\log x_i - 2\log\theta - \frac{x_i^2}{2\theta^2}\right]$

$\ell'(\theta) = -\frac{2n}{\theta} + \frac{1}{2\theta^3}\sum_{i=1}^{n} x_i^2 = 0$

$4n\hat{\theta}^2 = \sum x_i^2$

$\hat{\theta}^2 = \frac{1}{4n}\sum x_i^2$

$\hat{\theta} = \frac{1}{2}\sqrt{\sum x_i^2/n}$

c) $f'(x|\theta) = -\frac{2}{\theta} + \frac{x}{2\theta^3}$ $f''(\theta) = \frac{2}{\theta^2} - \frac{x}{2\theta^4}$

$I(\theta) = -E(f''(x|\theta)) = -\frac{2}{\theta^2} + \frac{1}{2\theta^4}\underbrace{E(X)}_{4\theta^2} = +\frac{4}{\theta^2}$

$\therefore \text{Var}(\hat{\theta}) = \frac{\theta^2}{4n}$

#43

$$l(\theta) = c - \sum_{i=1}^{2m+1} |x_i - \theta| \qquad c = \text{Constant}$$

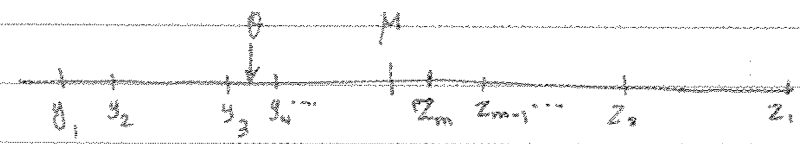
Maximizing $l(\theta)$ is equivalent to

$$\min \sum_{i=1}^{2m+1} |x_i - \theta|$$

Order $x_1, x_2, \ldots x_{2m+1}$ as follows (renaming)

$$y_1 \le y_2 \le \cdots \le y_m \le M \le z_m \le z_{m-1} \le \cdots \le z_1$$

where $y_i = x_{(i)}$, $z_i = x_{(2m-i+2)}$ $M = \text{median}$



$$\sum_{i=1}^{2m+1} |x_i - \theta| = \sum_{i=1}^{m} |y_i - \theta| + \sum_{i=1}^{m} |z_i - \theta| + |M - \theta|$$

Let $y_k \le \theta \le y_{k+1}$, for some k.

$|y_1 - \theta| + |z_1 - \theta| = z_1 - y_1$ | $|y_1 - M| + |z_1 - M| = z_1 - y_1$

$|y_2 - \theta| + |z_2 - \theta| = z_2 - y_2$ | $|y_2 - M| + |z_2 - M| = z_2 - y_2$

$|y_k - \theta| + |z_k - \theta| = z_k - y_k$ | $\vdots$

$|y_{k+1} - \theta| + |z_{k+1} - \theta| = z_k - y_{k+1} + 2(y_{k+1} - \theta)$

$\vdots$ | $|z_m - M| + |y_m - M| = z_m - y_m$

$|z_m - \theta| + |y_m - \theta| = z_m - y_m + 2(y_m - \theta)$

$|M - \theta|$ | $|M - M|$

Note that all the quantities on the right are less than or equal to those on the left. Therefore for any $y_k < \theta < y_{k+1}$ M beats θ. Other Cases Can be shown similarly

#45

a) $E(X_i) = \frac{\theta}{2}$ $\quad \bar{X} = \frac{\theta}{2} \implies \boxed{\hat{\theta} = 2\bar{X}}$

$E(\hat{\theta}) = 2E(\bar{X}) = 2\,\frac{\theta}{2} = \theta$

$Var(\hat{\theta}) = 4\,Var(\bar{X}) = 4\,\frac{Var(X_i)}{n} = \frac{4\theta^2}{12n} = \frac{\theta^2}{3n}$

b)
$$L(\theta) = \begin{cases} \frac{1}{\theta^n} & 0 \le X_i \le \theta \quad i=1,\dots,n \\ 0 & o.w. \end{cases}$$

$\therefore$ Under the constraints $\hat{\theta} = \max(X_1, \dots, X_n)$.

c) The probability density of $\max(X_1, \dots, X_n)$ is given by

$$f_{\max(X_1,\dots,X_n)}(x) = \frac{n x^{n-1}}{\theta^n} \qquad 0 \le x \le \theta$$

$\therefore$

$E(\hat{\theta}) = \int \frac{n x^n}{\theta^n}\,dx = \frac{n}{n+1}\,\theta$

$Var(\hat{\theta}) = E(\hat{\theta}^2) - \{E(\hat{\theta})\}^2$

$E(\hat{\theta}^2) = \int \frac{n x^{n+1}}{\theta^n}\,dx = \frac{n}{n+2}\,\theta^2$

$Var(\hat{\theta}) = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2 = \frac{n}{(n+2)(n+1)^2}\,\theta^2$

The mean squared error is

$$\begin{aligned} MSE &= Var + (bias)^2 \\ &= \frac{n}{(n+2)(n+1)^2}\theta^2 + \left(\frac{n}{n+1}\theta - \theta\right)^2 \\ &= \frac{n}{(n+2)(n+1)^2}\theta^2 + \frac{\theta^2}{(n+1)^2} = \frac{2}{(n+1)(n+2)}\,\theta^2 \end{aligned}$$

45c Cont)

For the method of moment

$MSE = Var(\hat{\theta}) = \frac{\theta^2}{3n}$

For all values of n for which

$\frac{2}{(n+1)(n+2)} < \frac{1}{3n}$ MLE has smaller MSE

$\iff 6n < n^2+3n+2$

$\iff n^2-3n+2 > 0$ $\quad n = \frac{3 \pm \sqrt{9-8}}{2}$ $\begin{cases} 2 \\ 1 \end{cases}$

$(n-2)(n-1) > 0$

So for $n > 2$ MLE has better MSE.

d) Use $\frac{n+1}{n} max(x_1, \_\_, x_n)$.

#52

a) $\ell(\tau) = \sum_{i=1}^{n} \left[\ln\left(\frac{1}{\tau}\right) - \frac{1}{\tau} x_i\right]$

$= -n \ln \tau - \frac{1}{\tau} \sum x_i$

$\ell'(\tau) = -\frac{n}{\tau} + \frac{1}{\tau^2} \sum x_i = 0$

$\Rightarrow \hat{\tau} = \bar{X}$

b) $X_i \sim \Gamma(1, \frac{1}{\tau})$

$\sum X_i \sim \Gamma(n, \frac{1}{\tau})$

$\hat{\tau} = \frac{1}{n} \sum x_i \sim \Gamma(n, \frac{n}{\tau})$

c) $E(\bar{X}) = \tau \qquad Var(\bar{X}) = \frac{\tau^2}{n}$

$\bar{X} \sim N(\tau, \frac{\tau^2}{n})$

d) $E(\hat{\tau}) = \tau$ using the fact that $\hat{\tau} \sim \Gamma(n, \frac{n}{\tau})$
$Var(\hat{\tau}) = \tau^2/n$

e) Using the Cramer-Rao lower bound
$Var(T) \geq \tau^2/n$ for any unbiased estimator τ
Note: $I(\tau) = \frac{1}{\tau^2}$

Therefore this is the best.

F) $\bar{X} \sim \Gamma(n, \frac{n}{\tau}) \Rightarrow \tau \bar{X} \sim \Gamma(n, n)$

$\Rightarrow P\left(\Gamma_{\alpha/2}(n,n) < \tau \bar{X} < \Gamma_{1-\alpha/2}(n,n)\right) = 1-\alpha$

$P\left(\frac{\Gamma_{\alpha/2}(n,n)}{\bar{X}} < \tau < \frac{\Gamma_{1-\alpha/2}(n,n)}{\bar{X}}\right) = 1-\alpha$

3.5.7 key solutions for chapter 9

Chapter 9

#3 $X \sim \mathcal{B}(100, p)$

$$\begin{cases} H_0: p = 0.5 \\ H_a: p \neq 0.5 \end{cases}$$

Rejection Region $\{x : |x - 50| > 10\}$

a. $\alpha = P(\text{rejecting } H_0 \mid H_0 \text{ true})$

$= P(|X - 50| > 10 \mid p = 0.5)$

$= 1 - P(40 < X < 60 \mid p = 0.5)$

$\approx 1 - P\left(\frac{40-50}{\sqrt{25}} < Z < \frac{60-50}{\sqrt{25}}\right)$

$= 1 - P(-2 < Z < 2) \approx 0.05$

b. Power $= 1 - \beta$

$= 1 - P(\text{accepting } H_0 \mid H_a \text{ true})$

$= 1 - P(|X - 50| \leq 10 \mid p \neq 0.5)$

$= 1 - P\left(\frac{40 - 100p}{\sqrt{100p(1-p)}} \leq Z \leq \frac{60 - 100p}{\sqrt{100p(1-p)}} \mid p \neq 0.5\right)$

see next page for the plot

7

$$\frac{f_0}{f_1} = \frac{e^{-\lambda_0} \lambda_0^{\sum x_i}}{e^{-\lambda_1} \lambda_1^{\sum x_i}}$$

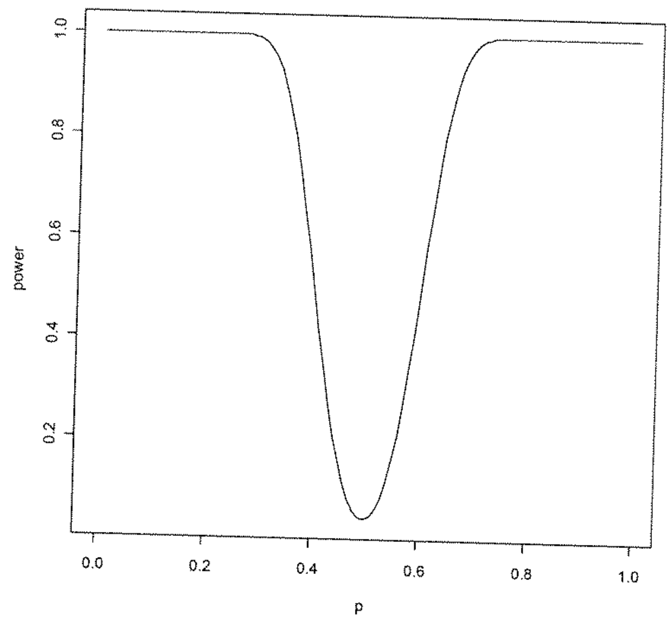
$$= e^{(\lambda_1 - \lambda_0)} \left(\frac{\lambda_0}{\lambda_1}\right)^{\sum x_i}$$

We reject H_0 if $\frac{f_0}{f_1}$ is small. since $\frac{\lambda_0}{\lambda_1} < 1$, this happens when $\sum x_i$ is large. $\sum x_i \sim P(n\lambda)$.

So we determine c such that

$$P(X > c) = \alpha \quad \text{where}$$

$$X \sim P(n\lambda_0).$$


power
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
p

#9

$$\frac{f_0}{f_A} = \frac{e^{-\frac{1}{2}\sum_{i=1}^{25}(x_i-0)^2/100}}{e^{-\frac{1}{2}\sum_{i=1}^{25}(x_i-1.5)^2/100}}$$

$$= e^{-1/200 \sum_{i=1}^{25}[(x_i-0)^2-(x_i-1.5)^2]}$$

$$= e^{-1/200 \sum_{i=1}^{25}[x_i^2-x_i^2+3x_i-(1.5)^2]}$$

$$= e^{-\frac{1}{200}\{3\sum_{i=1}^{25}(x_i-2.25)\}}$$

$$= e^{-\frac{1}{200}\{75(\bar{x}-2.25)\}}$$

The test would reject for large values of $\bar{x}$

$\bar{X} \sim N(0, \frac{100}{25})$ under H_0

$P(\bar{X}>c) = 0.10 \qquad P(Z > \frac{c-0}{\sqrt{4}})$

$\frac{c}{2} = 1.282 \Rightarrow$ Reject if $\bar{x} > 2.56$

Power = $P(\text{rejecting } H_0 \mid H_a = \text{True})$

$= P(\bar{X} > 2.56 \mid \mu = 1.5)$

$= P\left(Z > \frac{2.56-1.5}{2}\right) = P(Z > .53) = .298$

Calculations for $\alpha = .01$ is similar

#12

$$\Lambda = \frac{\max_{\theta=\theta_0} l(\theta)}{\max_{\theta\neq\theta_0} l(\theta)} = \frac{\theta_0^n e^{-\theta_0 \sum_{i=1}^n x_i}}{\max_{\theta\neq\theta_0}\{\theta^n e^{-\theta \sum x_i}\}}$$

The max in the denominator occures at $\hat{\theta} = \frac{1}{\bar{x}}$

$$\therefore \Lambda = \left(\frac{\bar{x}}{\theta_0}\right)^n e^{-n\theta_0 \bar{x} + n} = \frac{e^n}{\theta_0^n}\left[\bar{x} e^{-\theta_0 \bar{x}}\right]^n$$

We reject Ho when Λ is small i.e.

$$\bar{x} e^{-\theta_0 \bar{x}} < c.$$

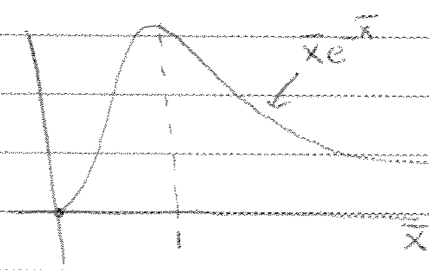
#13

a) The rejection region is of the form

$$\bar{x} e^{-\bar{x}} < c$$

To have significance level .05,

$$P(\bar{x} e^{-\bar{x}} < c) = .05$$



looking at the graph, it is clear that either small values of $\bar{x}$ or large values of $\bar{x}$ will lead to small values of the quantity $\bar{x}e^{-\bar{x}}$

b. $\alpha = P(\text{type I error}) = P(\text{rejecting } H_0 \mid H_0 \text{ true})$

$= P(\bar{X} e^{-\bar{X}} \le c \mid \theta_0 = 1)$

c. Using the mgf for the $X_i \sim \exp(\theta)$ it can be shown that

$$\sum_{i=1}^{n} X_i \sim \Gamma(\theta, n)$$

Specifically $\sum_{i=1}^{10} X_i \sim \Gamma(\lambda = \theta, \alpha = 10)$

We know that the rejection region is of the form $\{\bar{X} \le x_0\} \cup \{\bar{X} \ge x_1\}$

We need to ~~have~~ determine (given $\theta_0 = 1$) x_0 and x_1 such that

$$P\{\bar{X} \le x_0\} + P(\bar{X} \ge x_1\} = .05$$

But the knowledge that $n\bar{X} \sim \Gamma(\lambda = 1, \alpha = 10)$ will help finding appropriate x_0 and x_1

d. We generate many samples of size 10 from the exp(1) and obtain the empirical distn of $\bar{X} e^{-\bar{X}}$. We then choose x_0 and x_1 in such a way that overall 2.5% of the observed values of $\bar{X} e^{-\bar{X}}$ are above x_1 and ~~and~~ 2.5% are below x_0

#16 $X \sim bin(n,p)$

$H_0: p = 0.5 \qquad H_A: p \neq 0.5$

a. $$\Lambda = \frac{(0.5)^x (0.5)^{n-x}}{\left(\frac{x}{n}\right)^x \left(1-\frac{x}{n}\right)^{n-x}}$$

$$= \left(\frac{0.5n}{x}\right)^x \left(\frac{.5n}{n-x}\right)^{n-x}$$

$$= \left(\frac{.5n}{x}\right)\left(\frac{n-x}{0.5n}\right)^x \left(\frac{.5n}{n-x}\right)^n$$

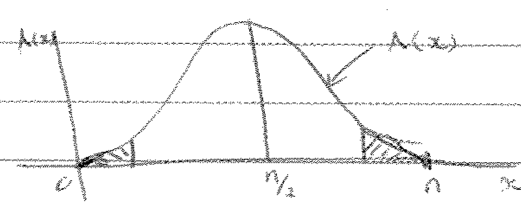
$$= \left(\frac{n-x}{x}\right)^x \left(\frac{.5n}{n-x}\right)^n$$

$$= (n-x)^{x-n}\, x^{-x} \left(n/2\right)^n$$

b. It can be shown that $\Lambda(x)$ is symmetric about $x = \frac{n}{2}$ and it attains its maximum at $x = \frac{n}{2}$

It follows that the rejection region is of the form

$|X - n/2| > k$.



c. Under H_0, $X \sim bin(n, 0.5)$ $\therefore$ k is determined such that

$P(|X - \frac{n}{2}| > k) = \alpha$

see section d for a specific example.

d. $X \sim bin(10, 0.5)$ under H_0

$P(|x-5|>2) = P(X>7) + P(X<3)$

$= P(X \geq 8) + P(X \leq 2)$

$= .0107 + .055 = .0654$

e. $P(|x-50|>10)$ $X \sim bin(100, .5)$

$= 1 - P(40 < X < 60)$ $\approx N(50, 25)$

$= 1 - P\left(\frac{40-50}{5} < Z < \frac{60-50}{5}\right)$

$= 1 - P(-2 < Z < 2) \approx .05$

#18

a. T

b. F

c. T

d. F

e. F

f. F

#20

a. P-value $= 2\,P(T > 1.5) =$

$= 2(.0668) = .1336$

b. P-value $= P(T > 1.5) = .0668$

#21

$P(T > t_0) = \alpha$ under H_0

$\Longrightarrow$

$P(S > g(t_0)) = P(g(T) > g(t_0))$

$= P(T > t_0) = \alpha$ since g is monotonically increasing

#25

$$-2\log\Lambda = 2n\sum_{i=1}^{m} \hat{p}_i \log\left(\frac{\hat{p}_i}{p_i(\hat{\theta})}\right)$$

$n = 1919$ $m = 2$ $\hat{p}_1 = \frac{1}{2}$ $\hat{p}_2 = \frac{1}{2}$

H_0: $p_1 = p_2 = \frac{1}{2}$

H_a: $p_1 < p_2$ p_1 = those who die before the holiday

$p_1(\hat{\theta}) = \frac{922}{1919}$ $p_2(\hat{\theta}) = \frac{997}{1919}$

$-2\log\Lambda = 2.933$ d.f. = 1

$P(\chi^2_{(1)} > 2.933) = .0868$

Do not reject H_0 at 5% level

For Japanese

$p_1(\hat{\theta}) = \frac{418}{852}$ $p_2(\hat{\theta}) = \frac{434}{852}$

$-2\log\Lambda = .3005$

P-value $= .583$

Do not reject H_0

#22

Under the null hypotheses $p_1 = p_2 = \cdots = p_{12} = \frac{1}{12}$
where p_i is the proportion of suicide in month i.

$n = 23480 \qquad m = 12$

$p_1(\hat{\theta}) = \frac{1867}{23480}, \ldots, p_{12}(\hat{\theta}) = \frac{1859}{23480}$

$$-2 \log \Lambda = +2n \sum_{i=1}^{12} \left[\log \left(\frac{\hat{p}_i}{p_i(\hat{\theta})} \right) \right] \hat{p}_i = 52.01$$

$P(\chi^2_{11} > 52.01) =$

P-value $< .0001$

The suicide rate does not appear to be constant

#31 Under the null hypothesis the log likelihood is maximized at $\hat{p}_1 = \hat{p}_2 = \cdots = \hat{p}_m = \frac{\sum x_i}{n} = \bar{x}$

Where $n = \sum_{i=1}^{n} n_i$

under the alternative hypothesis $\ell(\underset{\sim}{\theta})$ is maximized at $\hat{p}_i = \frac{x_i}{n_i}$

So

$$\Lambda = \frac{\bar{x}^{x_1}(1-\bar{x})^{n_1-x_1} \cdots \bar{x}^{x_m}(1-\bar{x})^{n-x_n}}{\prod \hat{p}_i^{x_i}(1-\hat{p}_i)^{n-x_i}}$$

$$= \frac{\bar{x}^{\sum x_i}(1-\bar{x})^{n-\sum x_i}}{\prod_{i=1}^{m} \hat{p}_i^{x_i}(1-\hat{p}_i)^{n-x_i}}$$

$$-2\log\Lambda \sim \chi^2_{m-1} \quad \text{Under } H_0$$

Chapter 4

Quizes

Local contents

4.1 Quiz 1

Local contents

Quiz 1 Math 502AB Fall 2007

Name (please print) Nasser Abbasi

1. Consider a sequence of days, and let R_i denote the event that it rains on day i. Let $P(R_0) = p$ (rain today), $P(R_i|R_{i-1}) = \alpha$, and $P(R_i^c|R_{i-1}^c) = \beta$. Suppose further that only today's weather is relevant to predicting tomorrow's; that is, $P(R_i|R_{i-1} \cap \cdots, \cap R_0) = P(R_i|R_{i-1})$. What is the probability that it rains n days from now? What happens as n approaches infinity?

Figure 4.1: Problem 1

Answer:

Given:

1. R_i ,Event that it rains on day i
2. R_i^c ,Event that it does not rain on day i
3. $P(R_0) = p$, Probability of rain on day 0
4. $P(R_i|R_{i-1}) = \alpha$,Probability of rain on day i given it rained on day $i-1$
5. $P\left(R_i^c|R_{i-1}^c\right) = \beta$,Probability of no rain on day i given it did not rain on day $i-1$
6. Only today's weather is relevant to predicting tomorrow rain

Find:

Probability of rain in n days and what happen as $n \to \infty$

Solution:

Consider the experiment that generates today's weather. Hence possible outcomes can be divided into 2 disjoint events: rain and no rain (A day can either be rainy or not, hence this division contains all possible outcomes).

Hence

$$\Omega = \{R_0, R_0^c\}$$

Now using the law of total probability, we write

$$P(R_1) = P(R_1|R_0)\,P(R_0) + P(R_1|R_0^c)\,P(R_0^c) \tag{1}$$

But

$$\begin{aligned} P(R_1|R_0^c) &= 1 - P(R_1|R_0) \\ &= 1 - \beta \end{aligned} \tag{2}$$

Note: To proof the above, we can utilize a simple state transition diagram as follows

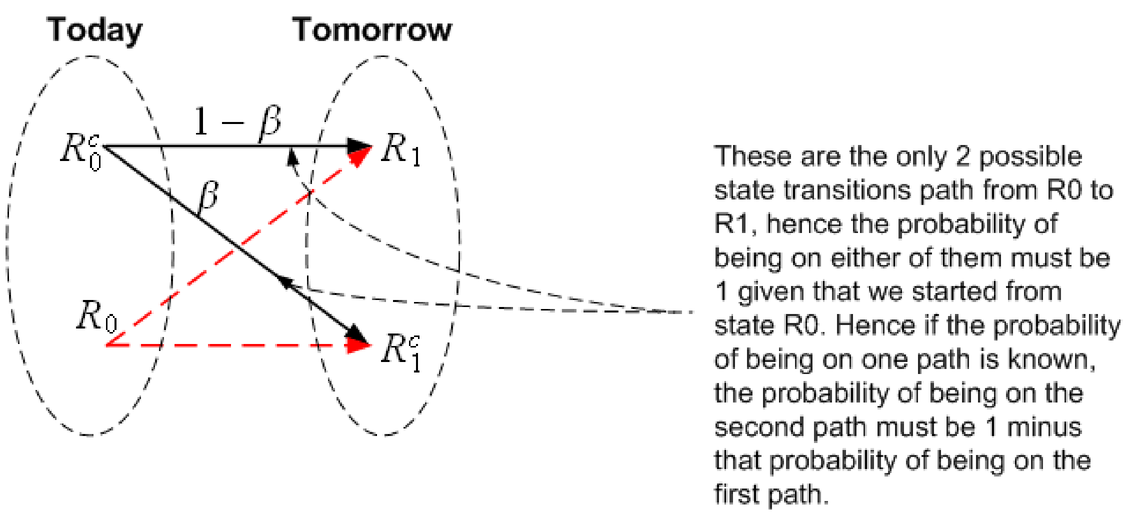


Figure 4.2: Problem 1 note

Now, substitute (2) into (1) and given that $P(R_1|R_0) = \alpha$ and $P(R_0) = p$ and $P(R_0^c) = 1-p$, then (1) becomes

$$P(R_1) = \alpha p + (1-\beta)(1-p) \tag{3}$$

Now we can recursively apply the above to find probability of rain on the day after tomorrow. Let $R_0 \to R_1$ and $R_1 \to R_2$, hence the above (1) becomes

$$P(R_2) = P(R_2|R_1)\,P(R_1) + P(R_2|R_1^c)\,P(R_1^c) \tag{4}$$

Now using (3) for $P(R_1)$, and given that $P(R_2|R_1) = \alpha$ (This probability does not change, since we are told only today's weather is relevant), and given that $P(R_1^c) = (1 - P(R_1))$ and that $P(R_2|R_1^c) = (1-\beta)$, then (4) becomes

$$\begin{aligned} P(R_2) &= \alpha\overbrace{[\alpha p + (1-\beta)(1-p)]}^{P(R_1)} + (1-\beta)\overbrace{(1 - [\alpha p + (1-\beta)(1-p)])}^{P(R_1^c)} \\ &= p + \alpha + \beta - 2p\alpha - 2p\beta - \beta^2 - \alpha\beta + p\alpha^2 + p\beta^2 + 2p\alpha\beta \\ &= p(1 - 2\alpha - 2\beta + 2\alpha\beta) + \alpha + \beta - \alpha\beta + (p\alpha^2 + p\beta^2 - \beta^2) \\ &= p(1 - 2\alpha - 2\beta + 2\alpha\beta) + \alpha + \beta - \alpha\beta + [\text{terms with higher powers in } \alpha \text{ and } \beta] \end{aligned}$$

We see that as we continue with the above process, terms will be generated with the form (something)$\times\alpha^m$ and (something)$\times\beta^r$, where the powers m,r are getting larger and larger as n gets larger. But since $\alpha,\beta<1$, hence all these terms go to zero. So we only need to look at the terms which do not contain a product of $\alpha's$ and product of $\beta's$

Hence the above reduces

$$P(R_2)\approx p(1-2\alpha-2\beta+2\alpha\beta)+\alpha+\beta-\alpha\beta$$

There is a pattern here, to see it more clearly, I generated more $P(R_i)$ for $i=3,4,5,6,7$ using a small piece of code and removed all terms of higher powers of α,β as described above, and I get the following table

i	$P(R_i)$
0	p
1	$1-\beta+p(-1+\alpha+\beta)$
2	$\alpha+\beta-\alpha\beta+p(1-2\alpha-2\beta+2\alpha\beta)$
3	$1-\alpha-2\beta+3\alpha\beta+p(-1+3\alpha+3\beta-6\alpha\beta)$
4	$2\alpha+2\beta-6\alpha\beta+p(1-4\alpha-4\beta+12\alpha\beta)$
5	$1-2\alpha-3\beta+10\alpha\beta+p(-1+5\alpha+5\beta-20\alpha\beta)$
6	$3\alpha+3\beta-15\alpha\beta+p(1-6\alpha-6\beta+30\alpha\beta)$

Hence the pattern can be seen as the following

$$\begin{aligned}P(R_n)=\operatorname{mod}(n,2)+(-1)^{(n)}\left\lfloor\frac{n}{2}\right\rfloor\alpha+(-1)^{(n)}\left\lceil\frac{n}{2}\right\rceil\beta+(-1)^{(n+1)}\left(\sum_{i=1}^{n-1}i\right)\alpha\beta+\\ p\left((-1)^n+(-1)^{n+1}n\alpha+(-1)^{n+1}n\beta+(-1)^n\left[n^2-n\right]\alpha\beta\right)\end{aligned}$$

Where $\operatorname{mod}(n,2)=0$ for even n and 1 for odd n, and $\left\lfloor\frac{n}{2}\right\rfloor$ means to round to nearest lower integer and $\left\lceil\frac{n}{2}\right\rceil$ means to round upper.

The above is valid for very large n.

As $n\to\infty$ $P(R_n)$ will reach a fixed value (I first though it will always go to 1, but that turned out not to be the case). I could not find an exact expression for $P(R_n)$ as $n\to\infty$, but I wrote a small program which simulates the above, and generates a table. Here is a table for few values as n gets large, these are all for $\alpha=.3,\beta=.6,p=.4$, notice that $P(R_n)$ fluctuates up and down from one day to the next as it converges to its limit.

```
In[29]:= TableForm[r, TableHeadings → {None, {"n", "P(Rn)"}}]

Out[29]//TableForm=
```

n	$P(R_n)$
0	0.4
1	0.28
2	0.316
3	0.3052
4	0.30844
5	0.307468
6	0.30776
7	0.307672
8	0.307698
9	0.30769
10	0.307693
11	0.307692
12	0.307692
13	0.307692
14	0.307692
15	0.307692
16	0.307692
17	0.307692

Figure 4.3: JNJY0F03

2 Show that if the conditional probabilities exist, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1 \cap A_2)\cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Figure 4.4: Problem 2

Given: Conditional probabilities exist

Show: $P(A_1 \cap A_2 \cap A_3 \cap \cdots A_n) = P(A_1)+P(A_2|A_1 \cap A_2)+\cdots+P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$

Solution:

Since Conditional probabilities exist, then we know that the following is true

$$P(X \cap Y) = P(X|Y)\, P(Y)$$

Let $X = A_n$ and $Y = A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}$ hence the above becomes

$$P(A_1 \cap A_2 \cap \cdots A_n) = P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\, P(A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Now apply the same idea to the last term above. In other words, we write

$$P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) = P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})\, P(A_1 \cap A_2 \cap \cdots \cap A_{n-2})$$

We repeat the process until we obtain $P(A_1 \cap A_2) = P(A_2|A_1)\, P(A_1)$

Hence, putting all the above together, we write

$$P(A_1 \cap A_2 \cap \cdots A_n) = P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\, P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})$$
$$P(A_{n-2}|A_1 \cap A_2 \cap \cdots A_{n-3})\cdots P(A_2|A_1)\, P(A_1)$$

The above is what is required to show (terms are just rewritten is reverse order from the problem statement, rearranging, we obtain

$$P(A_1 \cap A_2 \cap \cdots A_n) = P(A_1)\, P(A_2|A_1) \cdots P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})\, P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

3. Let A and B be arbitrary events. Use the three axioms of probability to show that

$$P(A \cup B) \leq P(A) + P(B).$$

Identify the axiom(s) that you use at each step. You are not allowed to use any theorems. [Hint: One way to show this is to first show that if C and D are events such that $C \subset D$, then $P(C) \leq P(D)$. Then, use this result to prove the result in the above display equation.]

Figure 4.5: Problem 3

Given:

Axioms of probability:

1. $P(\Omega) = 1$
2. if $A \subset \Omega$ then $P(A) \geq 0$
3. if A, B are disjoint events (i.e. $A \cap B = \varnothing$) then $P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$

Show that $P(A \cup B) \leq P(A) + P(B)$

Solution:

There are 4 possible cases.

1. A, B are disjoint
2. $A \subset B$
3. $B \subset A$
4. A, B have some common events between them. In other words $A \cap B = C \neq \varnothing$

Case 1: If A, B are disjoint then $A \cup B = A + B$ by set theory. Now apply the probability operator on both sides we obtain that

$$P(A \cup B) = P(A + B)$$

Now, by Axiom 3, $P(A + B) = P(A) + P(B)$ hence the above becomes

$$P(A \cup B) = P(A) + P(B)$$

Case 2: If $A \subset B$ then $A \cup B = B$ by set theory. Now apply the probability operator on both sides we obtain that

$$P(A \cup B) = P(B)$$

But $P(B) \leq P(B) + P(A)$ since $A \in \Omega$ and so $P(A) \geq 0$ by axiom 2. Hence the above becomes

$$P(A \cup B) \leq P(B) + P(A) \tag{0}$$

Case 3: This is the same as case 2, just exchange A and B

case 4: Since, by set theory
$$A = A \cap B + A \cap B^c$$

Then apply Probability operator on both sides

$$P(A) = P(A \cap B + A \cap B^c)$$

But by set theory $A \cap B$ is disjoint from $A \cap B^c$, then by axiom 3 the above becomes

$$P(A) = P(A \cap B) + P(A \cap B^c) \tag{1}$$

Similarly, by set theory

$$B = B \cap A + B \cap A^c$$

Then apply Probability operator on both sides

$$P(B) = P(B \cap A + B \cap A^c)$$

But $B \cap A$ is disjoint from $B \cap A^c$, by set theory, then by axiom 3 the above becomes

$$P(B) = P(B \cap A) + P(B \cap A^c) \tag{2}$$

Now by set theory
$$A \cup B = A \cap B + A \cap B^c + B \cap A^c$$

Apply the probability operator on the above

$$P(A \cup B) = P(A \cap B + A \cap B^c + B \cap A^c)$$

But $A \cap B, A \cap B^c$, and $B \cap A^c$ are disjoint by set theory, then above can be written using axiom 3 as

$$P(A \cup B) = P(A \cap B) + P(A \cap B^c) + P(B \cap A^c) \tag{3}$$

Add (1)+(2)

$$P(A) + P(B) = P(A \cap B) + P(A \cap B^c) + P(B \cap A) + P(B \cap A^c)$$

subtract the above from (3)

$$\begin{aligned} P(A \cup B) - [P(A) + P(B)] = {} & [P(A \cap B) + P(A \cap B^c) + P(B \cap A^c)] - \\ & [P(A \cap B) + P(A \cap B^c) + P(B \cap A) + P(B \cap A^c)] \end{aligned}$$

Cancel terms (Arithmetic)

$$P(A \cup B) - [P(A) + P(B)] = -P(B \cap A)$$

or (algebra)

$$P(A \cup B) = P(A) + P(B) - P(B \cap A)$$

Since $B \cap A$ is an event in Ω then $P(B \cap A) \geq 0$ by axiom 2, hence the above can be written as

$$P(A \cup B) \leq P(A) + P(B) \tag{4}$$

conclusion: We have looked at all 4 possible cases, and found that $P(A \cup B) = P(A) + P(B)$ or $P(A \cup B) \leq P(A) + P(B)$, hence $P(A \cup B) \leq P(A) + P(B)$

Note: I tried, really tried, to find a method which would require me to use the hint given in the problem that if $A \subset B$, then $P(A) \leq P(B)$ but I did not need to use such a relationship in the above. But I still show a proof for this identity below

Given: $A \subset B$, Show $P(A) \leq P(B)$

proof:

$B = A \cup A^c$ by set theory

$P(B) = P(A \cup A^c)$ by applying probability to each side.

But A, A^c are disjoint by set theory, hence $P(A \cup A^c) = P(A) + P(A^c)$ by axiom 3.

Hence $P(B) = P(A) + P(A^c)$, or $P(A) = P(B) - P(A^c)$

But by axiom 2, $P(A^c) \geq 0$, hence $P(A) \leq P(B)$, QED

4. Let $X \sim binomial(n, p)$. Derive the mode of the probability mass function of X.

Figure 4.6: Problem 4

Given: X binomial r.v., i.e. $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, Find the mode. This is the value k for which $P(X = k)$ is maximum

The mode is where $P(X)$ is maximum. Consider 2 terms, when $X = k$, and $X = k - 1$, hence $P(X)$ will be increasing when $\frac{P(X=k)}{P(X=k-1)} > 1$

But

$$P(X = k - 1) = \binom{n}{k-1} p^{(k-1)} (1-p)^{n-(k-1)}$$

Hence

$$\begin{aligned}
\frac{P(X = k)}{P(X = k - 1)} &= \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{(k-1)} (1-p)^{n-(k-1)}} = \frac{\frac{n!}{(n-k)!\ (k)!} p^k (1-p)^{n-k}}{\frac{n!}{(n-k+1)!\ (k-1)!} p^{(k-1)} (1-p)^{n-(k-1)}} \\
&= \frac{(n-k+1)!\ (k-1)!}{(n-k)!\ (k)!} \frac{(1-p)}{p} \\
&= \frac{(n-k)}{k} \frac{(1-p)}{p}
\end{aligned}$$

so $P(X)$ is getting larger when $\frac{(n-k)}{k}\frac{(1-p)}{p} > 1$ or

$$\begin{aligned}(n-k)(1-p) &> kp\\ n-np-k+kp &> kp\\ np+p &> k\\ p(1+n) &> k\end{aligned}$$

So as long as $k < p(1+n)$, pmf is increasing. Since k is an integer, then we need the largest integer such that it is $< p(1+n)$, hence

$$k = \lfloor p(1+n) \rfloor$$

5. Suppose that a rare disease has an incidence of 1 in 1000. Assuming that members of the population are affected independently, Find the probability that two individuals are affected in a population of 100,000 by (a) using the relevant binomial random variable, and a) using the relevant Poisson random variable. In each case identify the random variable and its distribution clearly. [Leave your solutions in expression forms].

Figure 4.7: Problem 5

Given:

$P(D) = 1/1000$

members are affected independently

Find: probability 2 individuals are affected in population of size 100,000

part(a)

In Binomial random variable we ask: How many are infected in a trial of length n given that the probability of being infected in each trial to be p. Here we view each trial as testing an individual. Consider it a 'hit' if the individual is infected. The number of trials is $100,000$, which is n, and $p = 1/1000$.

Therefore, $\boxed{X = \text{how many are infected in population of } 100000}$

Hence the probability of getting $k = 2$ hits is, using binomial r.v. is ($k = 2$ in this case)

$$P(X=2) = \binom{n}{k} p^k (1-p)^{n-k}$$

or numerically

$$P(X=2) = \binom{100000}{2} 0.001^2 (1-0.001)^{100000-2}$$

(b)Using Poisson r.v. Poisson is a generalization of Binomial. X is the number of successes in infinite number of trials, but with the probability of success in each one trial going to zero in such a way that $np = \lambda$ .We compute $p(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, \ldots.$

Hence here $X =$ how many are infected as n gets very large and p , the probability of infection in each individual goes very small in such a way to keep np fixed at a parameter λ. Since here n is large and p is small, we approximate binomial to Poisson using $\lambda = np = 100000 \times 0.001 = 100.0$

Hence

$$p(X=2)=\frac{100^2}{2!}e^{-100}$$

ps. computing a numerical value for the above, shows that using Binomial model, we obtain $P(X=2)$

```
In[5]:= n = 100000; k = 2; p = 0.001;
        Binomial[n, k] p^k (1 - p)^(n-k)

Out[6]= 1.77279 × 10^-40
```

Figure 4.8: Binomial model

and using Poisson model

```
In[12]:= λ = n p;
         λ^k/k! Exp[-λ]

Out[13]= 1.86004 × 10^-40
```

Figure 4.9: Poisson model

I am not sure, these are such small values, this means there is almost no chance of finding 2 individuals infected in a population of 100,000? I would have expected to see a much higher probability than the above. I do not see what I am doing wrong if anything.

4.1.1 Graded

16/20

16/20

QUIZ 1 MATH 502AB Fall 2007

Name (please print) Nasser Abbasi

1. Consider a sequence of days, and let R_i denote the event that it rains on day i. Let $P(R_0) = p$ (rain today), $P(R_i|R_{i-1}) = \alpha$, and $P(R_i^c|R_{i-1}^c) = \beta$. Suppose further that only today's weather is relevant to predicting tomorrow's; that is, $P(R_i|R_{i-1} \cap \cdots \cap R_0) = P(R_i|R_{i-1})$. What is the probability that it rains n days from now? What happens as n approaches infinity?

Answer:
Given:

1. R_i ,Event that it rains on day i
2. R_i^c ,Event that it does not rain on day i
3. $P(R_0) = p$, Probability of rain on day 0
4. $P(R_i|R_{i-1}) = \alpha$,Probability of rain on day i given it rained on day $i-1$
5. $P\left(R_i^c|R_{i-1}^c\right) = \beta$,Probability of no rain on day i given it did not rain on day $i-1$
6. Only today's weather is relevant to predicting tomorrow rain

Find:
Probability of rain in n days and what happen as $n \to \infty$
Solution:
Consider the experiment that generates today's weather. Hence possible outcomes can be divided into 2 disjoint events: rain and no rain (A day can either be rainy or not, hence this division contains all possible outcomes).
Hence

$$\Omega = \{R_0, R_0^c\}$$

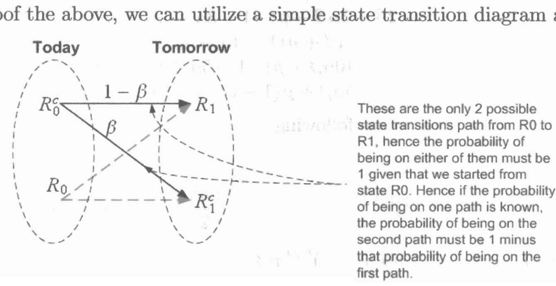
Now using the law of total probability, we write

$$P(R_1) = P(R_1|R_0)\,P(R_0) + P(R_1|R_0^c)\,P(R_0^c) \tag{1}$$

But

$$\begin{aligned} P(R_1|R_0^c) &= 1 - P(R_1|R_0) \\ &= 1 - \beta \end{aligned} \tag{2}$$

Note: To proof the above, we can utilize a simple state transition diagram as follows



1

Now, substitute (2) into (1) and given that $P(R_1|R_0) = \alpha$ and $P(R_0) = p$ and $P(R_0^c) = 1-p$, then (1) becomes

$$P(R_1) = \alpha p + (1-\beta)(1-p) \tag{3}$$

Now we can recursively apply the above to find probability of rain on the day after tomorrow. Let $R_0 \to R_1$ and $R_1 \to R_2$, hence the above (1) becomes

$$P(R_2) = P(R_2|R_1)\,P(R_1) + P(R_2|R_1^c)\,P(R_1^c) \tag{4}$$

Now using (3) for $P(R_1)$, and given that $P(R_2|R_1) = \alpha$ (This probability does not change, since we are told only today's weather is relevant), and given that $P(R_1^c) = (1 - P(R_1))$ and that $P(R_2|R_1^c) = (1-\beta)$, then (4) becomes

$$\begin{aligned}
P(R_2) &= \alpha\overbrace{[\alpha p + (1-\beta)(1-p)]}^{P(R_1)} + (1-\beta)\overbrace{(1 - [\alpha p + (1-\beta)(1-p)])}^{P(R_1^c)} \\
&= p + \alpha + \beta - 2p\alpha - 2p\beta - \beta^2 - \alpha\beta + p\alpha^2 + p\beta^2 + 2p\alpha\beta \\
&= p(1 - 2\alpha - 2\beta + 2\alpha\beta) + \alpha + \beta - \alpha\beta + (p\alpha^2 + p\beta^2 - \beta^2) \\
&= p(1 - 2\alpha - 2\beta + 2\alpha\beta) + \alpha + \beta - \alpha\beta + [\text{terms with higher powers in } \alpha \text{ and } \beta]
\end{aligned}$$

We see that as we continue with the above process, terms will be generated with the form (something)$\times\alpha^m$ and (something)$\times\beta^r$, where the powers m, r are getting larger and larger as n gets larger. But since $\alpha, \beta < 1$, hence all these terms go to zero. So we only need to look at the terms which do not contain a product of $\alpha's$ and product of $\beta's$

No!

Hence the above reduces

In a geometric series $\sum_{i=1}^{n} a^n$ $a<1$

The terms go to zero, but the sum doesn't!

$$P(R_2) \approx p(1 - 2\alpha - 2\beta + 2\alpha\beta) + \alpha + \beta - \alpha\beta$$

There is a pattern here, to see it more clearly, I generated more $P(R_i)$ for $i = 3,4,5,6,7$ using a small piece of code and removed all terms of higher powers of α, β as described above, and I get the following table

i	$P(R_i)$
0	p
1	$1 - \beta + p(-1 + \alpha + \beta)$
2	$\alpha + \beta - \alpha\beta + p(1 - 2\alpha - 2\beta + 2\alpha\beta)$
3	$1 - \alpha - 2\beta + 3\alpha\beta + p(-1 + 3\alpha + 3\beta - 6\alpha\beta)$
4	$2\alpha + 2\beta - 6\alpha\beta + p(1 - 4\alpha - 4\beta + 12\alpha\beta)$
5	$1 - 2\alpha - 3\beta + 10\alpha\beta + p(-1 + 5\alpha + 5\beta - 20\alpha\beta)$
6	$3\alpha + 3\beta - 15\alpha\beta + p(1 - 6\alpha - 6\beta + 30\alpha\beta)$

Hence the pattern can be seen as the following

$$\begin{aligned}
P(R_n) = \operatorname{mod}(n,2) + (-1)^{(n)}\left\lfloor\frac{n}{2}\right\rfloor\alpha + (-1)^{(n)}\left\lceil\frac{n}{2}\right\rceil\beta + (-1)^{(n+1)}\left(\sum_{i=1}^{n-1} i\right)\alpha\beta + \\
p\left((-1)^n + (-1)^{n+1}n\alpha + (-1)^{n+1}n\beta + (-1)^n\left[n^2 - n\right]\alpha\beta\right)
\end{aligned}$$

Where $\mathrm{mod}\,(n,2) = 0$ for even n and 1 for odd n, and $\lfloor \frac{n}{2} \rfloor$ means to round to nearest lower integer and $\lceil \frac{n}{2} \rceil$ means to round upper.

The above is valid for very large n.

As $n \to \infty$ $P\,(R_n)$ will reach a fixed value (I first though it will always go to 1, but that turned out not to be the case). I could not find an exact expression for $P\,(R_n)$ as $n \to \infty$, but I wrote a small program which simulates the above, and generates a table. Here is a table for few values as n gets large, these are all for $\alpha = .3, \beta = .6, p = .4$, notice that $P\,(R_n)$ fluctuates up and down from one day to the next as it converges to its limit.

```
In[29]:= TableForm[r, TableHeadings → {None, {"n", "P(Rn)"}}]
```

Out[29]//TableForm=

n	$P(R_n)$
0	0.4
1	0.28
2	0.316
3	0.3052
4	0.30844
5	0.307468
6	0.30776
7	0.307672
8	0.307698
9	0.30769
10	0.307693
11	0.307692
12	0.307692
13	0.307692
14	0.307692
15	0.307692
16	0.307692
17	0.307692

Need the theoretical
and limit
P(Rn)

13

$p(A_2 | A_1)$

2 Show that if the conditional probabilities exist, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

3

Given: Conditional probabilities exist
Show: $P(A_1 \cap A_2 \cap A_3 \cap \cdots A_n) = P(A_1) + P(A_2|A_1 \cap A_2) + \cdots + P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$
Solution:
Since Conditional probabilities exist, then we know that the following is true

$$P(X \cap Y) = P(X|Y)\, P(Y)$$

Let $X = A_n$ and $Y = A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}$ hence the above becomes

$$P(A_1 \cap A_2 \cap \cdots A_n) = P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\, P(A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Now apply the same idea to the last term above. In other words, we write

$$P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) = P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})\, P(A_1 \cap A_2 \cap \cdots \cap A_{n-2})$$

We repeat the process until we obtain $P(A_1 \cap A_2) = P(A_2|A_1)\, P(A_1)$
Hence, putting all the above together, we write

$$P(A_1 \cap A_2 \cap \cdots A_n) = P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})\, P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})$$
$$P(A_{n-2}|A_1 \cap A_2 \cap \cdots A_{n-3}) \cdots P(A_2|A_1)\, P(A_1)$$

The above is what is required to show (terms are just rewritten is reverse order from the problem statement, rearranging, we obtain

$$P(A_1 \cap A_2 \cap \cdots A_n) = P(A_1)\, P(A_2|A_1) \cdots P(A_{n-1}|A_1 \cap A_2 \cap \cdots \cap A_{n-2})\, P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

QED

need to give a inductive proof

3. Let A and B be arbitrary events. Use the three axioms of probability to show that

$$P(A \cup B) \leq P(A) + P(B).$$

Identify the axiom(s) that you use at each step. You are not allowed to use any theorems. [Hint: One way to show this is to first show that if C and D are events such that $C \subset D$, then $P(C) \leq P(D)$. Then, use this result to prove the result in the above display equation.]

Case 4 is sufficient other cases are not needed. You have some notation problems here!

Given:
Axioms of probability:

1. $P(\Omega) = 1$
2. if $A \subset \Omega$ then $P(A) \geq 0$
3. if A, B are disjoint events (i.e. $A \cap B = \varnothing$) then $P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$

Show that $P(A \cup B) \leq P(A) + P(B)$
Solution:
There are 4 possible cases.

1. A, B are disjoint
2. $A \subset B$
3. $B \subset A$
4. A, B have some common events between them. In other words $A \cap B = C \neq \varnothing$

Case 1: If A, B are disjoint then $A \cup B = A + B$ by set theory. Now apply the probability operator on both sides we obtain that

?

$$P(A \cup B) = P(A + B)$$

$P(A \cup B) = P(A) + P(B)$

Now, by Axiom 3, $P(A + B) = P(A) + P(B)$ hence the above becomes

$$P(A \cup B) = P(A) + P(B)$$

Case 2: If $A \subset B$ then $A \cup B = B$ by set theory. Now apply the probability operator on both sides we obtain that

b/2.

$$P(A \cup B) = P(B)$$

But $P(B) \leq P(B) + P(A)$ since $A \in \Omega$ and so $P(A) \geq 0$ by axiom 2. Hence the above becomes

$$\boxed{P(A \cup B) \leq P(B) + P(A)} \tag{0}$$

Case 3: This is the same as case 2, just exchange A and B
case 4: Since, by set theory

$$A = A \cap B + A \cap B^c$$

5

?

Then apply Probability operator on both sides

$$P(A) = P(A \cap B + A \cap B^c)$$

But by set theory $A \cap B$ is disjoint from $A \cap B^c$, then by axiom 3 the above becomes

$$P(A) = P(A \cap B) + P(A \cap B^c) \tag{1}$$

Similarly, by set theory

$$B = B \cap A + B \cap A^c$$

?

Then apply Probability operator on both sides

$$P(B) = P(B \cap A + B \cap A^c)$$

But $B \cap A$ is disjoint from $B \cap A^c$, by set theory, then by axiom 3 the above becomes

$$P(B) = P(B \cap A) + P(B \cap A^c) \tag{2}$$

Now by set theory

$$A \cup B = A \cap B + A \cap B^c + B \cap A^c$$

?

Apply the probability operator on the above

$$P(A \cup B) = P(A \cap B + A \cap B^c + B \cap A^c)$$

But $A \cap B, A \cap B^c$, and $B \cap A^c$ are disjoint by set theory, then above can be written using axiom 3 as

$$P(A \cup B) = P(A \cap B) + P(A \cap B^c) + P(B \cap A^c) \tag{3}$$

Add (1)+(2)

$$P(A) + P(B) = P(A \cap B) + P(A \cap B^c) + P(B \cap A) + P(B \cap A^c)$$

subtract the above from (3)

$$\begin{aligned} P(A \cup B) - [P(A) + P(B)] = [&P(A \cap B) + P(A \cap B^c) + P(B \cap A^c)] - \\ &[P(A \cap B) + P(A \cap B^c) + P(B \cap A) + P(B \cap A^c)] \end{aligned}$$

Cancel terms (Arithmetic)

$$P(A \cup B) - [P(A) + P(B)] = -P(B \cap A)$$

or (algebra)

$$P(A \cup B) = P(A) + P(B) - P(B \cap A)$$

Since $B \cap A$ is an event in Ω then $P(B \cap A) \geq 0$ by axiom 2, hence the above can be written as

$$\boxed{P(A \cup B) \leq P(A) + P(B)} \tag{4}$$

conclusion: We have looked at all 4 possible cases, and found that $P(A \cup B) = P(A) + P(B)$ or $P(A \cup B) \leq P(A) + P(B)$, hence $P(A \cup B) \leq P(A) + P(B)$

Note: I tried, really tried, to find a method which would require me to use the hint given in the problem that if $A \subset B$, then $P(A) \leq P(B)$ but I did not need to use such a relationship in the above. But I still show a proof for this identity below

Given: $A \subset B$, Show $P(A) \leq P(B)$

proof:

$B = A \cup A^c$ by set theory

$P(B) = P(A \cup A^c)$ by applying probability to each side.

But A, A^c are disjoint by set theory, hence $P(A \cup A^c) = P(A) + P(A^c)$ by axiom 3.

Hence $P(B) = P(A) + P(A^c)$, or $P(A) = P(B) - P(A^c)$

But by axiom 2, $P(A^c) \geq 0$, hence $P(A) \leq P(B)$, QED

4. Let $X \sim binomial(n, p)$. Derive the mode of the probability mass function of X.

Given: X binomial r.v., i.e. $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, Find the mode. This is the value k for which $P(X = k)$ is maximum

The mode is where $P(X)$ is maximum. Consider 2 terms, when $X = k$, and $X = k-1$, hence $P(X)$ will be increasing when $\frac{P(X=k)}{P(X=k-1)} > 1$

But

$$P(X = k-1) = \binom{n}{k-1} p^{(k-1)} (1-p)^{n-(k-1)}$$

Hence

$$\begin{aligned}
\frac{P(X=k)}{P(X=k-1)} &= \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{(k-1)} (1-p)^{n-(k-1)}} = \frac{\frac{n!}{(n-k)!\ (k)!} p^k (1-p)^{n-k}}{\frac{n!}{(n-k+1)!\ (k-1)!} p^{(k-1)} (1-p)^{n-(k-1)}} \\
&= \frac{(n-k+1)!\ (k-1)!}{(n-k)!\ (k)!} \frac{(1-p)}{p} \\
&= \frac{(n-k)}{k} \frac{(1-p)}{p}
\end{aligned}$$

so $P(X)$ is getting larger when $\frac{(n-k)}{k}\frac{(1-p)}{p} > 1$ or

$$\begin{aligned}
(n-k)(1-p) &> kp \\
n - np - k + kp &> kp \\
np + p &> k \\
p(1+n) &> k
\end{aligned}$$

So as long as $k < p(1+n)$, pmf is increasing. Since k is an integer, then we need the largest integer such that it is $< p(1+n)$, hence

$$\boxed{k = \lfloor p(1+n) \rfloor}$$

5. Suppose that a rare disease has an incidence of 1 in 1000. Assuming that members of the population are affected independently, Find the probability that two individuals are affected in a population of 100,000 by (a) using the relevant binomial random variable, and a) using the relevant Poisson random variable. In each case identify the random variable and its distribution clearly. [Leave your solutions in expression forms].

Given:
$P(D) = 1/1000$
members are affected independently
Find: probability 2 individuals are affected in population of size 100,000
part(a)

In Binomial random variable we ask: How many are infected in a trial of length n given that the probability of being infected in each trial to be p. Here we view each trial as testing an individual. Consider it a 'hit' if the individual is infected. The number of trials is $100,000$, which is n, and $p = 1/1000$.

Therefore, $\boxed{X = \text{how many are infected in population of 100000}}$

Hence the probability of getting $k = 2$ hits is, using binomial r.v. is ($k = 2$ in this case)

$$\boxed{P(X=2) = \binom{n}{k} p^k (1-p)^{n-k}}$$

or numerically

$$P(X=2) = \binom{100000}{2} 0.001^2 (1-0.001)^{100000-2}$$

(b)Using Poisson r.v. Poisson is a generalization of Binomial. X is the number of successes in infinite number of trials, but with the probability of success in each one trial going to zero in such a way that $np = \lambda$ .We compute $p(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, 2, \ldots.$

Hence here $X =$ how many are infected as n gets very large and p , the probability of infection in each individual goes very small in such a way to keep np fixed at a parameter λ. Since here n is large and p is small, we approximate binomial to Poisson using $\lambda = np = 100000 \times 0.001 = 100.0$

Hence

$$\boxed{p(X=2) = \frac{100^2}{2!} e^{-100}}$$

ps. computing a numerical value for the above, shows that using Binomial model, we obtain $P(X=2)$

```
In[5]:= n = 100000; k = 2; p = 0.001;
        Binomial[n, k] p^k (1 - p)^(n-k)

Out[6]= 1.77279 × 10^-40
```

and using Poisson model

```
In[12]:= λ = n p;
         λ^k
         --- Exp[-λ]
         k!
```

$$\text{Out[13]= } 1.86004 \times 10^{-40}$$

I am not sure, these are such small values, this means there is almost no chance of finding 2 individuals infected in a population of 100,000? I would have expected to see a much higher probability than the above. I do not see what I am doing wrong if anything.

4.2 Quiz 2

Local contents

QUIZ 2 MATH 502AB Fall 2007

Name (please print) Nasser Abbasi

1. Use the fact that $\Gamma(1/2)=\sqrt{\pi}$ to show that if n is an odd integer, then

$$\Gamma(n/2)=\frac{\sqrt{\pi}(n-1)!}{2^{n-1}\left(\frac{n-1}{2}\right)!}.$$

Figure 4.10: Problem one

By definition,

$$\Gamma(z)=\int_0^\infty t^{z-1}e^{-t}dt$$

Hence

$$\Gamma\left(\frac{n}{2}\right)=\int_0^\infty t^{\left(\frac{n}{2}-1\right)}e^{-t}dt$$

When $n=1$, we are told that

$$\Gamma\left(\frac{1}{2}\right)=\int_0^\infty t^{\left(\frac{1}{2}-1\right)}e^{-t}dt=\int_0^\infty t^{\left(-\frac{1}{2}\right)}e^{-t}dt=\sqrt{\pi}$$

For $n=3$, we have

$$\begin{aligned}\Gamma\left(\frac{3}{2}\right)&=\int_0^\infty t^{\left(\frac{3}{2}-1\right)}e^{-t}dt\\&=\int_0^\infty t^{\left(\frac{1}{2}\right)}e^{-t}dt\end{aligned}$$

Now do integration by parts,

$$\begin{aligned}\Gamma\left(\frac{3}{2}\right)&=\int_0^\infty \overbrace{t^{\left(\frac{1}{2}\right)}}^{u}\overbrace{e^{-t}dt}^{dv}\\&=\left[t^{\frac{1}{2}}\left(-e^{-t}\right)\right]_0^\infty-\int_0^\infty\frac{1}{2}t^{\left(-\frac{1}{2}\right)}\left(-e^{-t}\right)dt\\&=-\left[t^{\frac{1}{2}}e^{-t}\right]_0^\infty+\int_0^\infty\frac{1}{2}t^{\left(-\frac{1}{2}\right)}e^{-t}dt\end{aligned}$$

But $\left[t^{\frac{1}{2}}e^{-t}\right]_0^\infty=[0-0]=0$ and the above becomes

$$\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\int_0^\infty t^{\left(-\frac{1}{2}\right)}e^{-t}dt$$

But $\int_0^\infty t^{\left(-\frac{1}{2}\right)}e^{-t}dt=\Gamma\left(\frac{1}{2}\right)$, hence the above becomes

$$\boxed{\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)}$$

Now do the same for $n=5$

$$\begin{aligned}
\Gamma\left(\frac{5}{2}\right) &= \int_0^\infty t^{\left(\frac{5}{2}-1\right)}e^{-t}dt\\
&= \int_0^\infty \overbrace{t^{\left(\frac{3}{2}\right)}}^{u}\overbrace{e^{-t}dt}^{dv}\\
&= -\left[t^{\frac{3}{2}}e^{-t}\right]_0^\infty - \int_0^\infty \frac{3}{2}t^{\left(\frac{1}{2}\right)}\left(-e^{-t}\right)dt\\
&= \frac{3}{2}\int_0^\infty t^{\left(\frac{1}{2}\right)}e^{-t}dt
\end{aligned}$$

But $\int_0^\infty t^{\left(\frac{1}{2}\right)}e^{-t}dt$ we found from above to be $\Gamma\left(\frac{3}{2}\right)$, hence

$$\Gamma\left(\frac{5}{2}\right)=\frac{3}{2}\Gamma\left(\frac{3}{2}\right)$$

But $\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$ from above, hence

$$\begin{aligned}
\Gamma\left(\frac{5}{2}\right) &= \frac{3}{2}\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\\
&= \frac{3}{2}\frac{1}{2}\sqrt{\pi}
\end{aligned}$$

Continuing this way, we find that $\Gamma\left(\frac{7}{2}\right)=\frac{5}{2}\frac{3}{2}\frac{1}{2}\Gamma\left(\frac{1}{2}\right)$, and hence in general

$$\Gamma\left(\frac{n}{2}\right)=\frac{(n-2)}{2}\frac{(n-4)}{2}\frac{(n-6)}{2}\cdots\frac{5}{2}\frac{3}{2}\frac{1}{2}\sqrt{\pi} \tag{1}$$

Now,

$$(n-1)!=(n-1)(n-2)(n-3)(n-4)(n-5)(n-6)\cdots 5\times 4\times 3\times 2\times 1$$

Hence from above we see that

$$(n-2)(n-4)(n-6)\cdots 5\times 3\times 1=\frac{(n-1)!}{(n-1)(n-3)(n-5)\cdots 4\times 2\times 1}$$

Therefore (1) can be written as

$$\begin{aligned}
\Gamma\left(\frac{n}{2}\right) &= \left(\frac{(n-1)!}{(n-1)(n-3)(n-5)\cdots 4\times 2\times 1}\right)\overbrace{\left(\frac{1}{2}\frac{1}{2}\frac{1}{2}\cdots\frac{1}{2}\frac{1}{2}\right)}^{\text{There are } \frac{n-1}{2} \text{ such terms}}\sqrt{\pi}\\
&= \frac{(n-1)!}{(n-1)(n-3)(n-5)\cdots 4\times 2}\frac{1}{2^{\left(\frac{n-1}{2}\right)}}\sqrt{\pi}
\end{aligned} \tag{2}$$

But

$$2^{\left(\frac{n-1}{2}\right)}=\left(2^{n-1}\right)\left(2^{-\frac{n-1}{2}}\right)=\frac{2^{n-1}}{2^{\frac{n-1}{2}}}$$

Hence (2) becomes

$$\Gamma\left(\frac{n}{2}\right)=\left(\frac{(n-1)!}{\frac{1}{2^{\frac{n-1}{2}}}(n-1)(n-3)(n-5)\cdots 4\times 2}\right)\frac{1}{2^{n-1}}\sqrt{\pi}$$

But there are $\frac{n-1}{2}$ terms in the expression $(n-1)(n-3)(n-5)\cdots 4\times 2$ in the denominator above and we have $\frac{n-1}{2}$ number of $\frac{1}{2}$ sitting there, which we can distribute now below each terms to obtain

$$\Gamma\left(\frac{n}{2}\right)=\left(\frac{(n-1)!}{\frac{(n-1)}{2}\frac{(n-3)}{2}\frac{(n-5)}{2}\cdots\frac{4}{2}\times\frac{2}{2}}\right)\frac{1}{2^{n-1}}\sqrt{\pi} \tag{3}$$

But

$$\begin{aligned}\left(\frac{n-1}{2}\right)! &= \left(\frac{n-1}{2}\right)\left(\frac{n-1}{2}-1\right)\left(\frac{n-1}{2}-2\right)\cdots\times 4\times 3\times 2\times 1\\ &= \left(\frac{n-1}{2}\right)\left(\frac{n-3}{2}\right)\left(\frac{n-5}{2}\right)\cdots\times 4\times 2\end{aligned}$$

Compare the above to the denominator term in (3) we see it is the same. Hence (3) can be written as

$$\boxed{\Gamma\left(\frac{n}{2}\right)=\frac{(n-1)!}{\left(\frac{n-1}{2}\right)!}\frac{1}{2^{n-1}}\sqrt{\pi}}$$

Which is what we are asked to show.

2. If $U\sim Uniform[-1,1]$, find the density of $Z=U^2$.

Figure 4.11: Problem 2

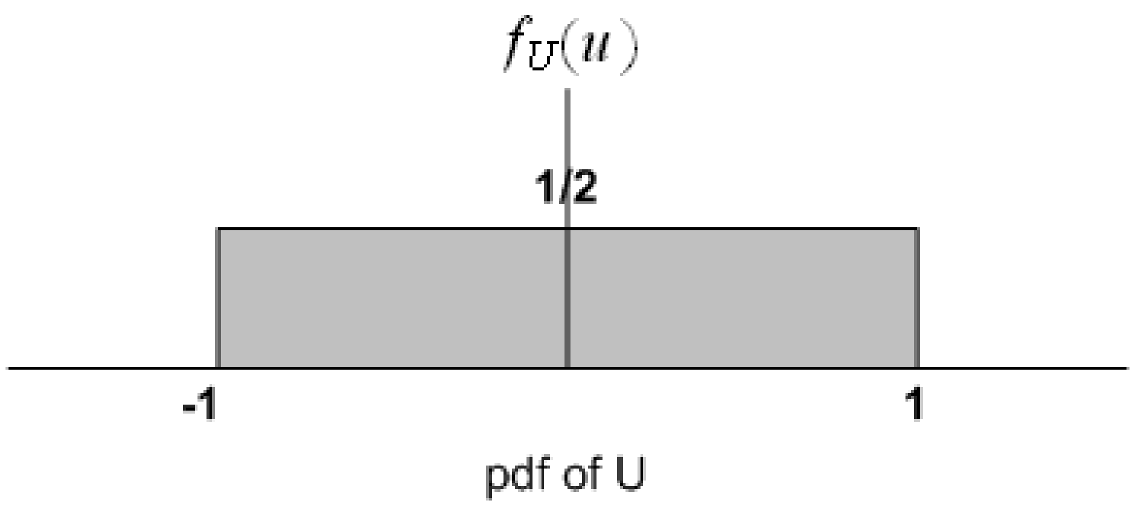


Figure 4.12: p2PDF

Since U is continuous r.v., we start with the CDF of Z

$$\begin{aligned}F_Z(z) &= P(Z\leq z)\\ &= P\left(U^2\leq z\right)\\ &= P\left(-\sqrt{z}\leq U\leq\sqrt{z}\right)\end{aligned} \tag{1}$$

But since $F_U'(u) = f_U(u)$, then we know that $P(a \leq U \leq b) = \int_a^b f_U(x)\,dx \rightarrow F_U(b) - F_U(a)$

Hence RHS of (1) becomes

$$F_Z(z) = F_U(\sqrt{z}) - F_U(-\sqrt{z})$$

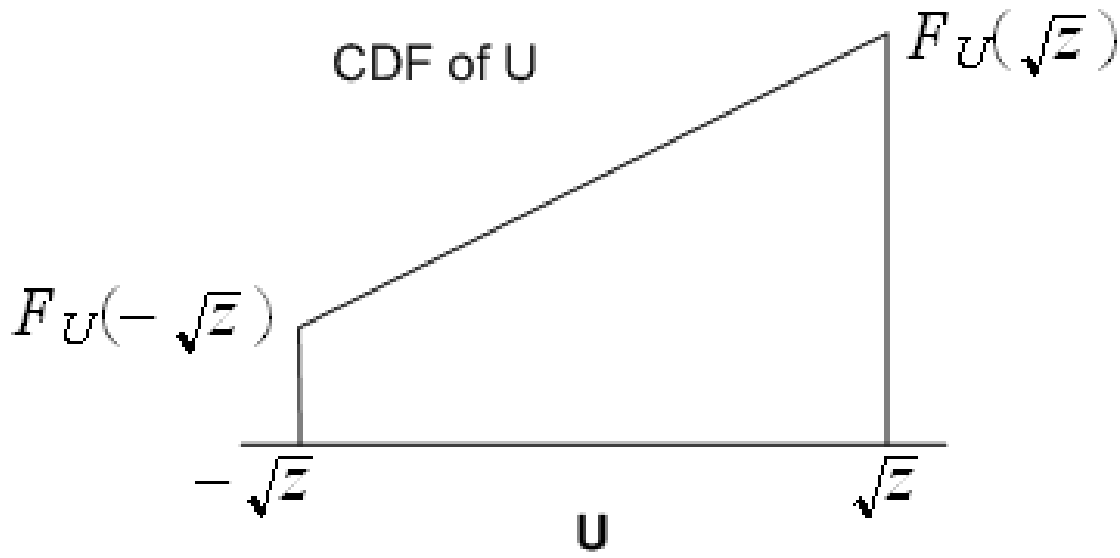


Figure 4.13: p2CDF

Therefore, taking derivatives with respect to z we obtain

$$\begin{aligned} f_Z(z) &= f_U(\sqrt{z})\,\frac{d}{dz}\sqrt{z} - f_U(-\sqrt{z})\,\frac{d}{dz}(-\sqrt{z}) \\ &= \frac{1}{2}z^{\frac{-1}{2}} f_U(\sqrt{z}) + f_U(-\sqrt{z})\,\frac{1}{2}z^{\frac{-1}{2}} \\ &= \frac{1}{2}z^{\frac{-1}{2}}\left(f_U(\sqrt{z}) + f_U(-\sqrt{z})\right) \end{aligned}$$

Since U is uniform, hence $f_U(a) = f_U(-a)$, hence the above becomes

$$\begin{aligned} f_Z(z) &= \frac{1}{2}z^{\frac{-1}{2}}\left(2f_U(\sqrt{z})\right) \\ &= \frac{1}{\sqrt{z}}\,f_U(\sqrt{z}) \end{aligned}$$

Now I need to determine the limits of $f_Z(z)$ and the shape. f_U is defined for real arguments from -1 to $+1$. i.e. f_U is real valued function of real arguments. Hence if z was negative then $\sqrt{z}$ will be complex, and so this will not be allowed. Hence we have to restrict $z \geq 0$. But now we observe that $z = 0$ is not possible, since we will have $\frac{1}{0}$ term, so this means z is strictly larger than zero. So

$$f_Z(z) = \frac{1}{\sqrt{z}}\,f_U(\sqrt{z}) \qquad z > 0$$

But we know that $f_U(x) = \frac{1}{2}$ for up to $x = 1$, hence this means when $\sqrt{z} > 1$ then $f_U(\sqrt{z}) = 0$, when means when $z > 1$ then $f_U(\sqrt{z}) = 0$

Hence we now write

$$f_Z(z) = \begin{cases} \frac{1}{\sqrt{z}}\frac{1}{2} & 0 < z \leq 1 \\ 0 & z > 1 \\ undefined & z \leq 0 \end{cases}$$

Here is a plot

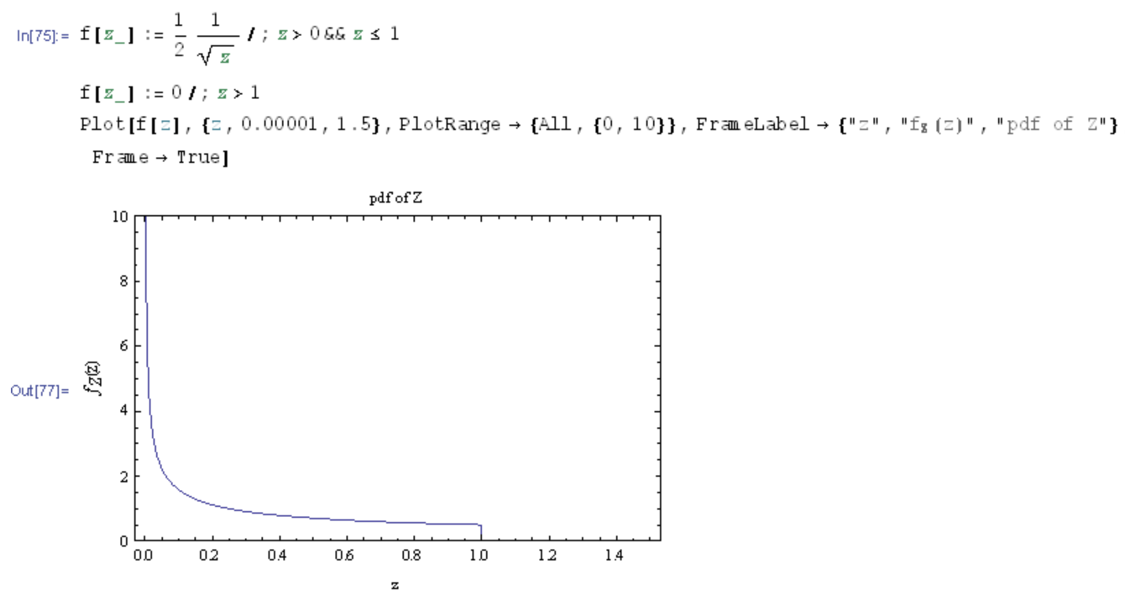


Figure 4.14: p2PDFfz

3. The following five numbers were randomly generated from the uniform random variable on (0,1):

0.0153 0.7468 0.4451 0.9318

Using these numbers generate five random numbers from the geometric random variable with parameter $p = 1/3$. Very briefly explain how you obtain your solution.

Figure 4.15: Problem 3

I explain the idea behind obtaining a discrete random number from a continues random number by the following diagram below. We assume that the discrete random number belongs to some distribution. In this example, we are told what the distribution is. We know that the CDF for geometric random variable is given by

$$F_K(k) = 1 - (1-p)^k$$

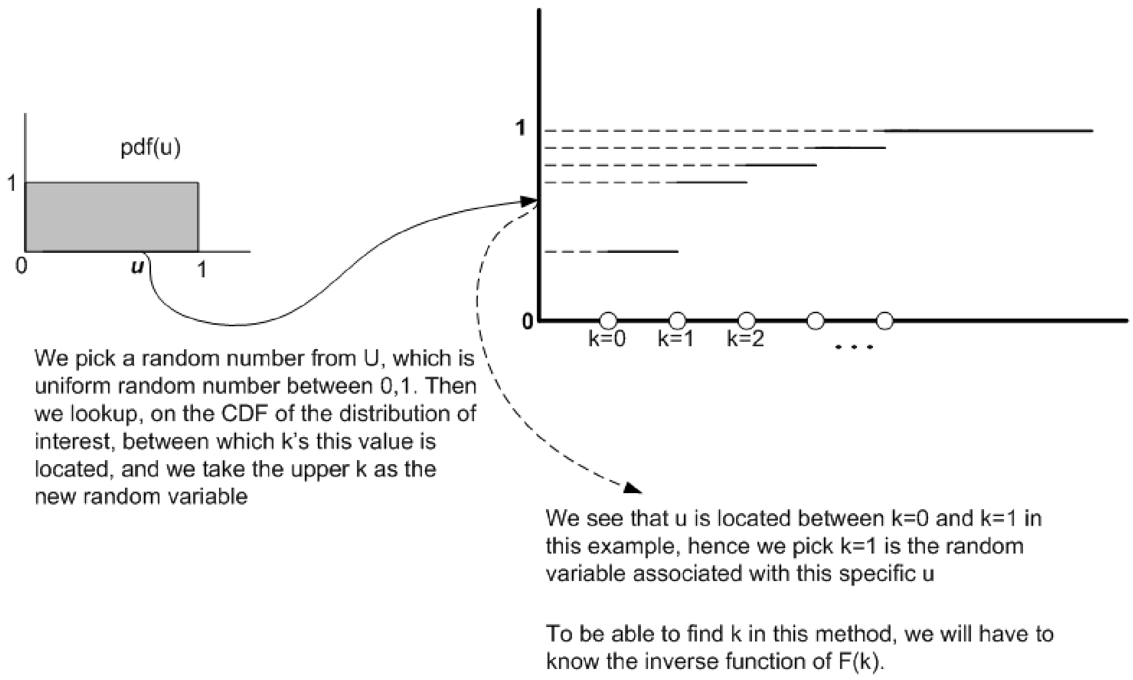


Figure 4.16: CDF for geometric random variable

We see from the above diagram, that once we are given u we need to find k which satisfy the following identity

$$F_K(k-1) < u \leq F_K(k)$$

Or in other words

$$1-(1-p)^{k-1} < u \leq 1-(1-p)^k$$

The specific discrete value k which will satisfy the above, is the random variable we want, which belong to the geometeric distribution.

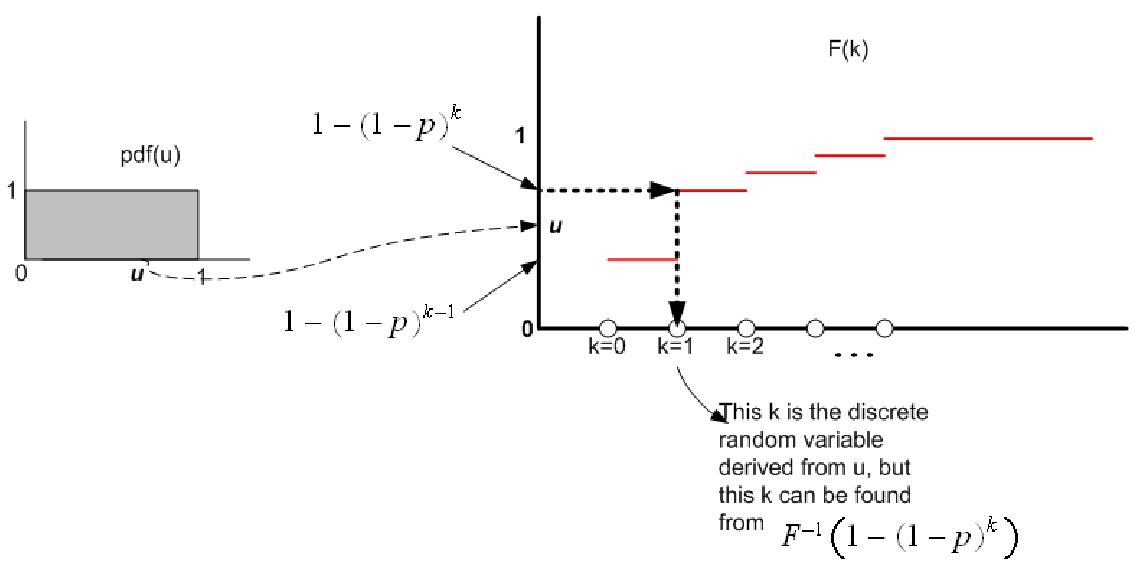


Figure 4.17: geometeric distribution

Now when $u = 0.0153$, and since $p = \frac{1}{3}$, we have

$$1-\left(1-\frac{1}{3}\right)^{k-1} < 0.0153 \leq 1-\left(1-\frac{1}{3}\right)^k$$

for $k = 1$, we have

$$0 < 0.0153 \leq 0.333\,33 \qquad YES$$

$\boxed{\text{Hence } k = 1 \text{ is the random variable associated with } u = 0.0153}$

Now let us do $u = 0.7468$

for $k = 1$, we have

$$0 < 0.7468 \leq 0.333\,33 \qquad NO$$

try $k = 2$

$$\begin{aligned} 1 - \left(1 - \frac{1}{3}\right)^1 &< 0.7468 \leq 1 - \left(1 - \frac{1}{3}\right)^2 \\ 0.333\,33 &< 0.7468 \leq 0.555\,56 \qquad NO \end{aligned}$$

try $k = 3$

$$\begin{aligned} 1 - \left(1 - \frac{1}{3}\right)^2 &< 0.7468 \leq 1 - \left(1 - \frac{1}{3}\right)^3 \\ 0.555\,56 &< 0.7468 \leq 0.703\,7 \qquad NO \end{aligned}$$

try $k = 4$

$$\begin{aligned} 1 - \left(1 - \frac{1}{3}\right)^3 &< 0.7468 \leq 1 - \left(1 - \frac{1}{3}\right)^4 \\ 0.703\,7 &< 0.7468 \leq 0.802\,47 \qquad YES \end{aligned}$$

$\boxed{\text{Hence } k = 4 \text{ is the random variable associated with } u = 0.7468}$

Now let us do $u = 0.4451$

We see from the above, that this will have $k = 2$ since for $k = 2$ the intervals is $0.333\,33 < u \leq 0.555\,56$

$\boxed{\text{Hence } k = 2 \text{ is the random variable associated with } u = 0.4451}$

Now let us do $u = 0.9318$

From above, we see that this will have a k larger than 4, so we do not need to try from the start, we can start trying from $k = 5$

try $k = 5$

$$\begin{aligned} 1 - \left(1 - \frac{1}{3}\right)^4 &< 0.9318 \leq 1 - \left(1 - \frac{1}{3}\right)^5 \\ 0.802\,47 &< 0.9318 \leq 0.868\,31 \qquad NO \end{aligned}$$

try $k = 6$

$$1-\left(1-\frac{1}{3}\right)^5 < 0.9318 \leq 1-\left(1-\frac{1}{3}\right)^6$$
$$0.868\,31 < 0.9318 \leq 0.912\,21 \qquad NO$$

try $k=7$

$$1-\left(1-\frac{1}{3}\right)^6 < 0.9318 \leq 1-\left(1-\frac{1}{3}\right)^7$$
$$0.912\,21 \ < 0.9318 \leq 0.941\,47 \qquad YES$$

$\boxed{\text{Hence } k=7 \text{ is the random variable associated with } u=0.9318}$

Hence result is

u	k
0.0153	1
0.4451	2
0.7468	4
0.9318	7

of course one would write a program to do this.

4. Three players play 10 independent rounds of a game, and each player has probability 1/3 of winning each round. **(a)** Find the joint distribution of the numbers of games won by each of the three players. **(b)** Identify the distribution of the number of games won by player one.

Figure 4.18: Problem 4

(a)

Let $P(x_i = n_i)$ means probability of player i winning n_i rounds.

We have 3 players, and a total of 10 rounds. Let the players be called x_1, x_2, x_3. Let the number of games WON by x_1 be n_1, and number of games won by x_2 be n_2, and number of games won by x_3 be n_3.

Since we have 10 rounds, then we must have 10 wins as well. (some one must win). Hence we have 10 wins and 3 ways to split it, where each 'bucket' is of different size. So this is a multi set selection. called multinomial in the book using proposition B in chapter 1, we see that the total number of ways the games can be won is $\binom{10}{n_1 n_2 n_3}$

But we need to find the probability of each one such combination. So we need to multiply the above by the probability each player wins the number of the games they happened to win, which is $P(x_i = n_i) = p^{n_i}$, but $p = \frac{1}{3}$ for each player to win a round. Hence we write

$$\begin{aligned}
P(x_1 = n_1, x_2 = n_2, x_3 = n_3) &= \binom{10}{n_1 n_2 n_3}\left(\frac{1}{3}\right)^{n_1}\left(\frac{1}{3}\right)^{n_2}\left(\frac{1}{3}\right)^{n_3} \\
&= \frac{10!}{n_1!\ n_2!\ n_3!}\left(\frac{1}{3}\right)^{n_1+n_2+n_3} \\
&= \frac{10!}{n_1!\ n_2!\ n_3!}\left(\frac{1}{3}\right)^{10}
\end{aligned}$$

So the above is the joint probability that p_1 wins n_1 rounds and p_2 wins n_2 rounds and p_3 wins n_3 rounds.

(b)We need to find $P(x_1 = n_1)$, i.e. the probability of first player winning n_1 rounds.

$$P(x_1 = n_1) = \sum_{\substack{n_2=0,1,\cdots 10\\ n_3=0,1,\cdots 10}} P(x_1 = n_1, x_2 = n_2, x_3 = n_3)$$

To simplify, let me write $P(n_1, n_2, n_3)$ instead, where the position of the n implies the player. So $p(0, 1, 9)$ means player one wins zero rounds and player 2 wins 1 round and player 3 wins 9 rounds.

So the above becomes

$$P(x_1 = n_1) = \sum_{\substack{n_2=0,1,\cdots 10\\ n_3=0,1,\cdots 10}} P(n_1, n_2, n_3)$$

But since $n_1 = 10 - (n_2 + n_3)$ we see that we only need to count those terms in the above sum when this is true. i.e. we do not need to count a term such as $p(1, 0, 0)$ since this is zero probability of happening. Now we write

$$P(x_1 = n_1) = P(n_1, 10 - n_1, 0) + P(n_1, 9 - n_1, 1) + P(n_1, 8 - n_1, 2) + \cdots P(n_1, 0, 10 - n_1)$$

For example,

$P(x_1 = 0) = P(0, 10, 0) + P(0, 9, 1) + P(0, 8, 2) + P(0, 7, 3) + P(0, 6, 4) + P(0, 5, 5) +$

$P(0, 4, 6) + P(0, 3, 7) + P(0, 2, 8) + P(0, 1, 9) + P(0, 0, 10)$

But $P(0, 10, 0) = P(0, 0, 10)$ and $P(0, 9, 1) = P(0, 1, 9)$, etc.. so the above can be written as

$$P(x_1 = 0) = 2P(0, 10, 0) + 2P(0, 9, 1) + 2P(0, 8, 2) + 2P(0, 7, 3) + 2P(0, 6, 4) + P(0, 5, 5)$$

$$\begin{aligned}
&= 2\frac{10!}{0!\ 10!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{0!\ 9!\ 1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{0!\ 8!\ 2!}\left(\frac{1}{3}\right)^{10} + \\
&2\frac{10!}{0!7!\ 3!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{0!\ 6!\ 4!}\left(\frac{1}{3}\right)^{10} + \frac{10!}{0!\ 5!\ 5!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{1.7342 \times 10^{-2}}
\end{aligned}$$

and

$$\begin{aligned}
P(x_1 = 1) &= P(1, 9, 0) + P(1, 8, 1) + P(1, 7, 2) + P(1, 6, 3) + P(1, 5, 4) + P(1, 4, 5) \\
&+ P(1, 3, 6) + P(1, 2, 7) + P(1, 1, 8) + P(1, 0, 9) \\
&= 2P(1, 9, 0) + 2P(1, 8, 1) + 2P(1, 7, 2) + 2P(1, 6, 3) + 2P(1, 5, 4) \\
&= 2\frac{10!}{1!\ 9!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!\ 8!\ 1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!\ 7!\ 2!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!6!\ 3!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!\ 5!\ 4!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{8.6708 \times 10^{-2}}
\end{aligned}$$

and

$$
\begin{aligned}
P(x_1=2) = {} & P(2,8,0)+P(2,7,1)+P(2,6,2)+P(2,5,3)+P(2,4,4)+P(2,3,5)+\\
& P(2,2,6)+P(2,1,7)+P(2,0,8)\\
&= 2P(2,8,0)+2P(2,7,1)+2P(2,6,2)+2P(2,5,3)+P(2,4,4)\\
&= 2\frac{10!}{2!\ 8!\ 0!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{2!\ 7!\ 1!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{2!\ 6!\ 2!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{2!5!\ 3!}\left(\frac{1}{3}\right)^{10}+\frac{10!}{2!\ 4!\ 4!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{0.19509}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=3) = {} & P(3,7,0)+P(3,6,1)+P(3,5,2)+P(3,4,3)+P(3,3,4)+\\
& P(3,2,5)+P(3,1,6)+P(3,0,7)\\
&= 2P(3,7,0)+2P(3,6,1)+2P(3,5,2)+2P(3,4,3)\\
&= 2\frac{10!}{3!\ 7!\ 0!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{3!\ 6!\ 1!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{3!\ 5!\ 2!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{3!4!\ 3!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{0.26012}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=4) &= P(4,6,0)+P(4,5,1)+P(4,4,2)+P(4,3,3)+P(4,2,4)+P(4,1,5)+P(4,0,6)\\
&= 2P(4,6,0)+2P(4,5,1)+2P(4,4,2)+P(4,3,3)\\
&= 2\frac{10!}{4!\ 6!\ 0!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{4!\ 5!\ 1!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{4!\ 4!\ 2!}\left(\frac{1}{3}\right)^{10}+\frac{10!}{4!3!\ 3!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{0.22761}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=5) &= P(5,5,0)+P(5,4,1)+P(5,3,2)+P(5,2,3)+P(5,1,4)+P(5,0,5)\\
&= 2P(5,5,0)+2P(5,4,1)+2P(5,3,2)\\
&= 2\frac{10!}{5!\ 5!\ 0!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{5!4!\ 1!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{5!\ 3!\ 2!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{0.13656}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=6) &= P(6,4,0)+P(6,3,1)+P(6,2,2)+P(6,1,3)+P(6,0,4)\\
&= 2P(6,4,0)+2P(6,3,1)+P(6,2,2)\\
&= 2\frac{10!}{6!\ 4!\ 0!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{6!3!\ 1!}\left(\frac{1}{3}\right)^{10}+\frac{10!}{6!\ 2!\ 2!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{5.6902\times 10^{-2}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=7) &= P(7,3,0)+P(7,2,1)+P(7,1,2)+P(7,0,3)\\
&= 2P(7,3,0)+2P(7,2,1)\\
&= 2\frac{10!}{7!\ 3!\ 0!}\left(\frac{1}{3}\right)^{10}+2\frac{10!}{7!2!\ 1!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{1.6258\times 10^{-2}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=8) &= P(8,2,0)+P(8,1,1)+P(8,0,2)\\
&= 2P(8,2,0)+P(8,1,1)\\
&= 2\frac{10!}{8!\ 2!\ 0!}\left(\frac{1}{3}\right)^{10}+\frac{10!}{8!1!\ 1!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{3.0483\times 10^{-3}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=9) &= P(9,1,0)+P(9,0,1)\\
&= 2P(9,1,0)\\
&= 2\frac{10!}{9!\ 1!\ 0!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{3.387\times 10^{-4}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1=10) &= P(10,0,0)\\
&= \frac{10!}{10!\ 0!\ 0!}\left(\frac{1}{3}\right)^{10}\\
&= \boxed{1.6935\times 10^{-5}}
\end{aligned}
$$

Here is a plot of the marginal probability for player 1 winning n rounds

```
In[115]:= data = {1.7342 * 10^-2, 8.6708 *10^-2, 0.19509, 0.260112, 0.22761, 0.13656, 5.6902*10^-2,
   1.6258*10^-2, 3.0483*10^-3, 3.3878*10^-4, 1.6935*10^-5};
TableForm[Table[{i - 1, data[[i]]}, {i, 1, Length[data]}], TableHeadings → {None, {"n", "p(n)"}}]
Total[data]
ListPlot[data, Filling → Axis, PlotStyle → {Red, Thick},
 PlotLabel → "Probability of number of rounds wins by firstplayer",
 AxesLabel → {"number of wins", "P(n)"}]
```

Out[116]//TableForm=

n	p(n)
0	0.017342
1	0.086708
2	0.19509
3	0.260112
4	0.22761
5	0.13656
6	0.056902
7	0.016258
8	0.0030483
9	0.00033878
10	0.000016935

Out[117]= 0.999986

Out[118]=

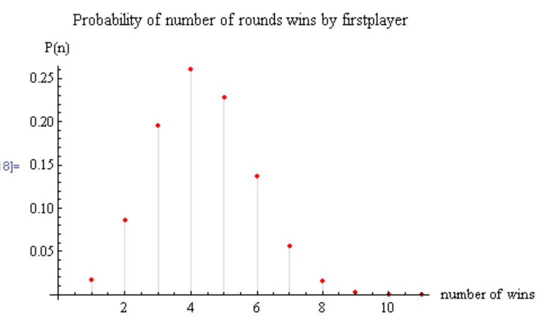


Figure 4.19: marginal probability for player 1

5. Let (X, Y) be jointly distributed random variables with pdf

$$f(x,y) = \frac{1}{8}(x^2 - y^2)e^{-x} \quad 0 \le x \le \infty \quad -x \le y \le x.$$

(a) Find the marginal density of Y. **(b)** Find $P(X+Y \le 1)$. For part (b) leave your solution as integrals, and do not calculate the integrals.

Figure 4.20: Problem 5

(a)

$$\begin{aligned} f_Y(y) &= \int_0^\infty f(x,y)\,dx \\ &= \int_0^\infty \frac{1}{8}\left(x^2 - y^2\right) e^{-x} dx \end{aligned}$$

Integrate by parts, $dv = e^{-x}dx, u = (x^2 - y^2)$, hence $du = 2x$ and $v = -e^{-x}$, so we obtain

$$\begin{aligned} f_Y(y) &= \frac{1}{8}\left\{ \left[\left(x^2 - y^2\right)\left(-e^{-x}\right)\right]_0^\infty - \int_0^\infty (2x)\left(-e^{-x}\right) dx \right\} \\ &= \frac{1}{8}\left\{ -\left[\left(x^2 - y^2\right)\left(e^{-x}\right)\right]_0^\infty + 2\int_0^\infty xe^{-x}dx \right\} \\ &= \frac{1}{8}\left\{ -\left[0 + y^2\right] + 2\int_0^\infty xe^{-x}dx \right\} \end{aligned}$$

Do integration by parts again, $dv = e^{-x}dx$, $u = x$, hence

$$
\begin{aligned}
f_Y(y) &= \frac{1}{8}\left\{-y^2+2\left[\left(x\left(-e^{-x}\right)\right)_0^{\infty}-\int_0^{\infty}-e^{-x}dx\right]\right\}\\
&= \frac{1}{8}\left\{-y^2+2\left[-\left(xe^{-x}\right)_0^{\infty}+\int_0^{\infty}e^{-x}dx\right]\right\}\\
&= \frac{1}{8}\left\{-y^2+2\left[0+\left[-e^{-x}\right]_0^{\infty}\right]\right\}\\
&= \frac{1}{8}\{-y^2+2[-[0-1]]\}\\
&= \frac{1}{8}\{-y^2+2\}
\end{aligned}
$$

Hence

$$\boxed{f_Y(y)=\tfrac{1}{8}(2-y^2)}$$

(b)The hard part is to determine the region to integrate. The following is the needed region which satisfy $P(X+Y\leq 1)$ and $0\leq x\leq\infty$ and $-x\leq y\leq x$

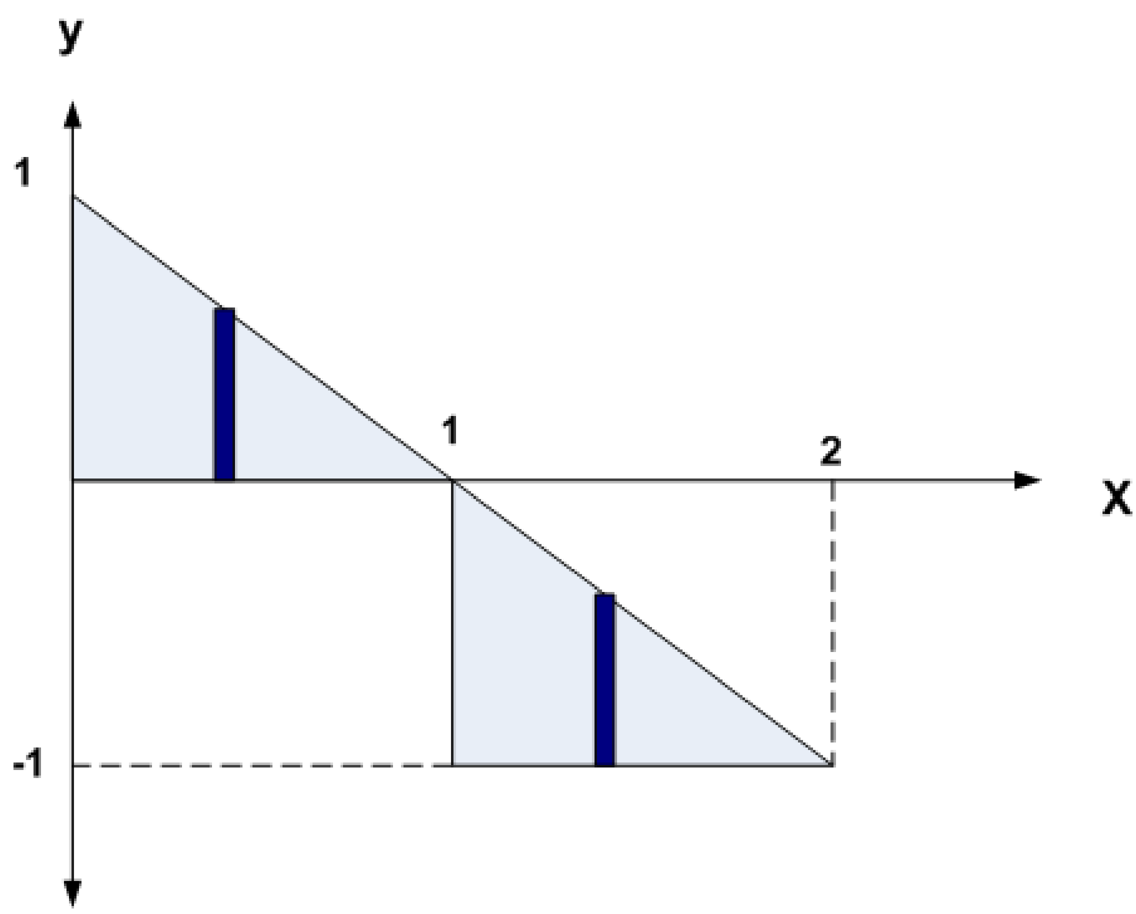


Figure 4.21: region

For the top region,

$$I_1=\int_{x=0}^{x=1}\int_{y=1-x}^{y=1}\frac{1}{8}\left(x^2-y^2\right)e^{-x}dydx$$

and for the bottom region

$$I_2=\int_{x=1}^{x=2}\int_{y=0}^{y=-x}\frac{1}{8}\left(x^2-y^2\right)e^{-x}dydx$$

Hence

$$\boxed{P(X+Y\leq 1)=\int_{x=0}^{x=1}\int_{y=1-x}^{y=1}\tfrac{1}{8}\left(x^2-y^2\right)e^{-x}dydx+\int_{x=1}^{x=2}\int_{y=0}^{y=-x}\tfrac{1}{8}\left(x^2-y^2\right)e^{-x}dydx}$$

4.2.1 Graded

14/20

14/20

QUIZ 2 MATH 502AB Fall 2007

Name (please print) Nasser Abbasi

1. Use the fact that $\Gamma(1/2)=\sqrt{\pi}$ to show that if n is an odd integer, then

$$\Gamma(n/2)=\frac{\sqrt{\pi}(n-1)!}{2^{n-1}\left(\frac{n-1}{2}\right)!}.$$

By definition,

$$\Gamma(z)=\int_0^\infty t^{z-1}e^{-t}dt$$

Hence

$$\Gamma\left(\frac{n}{2}\right)=\int_0^\infty t^{\left(\frac{n}{2}-1\right)}e^{-t}dt$$

When $n=1$, we are told that

$$\Gamma\left(\frac{1}{2}\right)=\int_0^\infty t^{\left(\frac{1}{2}-1\right)}e^{-t}dt=\int_0^\infty t^{\left(-\frac{1}{2}\right)}e^{-t}dt=\sqrt{\pi}$$

For $n=3$, we have

$$\begin{aligned}\Gamma\left(\frac{3}{2}\right)&=\int_0^\infty t^{\left(\frac{3}{2}-1\right)}e^{-t}dt\\&=\int_0^\infty t^{\left(\frac{1}{2}\right)}e^{-t}dt\end{aligned}$$

Now do integration by parts,

$$\begin{aligned}\Gamma\left(\frac{3}{2}\right)&=\int_0^\infty \overbrace{t^{\left(\frac{1}{2}\right)}}^{u}\overbrace{e^{-t}dt}^{dv}\\&=\left[t^{\frac{1}{2}}\left(-e^{-t}\right)\right]_0^\infty-\int_0^\infty\frac{1}{2}t^{\left(-\frac{1}{2}\right)}\left(-e^{-t}\right)dt\\&=-\left[t^{\frac{1}{2}}e^{-t}\right]_0^\infty+\int_0^\infty\frac{1}{2}t^{\left(-\frac{1}{2}\right)}e^{-t}dt\end{aligned}$$

But $\left[t^{\frac{1}{2}}e^{-t}\right]_0^\infty=[0-0]=0$ and the above becomes

$$\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\int_0^\infty t^{\left(-\frac{1}{2}\right)}e^{-t}dt$$

But $\int_0^\infty t^{\left(-\frac{1}{2}\right)}e^{-t}dt=\Gamma\left(\frac{1}{2}\right)$, hence the above becomes

$$\boxed{\Gamma\left(\frac{3}{2}\right)=\frac{1}{2}\Gamma\left(\frac{1}{2}\right)}$$

1

Now do the same for $n = 5$

$$
\begin{aligned}
\Gamma\left(\frac{5}{2}\right) &= \int_0^\infty t^{\left(\frac{5}{2}-1\right)} e^{-t} dt \\
&= \int_0^\infty \overbrace{t^{\left(\frac{3}{2}\right)}}^{u} \overbrace{e^{-t} dt}^{dv} \\
&= -\left[t^{\frac{3}{2}} e^{-t}\right]_0^\infty - \int_0^\infty \frac{3}{2} t^{\left(\frac{1}{2}\right)} \left(-e^{-t}\right) dt \\
&= \frac{3}{2} \int_0^\infty t^{\left(\frac{1}{2}\right)} e^{-t} dt
\end{aligned}
$$

But $\int_0^\infty t^{\left(\frac{1}{2}\right)} e^{-t} dt$ we found from above to be $\Gamma\left(\frac{3}{2}\right)$, hence

$$
\Gamma\left(\frac{5}{2}\right) = \frac{3}{2} \Gamma\left(\frac{3}{2}\right)
$$

But $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2} \Gamma\left(\frac{1}{2}\right)$ from above, hence

$$
\begin{aligned}
\Gamma\left(\frac{5}{2}\right) &= \frac{3}{2}\frac{1}{2} \Gamma\left(\frac{1}{2}\right) \\
&= \frac{3}{2}\frac{1}{2} \sqrt{\pi}
\end{aligned}
$$

Continuing this way, we find that $\Gamma\left(\frac{7}{2}\right) = \frac{5}{2}\frac{3}{2}\frac{1}{2} \Gamma\left(\frac{1}{2}\right)$, and hence in general

$$
\boxed{\Gamma\left(\frac{n}{2}\right) = \frac{(n-2)}{2}\frac{(n-4)}{2}\frac{(n-6)}{2} \cdots \frac{5}{2}\frac{3}{2}\frac{1}{2}\sqrt{\pi}} \tag{1}
$$

Now,

$$
(n-1)! = (n-1)(n-2)(n-3)(n-4)(n-5)(n-6) \cdots 5 \times 4 \times 3 \times 2 \times 1
$$

Hence from above we see that

$$
(n-2)(n-4)(n-6) \cdots 5 \times 3 \times 1 = \frac{(n-1)!}{(n-1)(n-3)(n-5) \cdots 4 \times 2 \times 1}
$$

Therefore (1) can be written as

$$
\begin{aligned}
\Gamma\left(\frac{n}{2}\right) &= \left(\frac{(n-1)!}{(n-1)(n-3)(n-5) \cdots 4 \times 2 \times 1}\right) \overbrace{\left(\frac{1}{2}\frac{1}{2}\frac{1}{2} \cdots \frac{1}{2}\frac{1}{2}\right)}^{\text{There are } \frac{n-1}{2} \text{ such terms}} \sqrt{\pi} \\
&= \frac{(n-1)!}{(n-1)(n-3)(n-5) \cdots 4 \times 2} \frac{1}{2^{\left(\frac{n-1}{2}\right)}} \sqrt{\pi}
\end{aligned} \tag{2}
$$

But

2^{n-1}

Hence (2) becomes

$$\Gamma\left(\frac{n}{2}\right)=\left(\frac{(n-1)!}{\frac{1}{2^{\frac{n-1}{2}}}(n-1)(n-3)(n-5)\cdots 4\times 2}\right)\frac{1}{2^{n-1}}\sqrt{\pi}$$

But there are $\frac{n-1}{2}$ terms in the expression $(n-1)(n-3)(n-5)\cdots 4\times 2$ in the denominator above and we have $\frac{n-1}{2}$ number of $\frac{1}{2}$ sitting there, which we can distribute now below each terms to obtain

$$\Gamma\left(\frac{n}{2}\right)=\left(\frac{(n-1)!}{\frac{(n-1)}{2}\frac{(n-3)}{2}\frac{(n-5)}{2}\cdots\frac{4}{2}\times\frac{2}{2}}\right)\frac{1}{2^{n-1}}\sqrt{\pi}\tag{3}$$

But

$$\begin{aligned}\left(\frac{n-1}{2}\right)! &= \left(\frac{n-1}{2}\right)\left(\frac{n-1}{2}-1\right)\left(\frac{n-1}{2}-2\right)\cdots\times 4\times 3\times 2\times 1\\ &= \left(\frac{n-1}{2}\right)\left(\frac{n-3}{2}\right)\left(\frac{n-5}{2}\right)\cdots\times 4\times 2\end{aligned}$$

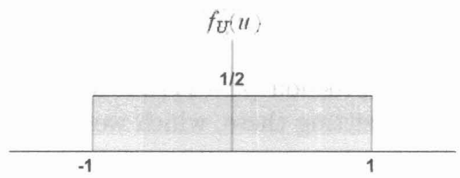
Compare the above to the denominator term in (3) we see it is the same. Hence (3) can be written as

$$\boxed{\Gamma\left(\frac{n}{2}\right)=\frac{(n-1)!}{\left(\frac{n-1}{2}\right)!}\frac{1}{2^{n-1}}\sqrt{\pi}}$$

Which is what we are asked to show.

An induction proof is much cleaner and technically more sound.

2. If $U \sim Uniform[-1,1]$, find the density of $Z = U^2$.



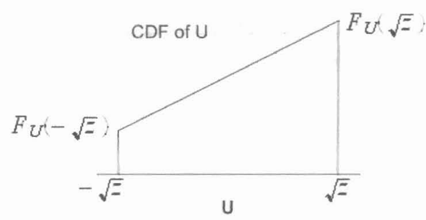
pdf of U

Since U is continuous r.v., we start with the CDF of Z

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P\left(U^2 \leq z\right) \\ &= P\left(-\sqrt{z} \leq U \leq \sqrt{z}\right) \end{aligned} \tag{1}$$

But since $F_U'(u) = f_U(u)$, then we know that $P(a \leq U \leq b) = \int_a^b f_U(x)\,dx \rightarrow F_U(b) - F_U(a)$
Hence RHS of (1) becomes

$$F_Z(z) = F_U\left(\sqrt{z}\right) - F_U\left(-\sqrt{z}\right)$$



Therefore, taking derivatives with respect to z we obtain

$$\begin{aligned} f_Z(z) &= f_U\left(\sqrt{z}\right)\frac{d}{dz}\sqrt{z} - f_U\left(-\sqrt{z}\right)\frac{d}{dz}\left(-\sqrt{z}\right) \\ &= \frac{1}{2}z^{\frac{-1}{2}} f_U\left(\sqrt{z}\right) + f_U\left(-\sqrt{z}\right)\frac{1}{2}z^{\frac{-1}{2}} \\ &= \frac{1}{2}z^{\frac{-1}{2}}\left(f_U\left(\sqrt{z}\right) + f_U\left(-\sqrt{z}\right)\right) \end{aligned}$$

Since U is uniform, hence $f_U(a) = f_U(-a)$, hence the above becomes

$$\begin{aligned} f_Z(z) &= \frac{1}{2}z^{\frac{-1}{2}}\left(2f_U\left(\sqrt{z}\right)\right) \\ &= \frac{1}{\sqrt{z}} f_U\left(\sqrt{z}\right) \end{aligned}$$

Now I need to determine the limits of $f_Z(z)$ and the shape. f_U is defined for real arguments from -1 to $+1$. i.e. f_U is real valued function of real arguments. Hence if z was negative then

$\sqrt{z}$ will be complex, and so this will not be allowed. Hence we have to restrict $z \geq 0$. But now we observe that $z = 0$ is not possible, since we will have $\frac{1}{0}$ term, so this means z is strictly larger than zero. So

$$f_Z(z) = \frac{1}{\sqrt{z}} f_U\left(\sqrt{z}\right) \qquad z > 0$$

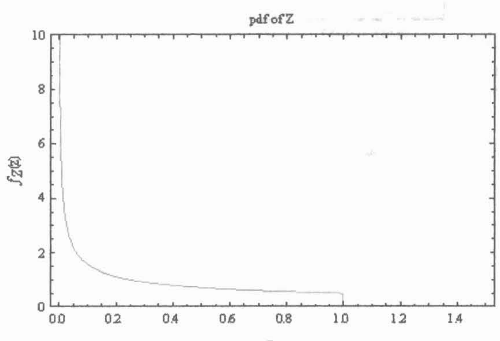
But we know that $f_U(x) = \frac{1}{2}$ for up to $x = 1$, hence this means when $\sqrt{z} > 1$ then $f_U\left(\sqrt{z}\right) = 0$, when means when $z > 1$ then $f_U\left(\sqrt{z}\right) = 0$

Hence we now write

$$\boxed{f_Z(z) = \begin{cases} \frac{1}{\sqrt{z}}\frac{1}{2} & 0 < z \leq 1 \\ 0 & z > 1 \\ undefined & z \leq 0 \end{cases}}$$

Here is a plot

```
In[75]:= f[z_] := 1/2 1/Sqrt[z] /; z > 0 && z ≤ 1
         f[z_] := 0 /; z > 1
         Plot[f[z], {z, 0.00001, 1.5}, PlotRange → {All, {0, 10}}, FrameLabel → {"z", "fz(z)", "pdf of Z"},
          Frame → True]
```

Out[77]=


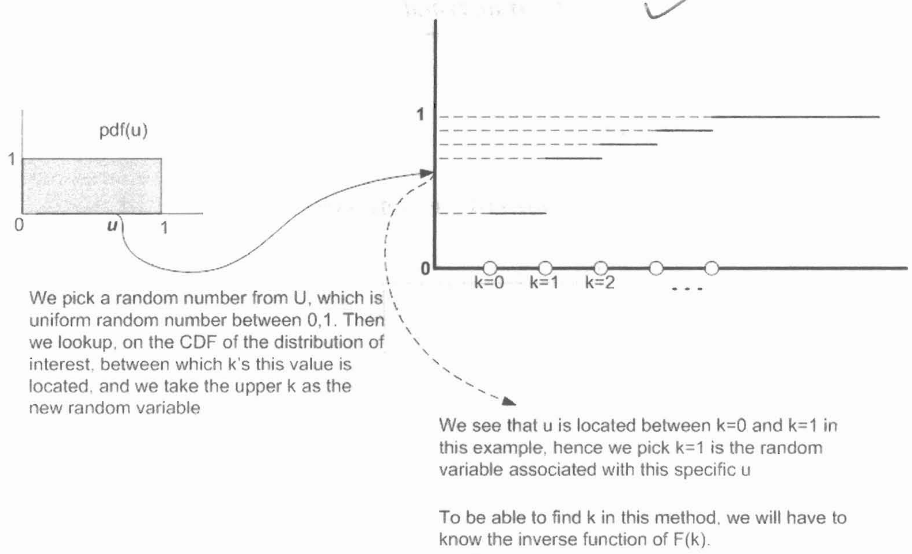
3. The following five numbers were randomly generated from the uniform random variable on (0,1):

0.0153 0.7468 0.4451 0.9318

Using these numbers generate five random numbers from the geometric random variable with parameter $p = 1/3$. Very briefly explain how you obtain your solution.

I explain the idea behind obtaining a discrete random number from a continues random number by the following diagram below. We assume that the discrete random number belongs to some distribution. In this example, we are told what the distribution is. We know that the CDF for geometric random variable is given by

$$F_K(k) = 1 - (1-p)^k$$



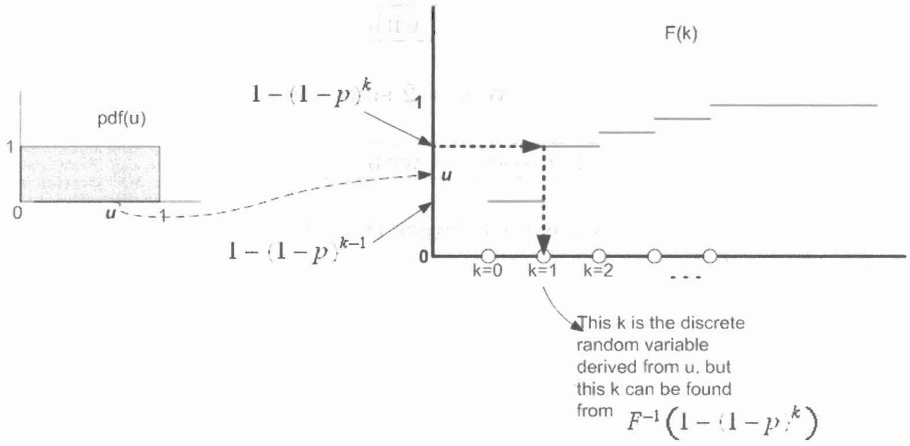
We see from the above diagram, that once we are given u we need to find k which satisfy the following identity

$$F_K(k-1) < u \leq F_K(k)$$

Or in other words

$$1-(1-p)^{k-1} < u \leq 1-(1-p)^k$$

The specific discrete value k which will satisfy the above, is the random variable we want, which belong to the geometeric distribution.



Now when $u = 0.0153$, and since $p = \frac{1}{3}$, we have

$$1-\left(1-\frac{1}{3}\right)^{k-1} < 0.0153 \leq 1-\left(1-\frac{1}{3}\right)^{k}$$

for $k = 1$, we have

$$0 < 0.0153 \leq 0.333\,33 \qquad YES$$

$\boxed{\text{Hence } k = 1 \text{ is the random variable associated with } u = 0.0153}$

Now let us do $u = 0.7468$
for $k = 1$, we have

$$0 < 0.7468 \leq 0.333\,33 \qquad NO$$

try $k = 2$

$$1-\left(1-\frac{1}{3}\right)^{1} < 0.7468 \leq 1-\left(1-\frac{1}{3}\right)^{2}$$
$$0.333\,33 < 0.7468 \leq 0.555\,56 \qquad NO$$

try $k = 3$

$$1-\left(1-\frac{1}{3}\right)^{2} < 0.7468 \leq 1-\left(1-\frac{1}{3}\right)^{3}$$
$$0.555\,56 < 0.7468 \leq 0.703\,7 \qquad NO$$

try $k = 4$

$$1-\left(1-\frac{1}{3}\right)^{3} < 0.7468 \leq 1-\left(1-\frac{1}{3}\right)^{4}$$
$$0.703\,7 < 0.7468 \leq 0.802\,47 \qquad YES$$

$\boxed{\text{Hence } k=4 \text{ is the random variable associated with } u=0.7468}$

Now let us do $u = 0.4451$

We see from the above, that this will have $k = 2$ since for $k = 2$ the intervals is $0.333\,33 < u \leq 0.555\,56$

$\boxed{\text{Hence } k=2 \text{ is the random variable associated with } u=0.4451}$

Now let us do $u = 0.9318$

From above, we see that this will have a k larger than 4, so we do not need to try from the start, we can start trying from $k = 5$

try $k = 5$

$$1-\left(1-\frac{1}{3}\right)^4 < 0.9318 \leq 1-\left(1-\frac{1}{3}\right)^5$$
$$0.802\,47 < 0.9318 \leq 0.868\,31 \qquad NO$$

try $k = 6$

$$1-\left(1-\frac{1}{3}\right)^5 < 0.9318 \leq 1-\left(1-\frac{1}{3}\right)^6$$
$$0.868\,31 < 0.9318 \leq 0.912\,21 \qquad NO$$

try $k = 7$

$$1-\left(1-\frac{1}{3}\right)^6 < 0.9318 \leq 1-\left(1-\frac{1}{3}\right)^7$$
$$0.912\,21 < 0.9318 \leq 0.941\,47 \qquad YES$$

$\boxed{\text{Hence } k=7 \text{ is the random variable associated with } u=0.9318}$

Hence result is

u	k
0.0153	1
0.4451	2
0.7468	4
0.9318	7

of course one would write a program to do this.

8

4. Three players play 10 independent rounds of a game, and each player has probability 1/3 of winning each round. **(a)** Find the joint distribution of the numbers of games won by each of the three players. **(b)** Identify the distribution of the number of games won by player one.

(a)
Let $P(x_i = n_i)$ means probability of player i winning n_i rounds.
We have 3 players, and a total of 10 rounds. Let the players be called x_1, x_2, x_3. Let the number of games WON by x_1 be n_1, and number of games won by x_2 be n_2, and number of games won by x_3 be n_3.

Since we have 10 rounds, then we must have 10 wins as well. (some one must win). Hence we have 10 wins and 3 ways to split it, where each 'bucket' is of different size. So this is a multi set selection. called multinomial in the book using proposition B in chapter 1, we see that the total number of ways the games can be won is $\binom{10}{n_1 n_2 n_3}$

But we need to find the probability of each one such combination. So we need to multiply the above by the probability each player wins the number of the games they happened to win, which is $P(x_i = n_i) = p^{n_i}$, but $p = \frac{1}{3}$ for each player to win a round. Hence we write

$$\begin{aligned} P(x_1 = n_1, x_2 = n_2, x_3 = n_3) &= \binom{10}{n_1 n_2 n_3} \left(\frac{1}{3}\right)^{n_1} \left(\frac{1}{3}\right)^{n_2} \left(\frac{1}{3}\right)^{n_3} \\ &= \frac{10!}{n_1!\, n_2!\, n_3!} \left(\frac{1}{3}\right)^{n_1+n_2+n_3} \\ &= \frac{10!}{n_1!\, n_2!\, n_3!} \left(\frac{1}{3}\right)^{10} \end{aligned}$$

-1 $n_1+n_2+n_3=?$

So the above is the joint probability that p_1 wins n_1 rounds and p_2 wins n_2 rounds and p_3 wins n_3 rounds.

(b)We need to find $P(x_1 = n_1)$, i.e. the probability of first player winning n_1 rounds.

$$P(x_1 = n_1) = \sum_{\substack{n_2=0,1,\cdots 10 \\ n_3=0,1,\cdots 10}} P(x_1 = n_1, x_2 = n_2, x_3 = n_3)$$

To simplify, let me write $P(n_1, n_2, n_3)$ instead, where the position of the n implies the player. So $p(0,1,9)$ means player one wins zero rounds and player 2 wins 1 round and player 3 wins 9 rounds.

So the above becomes

$$P(x_1 = n_1) = \sum_{\substack{n_2=0,1,\cdots 10 \\ n_3=0,1,\cdots 10}} P(n_1, n_2, n_3)$$

But since $n_1 = 10 - (n_2 + n_3)$ we see that we only need to count those terms in the above sum when this is true. i.e. we do not need to count a term such as $p(1,0,0)$ since this is zero probability of happening. Now we write

$$P(x_1 = n_1) = P(n_1, 10 - n_1, 0) + P(n_1, 9 - n_1, 1) + P(n_1, 8 - n_1, 2) + \cdots P(n_1, 0, 10 - n_1)$$

For example,
$P(x_1=0) = P(0,10,0)+P(0,9,1)+P(0,8,2)+P(0,7,3)+P(0,6,4)+P(0,5,5)+P(0,4,6)+P(0,3,7)+P(0,2,8)+P(0,1,9)+P(0,0,10)$
But $P(0,10,0)=P(0,0,10)$ and $P(0,9,1)=P(0,1,9)$, etc.. so the above can be written as

$$P(x_1=0) = 2P(0,10,0)+2P(0,9,1)+2P(0,8,2)+2P(0,7,3)+2P(0,6,4)+P(0,5,5)$$

$$\begin{aligned}
&= 2\frac{10!}{0!\,10!\,0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{0!\,9!\,1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{0!\,8!\,2!}\left(\frac{1}{3}\right)^{10} + \\
&2\frac{10!}{0!7!\,3!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{0!\,6!\,4!}\left(\frac{1}{3}\right)^{10} + \frac{10!}{0!\,5!\,5!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{1.7342\times10^{-2}}
\end{aligned}$$

and

$$\begin{aligned}
P(x_1=1) &= P(1,9,0)+P(1,8,1)+P(1,7,2)+P(1,6,3)+P(1,5,4)+P(1,4,5) \\
&\quad +P(1,3,6)+P(1,2,7)+P(1,1,8)+P(1,0,9) \\
&= 2P(1,9,0)+2P(1,8,1)+2P(1,7,2)+2P(1,6,3)+2P(1,5,4) \\
&= 2\frac{10!}{1!\,9!\,0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!\,8!\,1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!\,7!\,2!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!6!\,3!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{1!\,5!\,4!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{8.6708\times10^{-2}}
\end{aligned}$$

and

$$\begin{aligned}
P(x_1=2) &= P(2,8,0)+P(2,7,1)+P(2,6,2)+P(2,5,3)+P(2,4,4)+P(2,3,5)+ \\
&\quad P(2,2,6)+P(2,1,7)+P(2,0,8) \\
&= 2P(2,8,0)+2P(2,7,1)+2P(2,6,2)+2P(2,5,3)+P(2,4,4) \\
&= 2\frac{10!}{2!\,8!\,0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{2!\,7!\,1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{2!\,6!\,2!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{2!5!\,3!}\left(\frac{1}{3}\right)^{10} + \frac{10!}{2!\,4!\,4!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{0.19509}
\end{aligned}$$

and

$$P(x_1 = 3) = P(3,7,0) + P(3,6,1) + P(3,5,2) + P(3,4,3) + P(3,3,4) + P(3,2,5) + P(3,1,6) + P(3,0,7)$$

$$= 2P(3,7,0) + 2P(3,6,1) + 2P(3,5,2) + 2P(3,4,3)$$

$$= 2\frac{10!}{3!\ 7!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{3!\ 6!\ 1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{3!\ 5!\ 2!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{3!4!\ 3!}\left(\frac{1}{3}\right)^{10}$$
$$= \boxed{0.260\,12}$$

and

$$P(x_1 = 4) = P(4,6,0) + P(4,5,1) + P(4,4,2) + P(4,3,3) + P(4,2,4) + P(4,1,5) + P(4,0,6)$$
$$= 2P(4,6,0) + 2P(4,5,1) + 2P(4,4,2) + P(4,3,3)$$

$$= 2\frac{10!}{4!\ 6!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{4!\ 5!\ 1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{4!\ 4!\ 2!}\left(\frac{1}{3}\right)^{10} + \frac{10!}{4!3!\ 3!}\left(\frac{1}{3}\right)^{10}$$
$$= \boxed{0.227\,61}$$

and

$$P(x_1 = 5) = P(5,5,0) + P(5,4,1) + P(5,3,2) + P(5,2,3) + P(5,1,4) + P(5,0,5)$$
$$= 2P(5,5,0) + 2P(5,4,1) + 2P(5,3,2)$$
$$= 2\frac{10!}{5!\ 5!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{5!4!\ 1!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{5!\ 3!\ 2!}\left(\frac{1}{3}\right)^{10}$$
$$= \boxed{0.136\,56}$$

and

$$P(x_1 = 6) = P(6,4,0) + P(6,3,1) + P(6,2,2) + P(6,1,3) + P(6,0,4)$$
$$= 2P(6,4,0) + 2P(6,3,1) + P(6,2,2)$$
$$= 2\frac{10!}{6!\ 4!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{6!3!\ 1!}\left(\frac{1}{3}\right)^{10} + \frac{10!}{6!\ 2!\ 2!}\left(\frac{1}{3}\right)^{10}$$
$$= \boxed{5.690\,2 \times 10^{-2}}$$

and

$$
\begin{aligned}
P(x_1 = 7) &= P(7,3,0) + P(7,2,1) + P(7,1,2) + P(7,0,3) \\
&= 2P(7,3,0) + 2P(7,2,1) \\
&= 2\frac{10!}{7!\ 3!\ 0!}\left(\frac{1}{3}\right)^{10} + 2\frac{10!}{7!2!\ 1!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{1.6258 \times 10^{-2}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1 = 8) &= P(8,2,0) + P(8,1,1) + P(8,0,2) \\
&= 2P(8,2,0) + P(8,1,1) \\
&= 2\frac{10!}{8!\ 2!\ 0!}\left(\frac{1}{3}\right)^{10} + \frac{10!}{8!1!\ 1!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{3.0483 \times 10^{-3}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1 = 9) &= P(9,1,0) + P(9,0,1) \\
&= 2P(9,1,0) \\
&= 2\frac{10!}{9!\ 1!\ 0!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{3.387 \times 10^{-4}}
\end{aligned}
$$

and

$$
\begin{aligned}
P(x_1 = 10) &= P(10,0,0) \\
&= \frac{10!}{10!\ 0!\ 0!}\left(\frac{1}{3}\right)^{10} \\
&= \boxed{1.6935 \times 10^{-5}}
\end{aligned}
$$

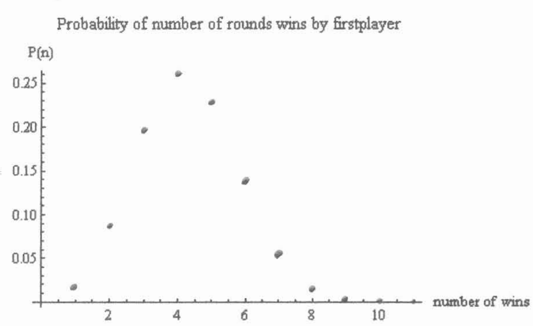
Here is a plot of the marginal probability for player 1 winning n rounds

```
In[115]:= data = {1.7342*10^-2, 8.6708*10^-2, 0.19509, 0.260112, 0.22761, 0.13656, 5.6902*10^-2,
    1.6258*10^-2, 3.0483*10^-3, 3.3878*10^-4, 1.6935*10^-5};
TableForm[Table[{i - 1, data[[i]]}, {i, 1, Length[data]}], TableHeadings -> {None, {"n", "p(n)"}}]
Total[data]
ListPlot[data, Filling -> Axis, PlotStyle -> {Red, Thick},
 PlotLabel -> "Probability of number of rounds wins by firstplayer",
 AxesLabel -> {"number of wins", "P(n)"}]
```

n	p (n)
0	0.017342
1	0.086708
2	0.19509
3	0.260112
4	0.22761
5	0.13656
6	0.056902
7	0.016258
8	0.0030483
9	0.00033878
10	0.000016935

Out[117]= 0.999986

Out[118]=



PS. we are asked to "identify" this distribution.
I would then have to say that this distribution is "Binomial", since in Binomial we ask: how many wins in "n" trials when the probability of win is p in each.
So here it is like that for asking how many wins for player ONE.
(PS. this part is the Hardest part! "with what parameters"
i.e identify the type of

5. Let (X,Y) be jointly distributed random variables with pdf

$$f(x,y)=\frac{1}{8}(x^2-y^2)e^{-x}\quad 0\le x\le\infty\quad -x\le y\le x.$$

(a) Find the marginal density of Y. **(b)** Find $P(X+Y\le 1)$. For part (b) leave your solution as integrals, and do not calculate the integrals.

(a)

$$\begin{aligned} f_Y(y) &= \int_0^\infty f(x,y)\,dx \\ &= \int_0^\infty \frac{1}{8}\left(x^2-y^2\right)e^{-x}dx \end{aligned}$$

Integrate by parts, $dv=e^{-x}dx, u=(x^2-y^2)$, hence $du=2x$ and $v=-e^{-x}$, so we obtain

$$\begin{aligned} f_Y(y) &= \frac{1}{8}\left\{\left[\left(x^2-y^2\right)\left(-e^{-x}\right)\right]_0^\infty-\int_0^\infty(2x)\left(-e^{-x}\right)dx\right\} \\ &= \frac{1}{8}\left\{-\left[\left(x^2-y^2\right)\left(e^{-x}\right)\right]_0^\infty+2\int_0^\infty xe^{-x}dx\right\} \\ &= \frac{1}{8}\left\{-\left[0+y^2\right]+2\int_0^\infty xe^{-x}dx\right\} \end{aligned}$$

Do integration by parts again, $dv=e^{-x}dx$, $u=x$, hence

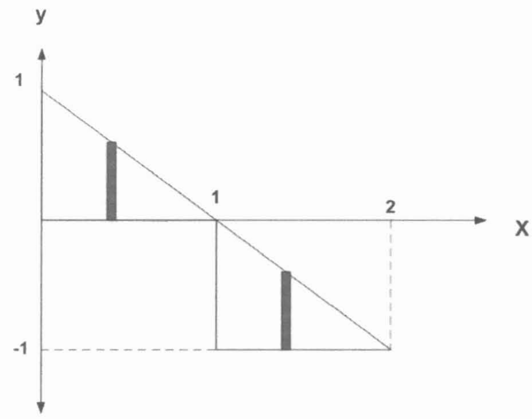
-4

$$\begin{aligned} f_Y(y) &= \frac{1}{8}\left\{-y^2+2\left[\left(x\left(-e^{-x}\right)\right)_0^\infty-\int_0^\infty -e^{-x}dx\right]\right\} \\ &= \frac{1}{8}\left\{-y^2+2\left[-\left(xe^{-x}\right)_0^\infty+\int_0^\infty e^{-x}dx\right]\right\} \\ &= \frac{1}{8}\left\{-y^2+2\left[0+\left[-e^{-x}\right]_0^\infty\right]\right\} \\ &= \frac{1}{8}\left\{-y^2+2\left[-\left[0-1\right]\right]\right\} \\ &= \frac{1}{8}\left\{-y^2+2\right\} \end{aligned}$$

Hence

$$\boxed{f_Y(y)=\tfrac{1}{8}\left(2-y^2\right)}$$

(b)The hard part is to determine the region to integrate. The following is the needed region which satisfy $P(X+Y\le 1)$ and $0\le x\le\infty$ and $-x\le y\le x$



For the top region,

$$I_1 = \int_{x=0}^{x=1} \int_{y=1-x}^{y=1} \frac{1}{8} \left(x^2 - y^2\right) e^{-x} dy dx$$

and for the bottom region

$$I_2 = \int_{x=1}^{x=2} \int_{y=0}^{y=-x} \frac{1}{8} \left(x^2 - y^2\right) e^{-x} dy dx$$

Hence

$$\boxed{P\left(X+Y \leq 1\right) = \int_{x=0}^{x=1} \int_{y=1-x}^{y=1} \frac{1}{8} \left(x^2 - y^2\right) e^{-x} dy dx + \int_{x=1}^{x=2} \int_{y=0}^{y=-x} \frac{1}{8} \left(x^2 - y^2\right) e^{-x} dy dx}$$

4.3 Quiz 3

Local contents

4.3.1 short version

Quiz 3 Math 502AB Fall 2007

Name (please print) NASSER ABBASI

1. Suppose that two components have independent exponentially distributed lifetimes T_1 and T_2, with parameters α and β, respectively. Find (a) $P(T_1 > T_2)$, (b) identify the distribution of $W = 2T_2$, and (c) use the results in parts (a) and (b) to obtain $P(T_1 > 2T_2)$.

Figure 4.22: Problem 1

(a)Problem review:

T_1 is a random variable and T_2 is a random variable, where $T_1 = \alpha e^{-\alpha t_1}$ and $T_2 = \beta e^{-\beta t_2}$

α and β can be thought of as the failure rate for each respective component. T_i is the lifetime of component i. Hence $P(T_1 = t_1)$ means to ask for the probability of the first component to have a lifetime of t_1 given that the failure rate of this kind of components is α.

solution:

Now we know that

$$P(T_1 > T_2) = \int\int f_{T_1,T_2}(t_1,t_2)\,dt_2 dt_1$$

Looking at the following diagram to help determine the region to integrate:

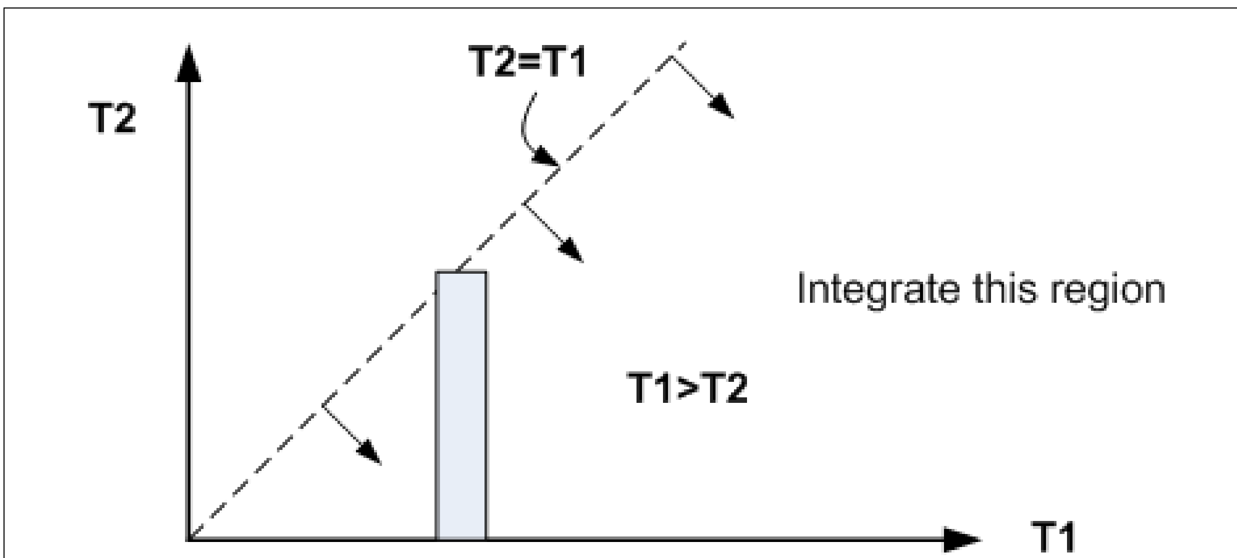


Figure 4.23: region to integrate

Hence

$$P(T_1 > T_2) = \int_{t_1=0}^{t_1=\infty}\int_{t_2=0}^{t_2=t_1} f_{T_1,T_2}(t_1,t_2)\,dt_2\,dt_1$$

But since $T_1 \perp T_2$, then the joint density is the product of the marginal densities.

Hence

$$\begin{aligned} f_{T_1,T_2}(t_1,t_2) &= f_{T_1}(t_1)\, f_{T_2}(t_2) \\ &= \alpha e^{-\alpha t_1} \beta e^{-\beta t_2} \end{aligned}$$

Therefore

$$\begin{aligned} P(T_1 > T_2) &= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \beta e^{-\beta t_2}\, dt_2\, dt_1 \\ &= \beta\alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta t_2}\, dt_2 \right) dt_1 \\ &= \beta\alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{1}{\beta} \left[e^{-\beta t_2} \right]_{t_2=0}^{t_2=t_1} \right) dt_1 \\ &= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta t_1} - 1 \right] dt_1 \\ &= -\alpha \int_0^\infty e^{t_1(-\alpha-\beta)} - e^{-\alpha t_1}\, dt_1 \\ &= -\alpha \left(\left[\frac{1}{(-\alpha-\beta)} e^{t_1(-\alpha-\beta)} \right]_0^\infty - \frac{1}{-\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \end{aligned}$$

We take $\alpha, \beta \geq 0$ since we expect the lifetime to go to zero eventually. Also this is a requirement for the integrals to not diverge.

Hence the above becomes

$$\begin{aligned} P(T_1 > T_2) &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{t_1(-\alpha-\beta)} \right]_0^\infty + \frac{1}{\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \\ &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{-\infty} - 1 \right] + \frac{1}{\alpha} \left[e^{-\infty} - 1 \right] \right) \\ &= -\alpha \left(\frac{1}{(-\alpha-\beta)} [0-1] + \frac{1}{\alpha} [0-1] \right) \\ &= -\alpha \left(\frac{1}{(\alpha+\beta)} - \frac{1}{\alpha} \right) \\ &= -\alpha \left(\frac{\alpha - (\alpha+\beta)}{\alpha(\alpha+\beta)} \right) \\ &= -\left(\frac{\alpha-\alpha-\beta}{(\alpha+\beta)} \right) \end{aligned}$$

Hence

$$\boxed{P(T_1 > T_2) = \frac{\beta}{(\alpha+\beta)}}$$

(b)

$$\begin{aligned} F_W(w) &= P(W \leq w) \\ &= P(2T_2 \leq w) \\ &= P\left(T_2 \leq \frac{w}{2} \right) \\ &= F_{T_2}\left(\frac{w}{2} \right) \end{aligned}$$

Hence

$$f_W(w) = f_{T_2}\left(\frac{w}{2}\right) \times \frac{d}{dw}\left(\frac{w}{2}\right)$$

Hence

$$\boxed{f_W(w) = \frac{1}{2} f_{T_2}\left(\frac{w}{2}\right)}$$

(c)Need to find $P(T_1 > 2T_2)$ which is the same as $P(T_1 > W)$, hence this is the same as part(a) but replace T_2 by W as show in the following diagram

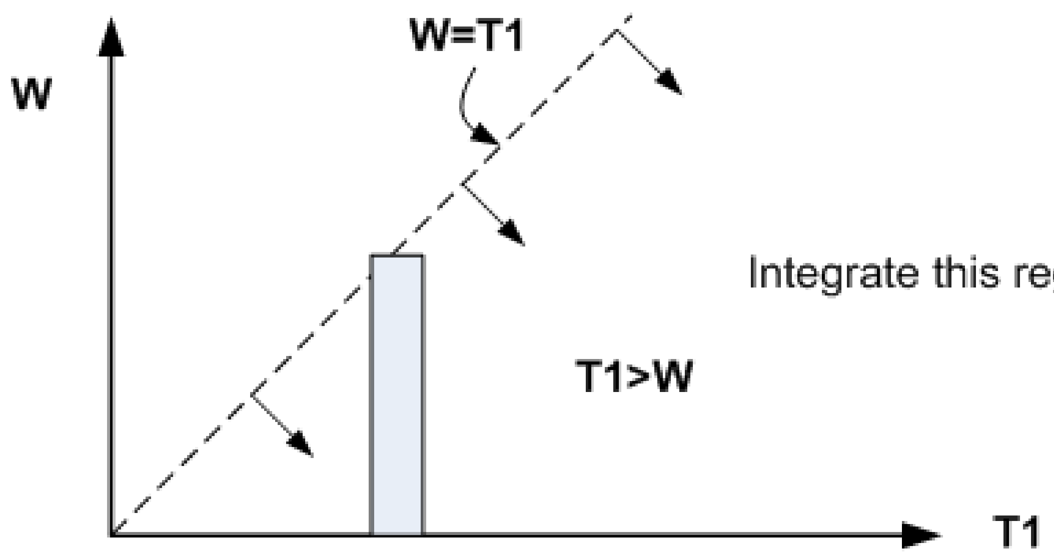


Figure 4.24: diagram

Hence

$$\begin{aligned}
P(T_1 > W) &= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1)\, f_W(w)\ dw\ dt_1 \\
&= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1) \left[\frac{1}{2} f_{T_2}\left(\frac{w}{2}\right)\right]\ dw\ dt_1 \\
&= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \left[\frac{1}{2}\beta e^{-\beta\left(\frac{w}{2}\right)}\right]\ dw\ dt_1 \\
&= \frac{1}{2}\beta\alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta\left(\frac{w}{2}\right)}\ dw\right)\ dt_1 \\
&= \frac{1}{2}\beta\alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{2}{\beta}\left[e^{-\beta\left(\frac{w}{2}\right)}\right]_{w=0}^{w=t_1}\right)\ dt_1 \\
&= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta\left(\frac{t_1}{2}\right)} - 1\right]\ dt_1 \\
&= -\alpha \int_0^\infty e^{t_1\left(-\alpha - \frac{\beta}{2}\right)} - e^{-\alpha t_1}\ dt_1 \\
&= -\alpha \int_0^\infty e^{t_1\left(\frac{-2\alpha-\beta}{2}\right)} - e^{-\alpha t_1}\ dt_1 \\
&= -\alpha \left(\left[\frac{2}{(-2\alpha-\beta)} e^{t_1\left(\frac{-2\alpha-\beta}{2}\right)}\right]_0^\infty - \frac{1}{-\alpha}\left[e^{-\alpha t_1}\right]_0^\infty\right) \\
&= -\alpha \left(\frac{2}{(-2\alpha-\beta)}[0-1] + \frac{1}{\alpha}[0-1]\right) \\
&= -\alpha \left(\frac{2}{(2\alpha+\beta)} - \frac{1}{\alpha}\right)
\end{aligned}$$

Hence

$$
\begin{aligned}
P(T_1 > W) &= -\left(\frac{2\alpha - (2\alpha+\beta)}{(2\alpha+\beta)}\right) \\
&= -\left(\frac{2\alpha - 2\alpha - \beta}{(2\alpha+\beta)}\right)
\end{aligned}
$$

Then

$$\boxed{P(T_1 > W) = \frac{\beta}{(2\alpha+\beta)}}$$

2. Consider a Poisson process on the real line, and denote by $N(t_1, t_2)$ the number of events in the interval (t_1, t_2). If $t_0 < t_1 < t_2$, find the conditional distribution of $N(t_0, t_1)$ given that $N(t_0, t_2) = n$, and identify the distribution.

Figure 4.25: Problem 2

Problem review: Poisson probability density is a discrete probability function (We normally call it the probability mass function pmf). This means the random variable is a discrete random variable.

The random variable X in this case is the number of success in n trials where the probability of success in each one trial is p and the trials are independent from each others. The difference between Poisson and Binomial is that in Poisson we are looking at the problem as n becomes very large and p becomes very small in such a way that the product np goes to a fixed value which is called λ, the Poisson parameter. And then we write $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ where $k = 0, 1, 2, \cdots$ The following diagram illustrates this problem, showing the three r.v. we need to analyze and the time line.

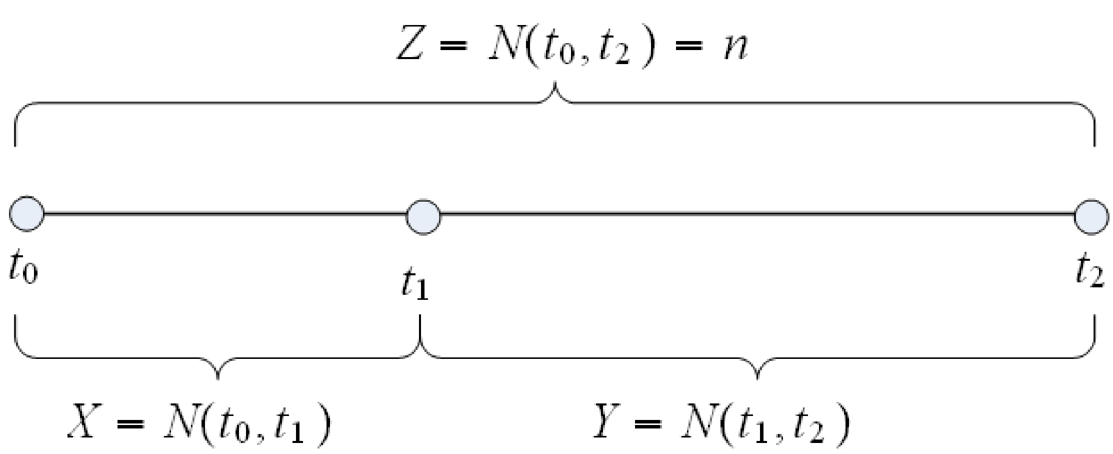


Figure 4.26: diagram illustrates this problem

But what is "trials" in this problem? If we divide the time line itself into very small time intervals δt then the number of time intervals is the number of trials, and we assume that at most one event will occur in this time interval (since it is too small). The probability p of event occurring in this δt is the same in the interval $[t_0, t_1]$ and in the interval $[t_1, t_2]$. Now let us find λ for X and Y and Z based on this. Since $\lambda = np$ where n is the number of trials, then for X we have $\lambda_x = n_x p = \frac{(t_1 - t_0)}{\delta t} p$ where we divided the time interval by the time width δt to obtain the number of time slots for X. We do the same for Y and obtain that $\boxed{\lambda_y = \frac{(t_2 - t_1)}{\delta t} p}$

Similarly, $\lambda_Z = \frac{(t_2-t_0)}{\delta t}p = \frac{(t_2-t_1)+(t_1-t_0)}{\delta t}p = \frac{(t_2-t_1)}{\delta t}p + \frac{(t_1-t_0)}{\delta t}P$, hence $\boxed{\lambda_z = \lambda_x + \lambda_y}$

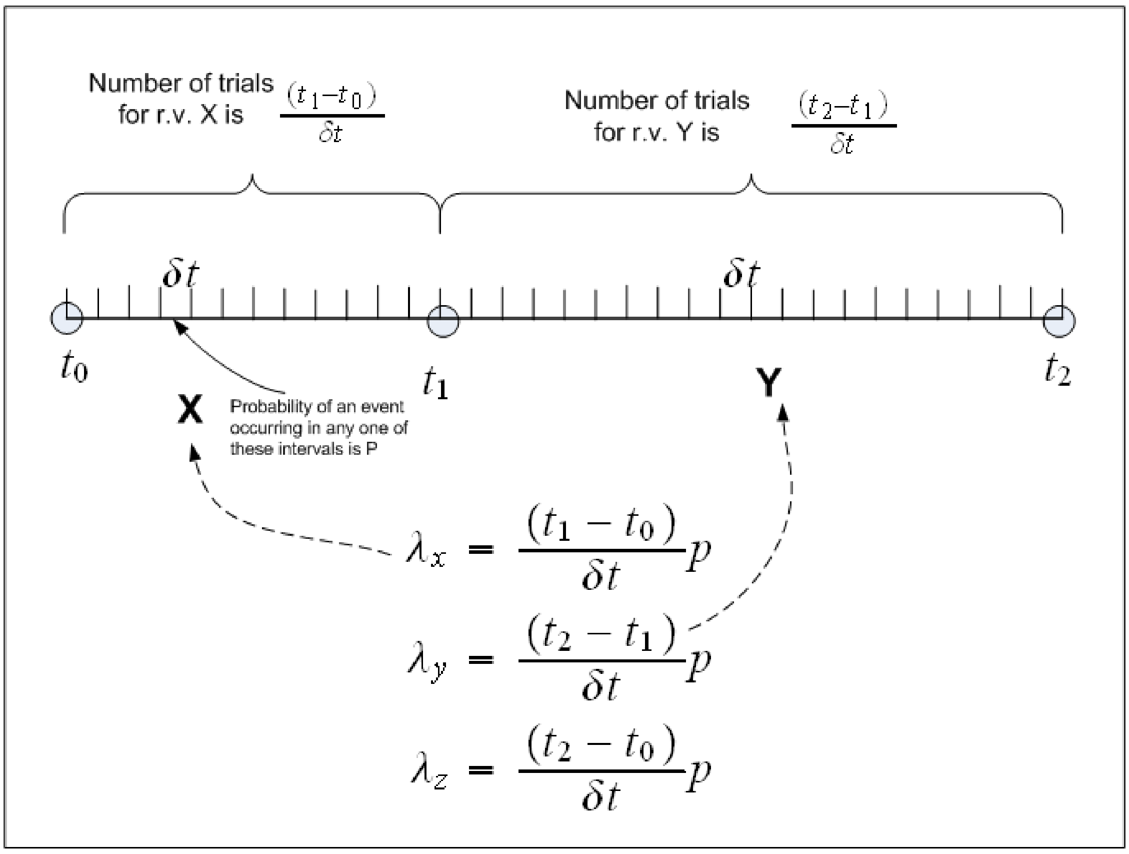


Figure 4.27: delta

Let us refer to the random variable $N(t_1,t_2)$ as Y and the r.v. $N(t_0,t_1)$ as X and the r.v. $N(t_0,t_2)$ as Z

The problem is then asking to find $P(X=x|Z=n)$ and to identify $pmf(X|Z)$

To help in the solution, we first draw a diagram to make it more clear.

We take λ to the same for the 3 random variables X, Y, Z.

$$P(X=x|Z=n) = \frac{P(X=x, Z=n)}{P(Z=n)}$$

But $Z=n$ is the same as $X+Y=n$ hence

$$\begin{aligned} P(X=x|Z=n) &= \frac{P(X=x,(X+Y)=n)}{P(Z=n)} \\ &= \frac{P(X=x, Y=n-x)}{P(Z=n)} \end{aligned}$$

Now r.v. $X \perp Y$, since the number of events in $[t_0,t_1]$ is independent from the number of events that could occur in $[t_1,t_2]$.

Given this, we can now write the joint probability of X, Y as the product of the marginal probabilities. Hence the numerator in the above can be rewritten and we obtain

$$P(X=x|Z=n) = \frac{P(X=x)\,P(Y=n-x)}{P(Z=n)} \tag{1}$$

Now since each of the above is a Poisson process, then

$$
\begin{aligned}
P(X=x) &= \frac{(\lambda_x)^x}{x!} e^{-\lambda_x} \\
P(Y=n-x) &= \frac{(\lambda_y)^{n-x}}{(n-x)!} e^{-\lambda_y} \\
P(Z=n) &= \frac{(\lambda_z)^n}{n!} e^{-\lambda_z}
\end{aligned}
$$

Hence (1) becomes

$$
P(X=x|Z=n) = \left(\frac{(\lambda_x)^x}{x!} e^{-\lambda_x}\right)\left(\frac{(\lambda_y)^{n-x}}{(n-x)!} e^{-\lambda_y}\right)\frac{1}{\frac{(\lambda_z)^n}{n!} e^{-\lambda_z}} \tag{2}
$$

Hence

$$
P(X=x|Z=n) = \frac{n!}{x!\,(n-x)!}\left((\lambda_x)^x e^{-\lambda_x}\right)\left((\lambda_y)^{n-x} e^{-\lambda_y}\right)\frac{e^{\lambda_z}}{(\lambda_z)^n}
$$

But we found that $\lambda_z = \lambda_x + \lambda_y$, hence the exponential term above vanish and we get

$$
\begin{aligned}
P(X=x|Z=n) &= \frac{n!}{x!\,(n-x)!}\frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_z)^n} \\
&= \binom{n}{x}\frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_z)^n} \\
&= \binom{n}{x}\frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^n} \\
&= \binom{n}{x}\frac{(\lambda_x)^x (\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^x (\lambda_x+\lambda_y)^{n-x}} \\
&= \binom{n}{x}\frac{(\lambda_x)^x}{(\lambda_x+\lambda_y)^x}\frac{(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^{n-x}} \\
&= \binom{n}{x}\left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x\left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{n-x}
\end{aligned}
$$

Let $k = \frac{\lambda_x}{\lambda_x+\lambda_y}$, then $1-k = 1-\frac{\lambda_x}{\lambda_x+\lambda_y} = \frac{\lambda_x+\lambda_y-\lambda_x}{\lambda_x+\lambda_y} = \frac{\lambda_y}{\lambda_x+\lambda_y}$ hence the last line above can be written as

$$
\begin{aligned}
P(X=x|Z=n) &= \binom{n}{x}\left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x\left(1-\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^{n-x} \\
&= \binom{n}{x}(k)^x (1-k)^{n-x}
\end{aligned}
$$

But this is a Binomial with parameters n, k, hence

$$
\boxed{P(X=x|Z=n) = Binomial\left(n, \frac{\lambda_x}{\lambda_x+\lambda_y}\right)}
$$

3. Suppose that the probability Θ of getting heads for a coin is unknown, and let the prior opinion about Θ be represented by the uniform distribution on [0,1]. You spin the coin repeatedly and record the number of times N until a heads comes up. (a) Find the posterior density of Θ given N. (b) Use Matlab or any other software to plot the posterior for cases where $N = 1$, $N = 2$, and $N = 6$. Using your plots, explain what you infer about the probability of heads in each circumstance.

Figure 4.28: Problem 3

part (a)

Let θ, the probability of getting heads, be the specific value that the random number Θ can take.

Let $g(\theta)$ be the probability density of Θ, which we are told to be $U[0,1]$, and let $pmf_X(x)$ be the probability mass function of the random variable X where X is the number of times until a head first comes up. $\boxed{X \text{ is then a geometric random variable with parameter } \theta}$, hence

$$pmf_X(N) = P(X = N) = (1-\theta)^{N-1}\,\theta \qquad N = 1, 2, 3, \cdots$$

The posterior density of Θ given N is then

$$\boxed{h(\Theta = \theta | X = N) = \frac{pmf_X(N|\Theta=\theta)g(\theta)}{\int_0^1 pmf_X(N|\Theta=\theta)g(\theta)d\theta}}$$

But

$$pmf_X(N|\Theta = \theta) = (1-\theta)^{N-1}\,\theta$$

and $g(\theta) = 1$ since $\Theta = U[0,1]$

Hence

$$h(\Theta = \theta | X = N) = \frac{(1-\theta)^{N-1}\,\theta}{\int_0^1 (1-\theta)^{N-1}\,\theta\; d\theta} \tag{1}$$

But Θ is a random continuous variable from $[0,1]$, so how to evaluate the above? I can evaluate the above for different values of Θ on the real line from $[0,1]$, and the more values I take between $0, 1$ the more accurate $h(\Theta = \theta|X = N)$ will become.

Part(b)

First let me evaluate eq (1) for $N = 1, N = 2, N = 6$

For $N = 1$

$$h(\Theta = \theta | X = 1) = \frac{\theta}{\int_0^1 \theta\; d\theta} = \frac{\theta}{\left[\frac{\theta^2}{2}\right]_0^1} = \boxed{2\theta}$$

For $N = 2$

$$\begin{aligned} h(\Theta = \theta | X = 2) &= \frac{(1-\theta)\,\theta}{\int_0^1 (1-\theta)\,\theta\; d\theta} = \frac{(1-\theta)\,\theta}{\int_0^1 (\theta-\theta^2)\; d\theta} = \frac{(1-\theta)\,\theta}{\left[\frac{\theta^2}{2}\right]_0^1 - \left[\frac{\theta^3}{3}\right]_0^1} \\ &= \frac{(1-\theta)\,\theta}{\frac{1}{2}-\frac{1}{3}} = \boxed{6(1-\theta)\,\theta} \end{aligned}$$

For $N = 6$

$$h(\Theta = \theta | X = 6) = \frac{(1-\theta)^{6-1}\theta}{\int_0^1 (1-\theta)^{6-1}\theta\, d\theta} = \frac{(1-\theta)^5\theta}{\int_0^1 (1-\theta)^5\theta\, d\theta}$$

We can use integration by parts for the denominator, where $u = \theta, dv = (1-\theta)^5$, when we do this we obtain

$$h(\Theta = \theta | X = 6) = \boxed{42(1-\theta)^5\theta}$$

Now we plot the above 3 cases on the same plot:

```
In[93]:= f[n_, x_] := Which[n == 1, 2 x, n == 2, 6 (1 - x) x, n == 6, 42 (1 - x)^5 x]
Plot[{f[1, x], f[2, x], f[6, x]}, {x, 0, 1},
 Frame → True, PlotStyle → {Red, Black, Blue},
 FrameLabel → {Style["θ", 14], Style["h(θ|X=N)", 16],
   Style["Posterior probability distribution for θ for different N", 16]},
 ImageSize → 600]
```

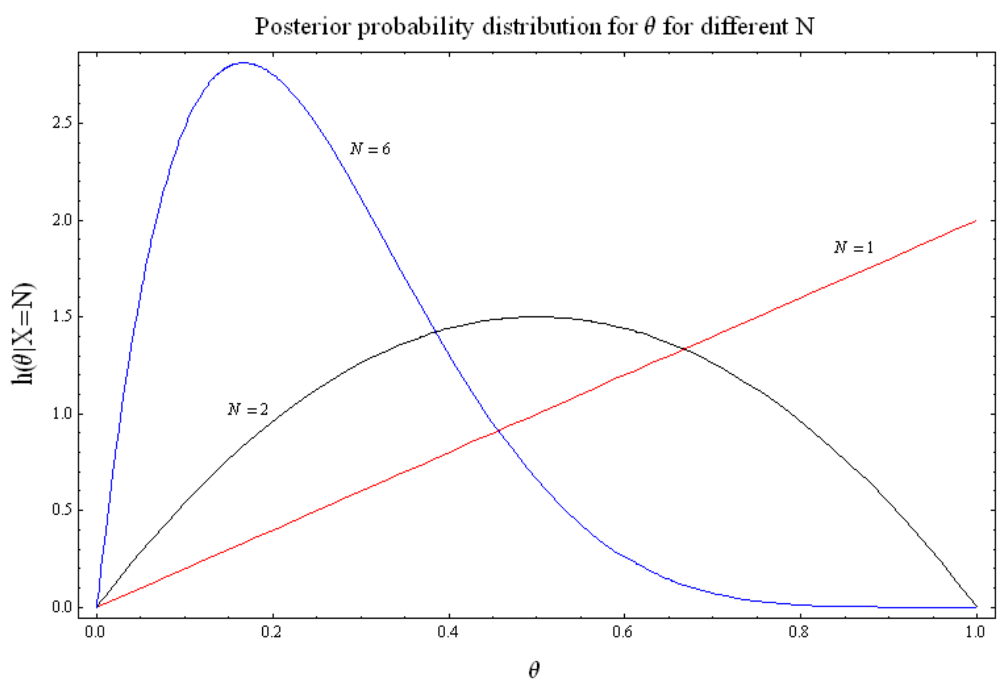


Figure 4.29: posterior

What the above plot is saying is the following:

If it takes 'longer' to see a head comes up ($N = 6$), then the coin is taken as biased towards a tail, and the probability of getting a head becomes smaller, this is why we see that the most likely probability in this case to be around 0.15 (looking at the N=6 curve). We say that based on the observation of $N = 6$, then the coin has a higher probability of having its probability of getting a head to be about 0.15 than any other value. (The area around $\theta = 1.5$ is larger than any other area for the same $\delta\theta$)

Now, when $N = 2$, i.e. we flipped the coin 2 times, and got a head on the second time, then we see from the $N = 2$ curve that the coin has a most likelihood of having a probability of getting a head to be 0.5.

This is what we would expect, since in an unbiased coin, the probability of getting a head is $\frac{1}{2}$, and hence with a fair coin, we expect to see a head half of the times it is flipped, and since we flipped 2 times, and saw a head the second time, this posterior probability has its most likely value to be around .5 as well.

When $N = 1$, this says that we got a head in the first time we flipped the coin. We see that the posterior probability of getting a head now has it maximum around 1. This means the posterior probability is saying this coin is biased towards a head.

The above is a method to estimate the probability distribution of the probability itself of getting a head based on the observed events and based on the prior known probability of getting a head. Hence the events observed allow us to estimate the probability of getting a head. Hence the posterior probability is conditioned on each event as in this problem.

4.3.2 Graded

18/20

Quiz 3 Math 502AB Fall 2007

Name (please print) NASSER ABBASI

1. Suppose that two components have independent exponentially distributed lifetimes T_1 and T_2, with parameters α and β, respectively. Find (a) $P(T_1 > T_2)$, (b) identify the distribution of $W = 2T_2$, and (c) use the results in parts (a) and (b) to obtain $P(T_1 > 2T_2)$.

(a)Problem review:

T_1 is a random variable and T_2 is a random variable, where $T_1\tilde{}\alpha e^{-\alpha t_1}$ and $T_2\tilde{}\beta e^{-\beta t_2}$.

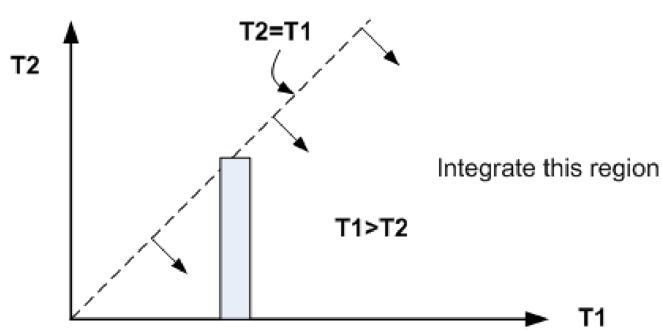
α and β can be thought of as the failure rate for each respective component. T_i is the lifetime of component i. Hence $P(T_1 = t_1)$ means to ask for the probability of the first component to have a lifetime of t_1 given that the failure rate of this kind of components is α.

solution:

Now we know that

$$P(T_1 > T_2) = \int\int f_{T_1,T_2}(t_1,t_2)\, dt_2 dt_1$$

Looking at the following diagram to help determine the region to integrate:



Hence

$$P(T_1 > T_2) = \int_{t_1=0}^{t_1=\infty} \int_{t_2=0}^{t_2=t_1} f_{T_1,T_2}(t_1,t_2)\, dt_2\, dt_1$$

But since $T_1 \perp T_2$, then the joint density is the product of the marginal densities.

Hence

$$\begin{aligned} f_{T_1,T_2}(t_1,t_2) &= f_{T_1}(t_1)\, f_{T_2}(t_2) \\ &= \alpha e^{-\alpha t_1} \beta e^{-\beta t_2} \end{aligned}$$

Therefore

$$\begin{aligned} P(T_1 > T_2) &= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \beta e^{-\beta t_2}\, dt_2\, dt_1 \\ &= \beta\alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta t_2}\, dt_2 \right) dt_1 \\ &= \beta\alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{1}{\beta} \left[e^{-\beta t_2} \right]_{t_2=0}^{t_2=t_1} \right) dt_1 \\ &= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta t_1} - 1 \right] dt_1 \\ &= -\alpha \int_0^\infty e^{t_1(-\alpha-\beta)} - e^{-\alpha t_1}\, dt_1 \\ &= -\alpha \left(\left[\frac{1}{(-\alpha-\beta)} e^{t_1(-\alpha-\beta)} \right]_0^\infty - \frac{1}{-\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \end{aligned}$$

We take $\alpha, \beta \geq 0$ since we expect the lifetime to go to zero eventually. Also this is a requirement for the integrals to not diverge.

Hence the above becomes

$$\begin{aligned} P(T_1 > T_2) &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{t_1(-\alpha-\beta)} \right]_0^\infty + \frac{1}{\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \\ &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{-\infty} - 1 \right] + \frac{1}{\alpha} \left[e^{-\infty} - 1 \right] \right) \\ &= -\alpha \left(\frac{1}{(-\alpha-\beta)} [0-1] + \frac{1}{\alpha} [0-1] \right) \\ &= -\alpha \left(\frac{1}{(\alpha+\beta)} - \frac{1}{\alpha} \right) \\ &= -\alpha \left(\frac{\alpha - (\alpha+\beta)}{\alpha(\alpha+\beta)} \right) \\ &= -\left(\frac{\alpha - \alpha - \beta}{(\alpha+\beta)} \right) \end{aligned}$$

Hence

$$\boxed{P(T_1 > T_2) = \frac{\beta}{(\alpha+\beta)}}$$

(b)

$$\begin{aligned} F_W(w) &= P(W \leq w) \\ &= P(2T_2 \leq w) \\ &= P\left(T_2 \leq \frac{w}{2}\right) \\ &= F_{T_2}\left(\frac{w}{2}\right) \end{aligned}$$

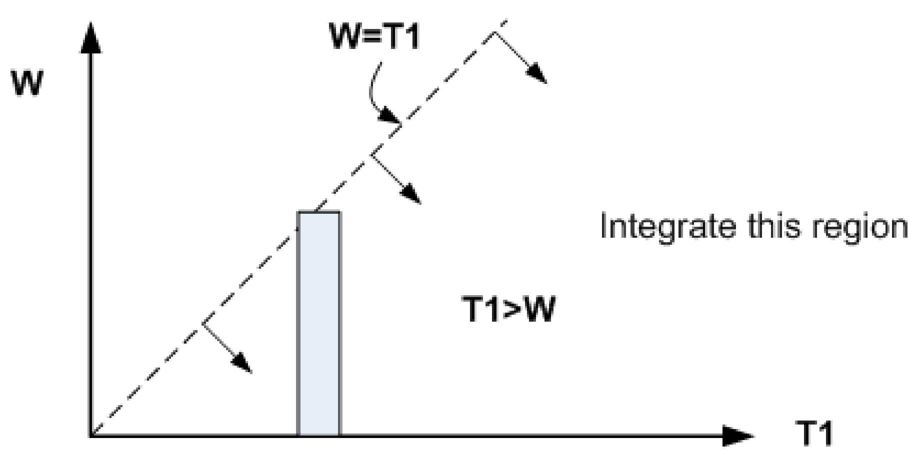
Hence

$$f_W(w) = f_{T_2}\left(\frac{w}{2}\right) \times \frac{d}{dw}\left(\frac{w}{2}\right)$$

Hence

$$\boxed{f_W(w) = \frac{1}{2} f_{T_2}\left(\frac{w}{2}\right)}$$

(c)Need to find $P(T_1 > 2T_2)$ which is the same as $P(T_1 > W)$, hence this is the same as part(a) but replace T_2 by W as show in the following diagram



Hence

$$\begin{aligned}
P(T_1 > W) &= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1)\, f_W(w)\ dw\ dt_1 \\
&= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1) \left[\frac{1}{2} f_{T_2}\left(\frac{w}{2}\right)\right]\ dw\ dt_1 \\
&= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \left[\frac{1}{2}\beta e^{-\beta\left(\frac{w}{2}\right)}\right]\ dw\ dt_1 \\
&= \frac{1}{2}\beta\alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta\left(\frac{w}{2}\right)}\ dw\right)\ dt_1 \\
&= \frac{1}{2}\beta\alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{2}{\beta}\left[e^{-\beta\left(\frac{w}{2}\right)}\right]_{w=0}^{w=t_1}\right)\ dt_1 \\
&= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta\left(\frac{t_1}{2}\right)} - 1\right]\ dt_1 \\
&= -\alpha \int_0^\infty e^{t_1\left(-\alpha - \frac{\beta}{2}\right)} - e^{-\alpha t_1}\ dt_1 \\
&= -\alpha \int_0^\infty e^{t_1\left(\frac{-2\alpha-\beta}{2}\right)} - e^{-\alpha t_1}\ dt_1 \\
&= -\alpha\left(\left[\frac{2}{(-2\alpha-\beta)} e^{t_1\left(\frac{-2\alpha-\beta}{2}\right)}\right]_0^\infty - \frac{1}{-\alpha}\left[e^{-\alpha t_1}\right]_0^\infty\right) \\
&= -\alpha\left(\frac{2}{(-2\alpha-\beta)}[0-1] + \frac{1}{\alpha}[0-1]\right) \\
&= -\alpha\left(\frac{2}{(2\alpha+\beta)} - \frac{1}{\alpha}\right)
\end{aligned}$$

Hence

$$\begin{aligned}
P(T_1 > W) &= -\left(\frac{2\alpha - (2\alpha+\beta)}{(2\alpha+\beta)}\right) \\
&= -\left(\frac{2\alpha - 2\alpha - \beta}{(2\alpha+\beta)}\right)
\end{aligned}$$

Then

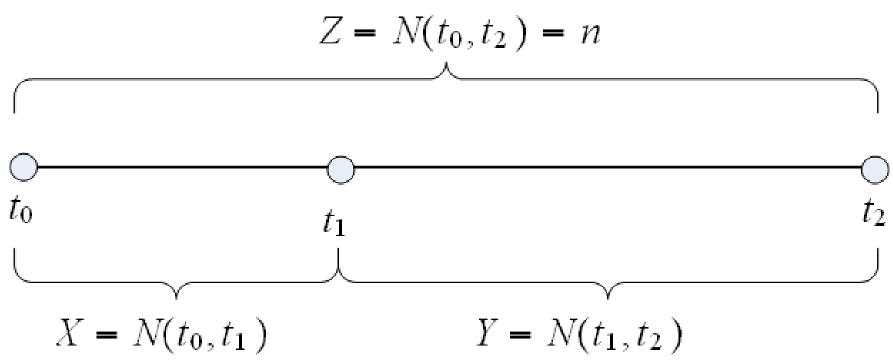
$$\boxed{P(T_1 > W) = \frac{\beta}{(2\alpha+\beta)}}$$

2. Consider a Poisson process on the real line, and denote by $N(t_1, t_2)$ the number of events in the interval (t_1, t_2). If $t_0 < t_1 < t_2$, find the conditional distribution of $N(t_0, t_1)$ given that $N(t_0, t_2) = n$, and identify the distribution.

4

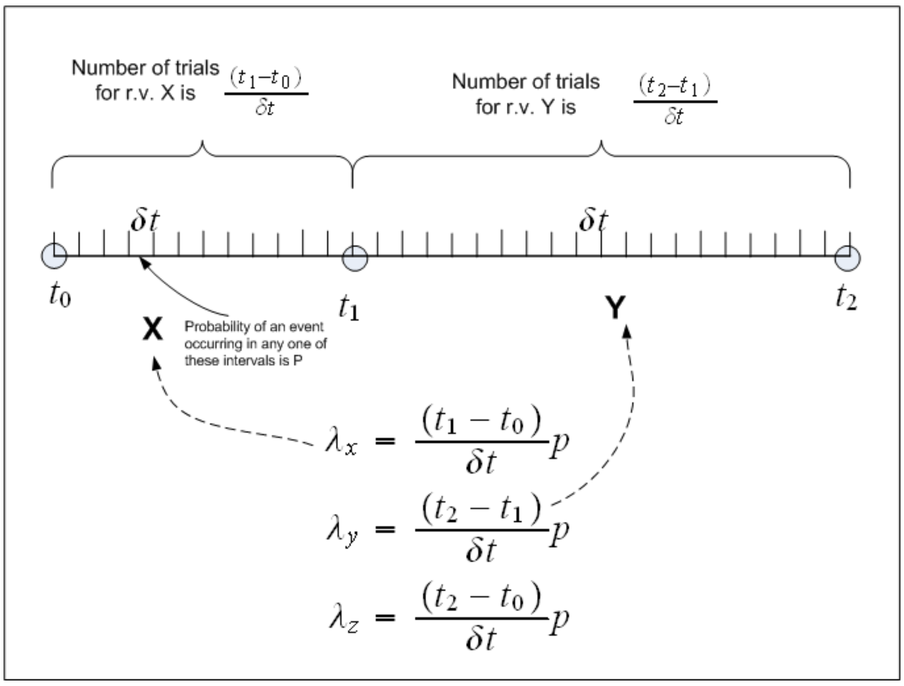
Problem review: Poisson probability density is a discrete probability function (We normally call it the probability mass function pmf). This means the random variable is a discrete random variable.

The random variable X in this case is the number of success in n trials where the probability of success in each one trial is p and the trials are independent from each others. The difference between Poisson and Binomial is that in Poisson we are looking at the problem as n becomes very large and p becomes very small in such a way that the product np goes to a fixed value which is called λ, the Poisson parameter. And then we write $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ where $k=0,1,2,\cdots$ The following diagram illustrates this problem, showing the three r.v. we need to analyze and the time line.



But what is "trials" in this problem? If we divide the time line itself into very small time intervals δt then the number of time intervals is the number of trials, and we assume that at most one event will occur in this time interval (since it is too small). The probability p of event occurring in this δt is the same in the interval $[t_0,t_1]$ and in the interval $[t_1,t_2]$. Now let us find λ for X and Y and Z based on this. Since $\lambda=np$ where n is the number of trials, then for X we have $\lambda_x=n_x p=\frac{(t_1-t_0)}{\delta t}p$ where we divided the time interval by the time width δt to obtain the number of time slots for X. We do the same for Y and obtain that $\boxed{\lambda_y=\frac{(t_2-t_1)}{\delta t}p}$

Similarly, $\lambda_Z=\frac{(t_2-t_0)}{\delta t}p=\frac{(t_2-t_1)+(t_1-t_0)}{\delta t}p=\frac{(t_2-t_1)}{\delta t}p+\frac{(t_1-t_0)}{\delta t}P$, hence $\boxed{\lambda_z=\lambda_x+\lambda_y}$



Let us refer to the random variable $N(t_1, t_2)$ as Y and the r.v. $N(t_0, t_1)$ as X and the r.v. $N(t_0, t_2)$ as Z

The problem is then asking to find $P(X = x|Z = n)$ and to identify $pmf(X|Z)$

To help in the solution, we first draw a diagram to make it more clear.

We take λ to the same for the 3 random variables X, Y, Z.

$$P(X = x|Z = n) = \frac{P(X = x, Z = n)}{P(Z = n)}$$

But $Z = n$ is the same as $X + Y = n$ hence

$$\begin{aligned} P(X = x|Z = n) &= \frac{P(X = x, (X + Y) = n)}{P(Z = n)} \\ &= \frac{P(X = x, Y = n - x)}{P(Z = n)} \end{aligned}$$

Now r.v. $X \perp Y$, since the number of events in $[t_0, t_1]$ is independent from the number of events that could occur in $[t_1, t_2]$.

Given this, we can now write the joint probability of X, Y as the product of the marginal probabilities. Hence the numerator in the above can be rewritten and we obtain

6

$$P(X=x|Z=n)=\frac{P(X=x)\,P(Y=n-x)}{P\,(Z=n)} \tag{1}$$

Now since each of the above is a Poisson process, then

$$\begin{aligned}
P(X=x) &= \frac{(\lambda_x)^x}{x!}e^{-\lambda_x}\\
P(Y=n-x) &= \frac{(\lambda_y)^{n-x}}{(n-x)!}e^{-\lambda_y}\\
P(Z=n) &= \frac{(\lambda_z)^n}{n!}e^{-\lambda_z}
\end{aligned}$$

Hence (1) becomes

$$P(X=x|Z=n)=\left(\frac{(\lambda_x)^x}{x!}e^{-\lambda_x}\right)\left(\frac{(\lambda_y)^{n-x}}{(n-x)!}e^{-\lambda_y}\right)\frac{1}{\frac{(\lambda_z)^n}{n!}e^{-\lambda_z}} \tag{2}$$

Hence

$$P(X=x|Z=n)=\frac{n!}{x!\,(n-x)!}\left((\lambda_x)^x\,e^{-\lambda_x}\right)\left((\lambda_y)^{n-x}\,e^{-\lambda_y}\right)\frac{e^{\lambda_z}}{(\lambda_z)^n}$$

But we found that $\lambda_z=\lambda_x+\lambda_y$, hence the exponential term above vanish and we get

$$\begin{aligned}
P(X=x|Z=n) &= \frac{n!}{x!\,(n-x)!}\frac{(\lambda_x)^x\,(\lambda_y)^{n-x}}{(\lambda_z)^n}\\
&= \binom{n}{x}\frac{(\lambda_x)^x\,(\lambda_y)^{n-x}}{(\lambda_z)^n}\\
&= \binom{n}{x}\frac{(\lambda_x)^x\,(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^n}\\
&= \binom{n}{x}\frac{(\lambda_x)^x\,(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^x\,(\lambda_x+\lambda_y)^{n-x}}\\
&= \binom{n}{x}\frac{(\lambda_x)^x}{(\lambda_x+\lambda_y)^x}\frac{(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^{n-x}}\\
&= \binom{n}{x}\left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x\left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{n-x}
\end{aligned}$$

Let $k=\frac{\lambda_x}{\lambda_x+\lambda_y}$, then $1-k=1-\frac{\lambda_x}{\lambda_x+\lambda_y}=\frac{\lambda_x+\lambda_y-\lambda_x}{\lambda_x+\lambda_y}=\frac{\lambda_y}{\lambda_x+\lambda_y}$ hence the last line above can be written as

$$\begin{aligned}
P(X=x|Z=n) &= \binom{n}{x}\left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x\left(1-\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^{n-x} \\
&= \binom{n}{x}(k)^x(1-k)^{n-x}
\end{aligned}$$

But this is a Binomial with parameters n, k, hence

$$\boxed{P(X=x|Z=n)\,\tilde{}\,Binomial\left(n,\frac{\lambda_x}{\lambda_x+\lambda_y}\right)}$$

3. Suppose that the probability Θ of getting heads for a coin is unknown, and let the prior opinion about Θ be represented by the uniform distribution on [0,1]. You spin the coin repeatedly and record the number of times N until a heads comes up. (a) Find the posterior density of Θ given N. (b) Use Matlab or any other software to plot the posterior for cases where $N=1$, $N=2$, and $N=6$. Using your plots, explain what you infer about the probability of heads in each circumstance.

part (a)

Let θ, the probability of getting heads, be the specific value that the random number Θ can take.

Let $g(\theta)$ be the probability density of Θ, which we are told to be $U[0,1]$, and let $pmf_X(x)$ be the probability mass function of the random variable X where X is the number of times until a head first comes up. $\boxed{X \text{ is then a geometric random variable with parameter } \theta}$, hence

$$pmf_X(N)=P(X=N)=(1-\theta)^{N-1}\theta \qquad N=1,2,3,\cdots$$

The posterior density of Θ given N is then

$$\boxed{h(\Theta=\theta|X=N)=\frac{pmf_X(N|\Theta=\theta)g(\theta)}{\int_0^1 pmf_X(N|\Theta=\theta)g(\theta)d\theta}}$$

But

$$pmf_X(N|\Theta=\theta)=(1-\theta)^{N-1}\theta$$

and $g(\theta)=1$ since $\Theta\,\tilde{}\,U[0,1]$

8

Hence

$$h(\Theta = \theta | X = N) = \frac{(1-\theta)^{N-1}\,\theta}{\int_0^1 (1-\theta)^{N-1}\,\theta\; d\theta} \tag{1}$$

But Θ is a random continuous variable from $[0,1]$, so how to evaluate the above? I can evaluate the above for different values of Θ on the real line from $[0,1]$, and the more values I take between $0,1$ the more accurate $h(\Theta = \theta | X = N)$ will become.

Part(b)

First let me evaluate eq (1) for $N=1, N=2, N=6$

For $N=1$

$$h(\Theta = \theta | X = 1) = \frac{\theta}{\int_0^1 \theta\; d\theta} = \frac{\theta}{\left[\frac{\theta^2}{2}\right]_0^1} = \boxed{2\theta}$$

For $N=2$

$$\begin{aligned} h(\Theta = \theta | X = 2) &= \frac{(1-\theta)\,\theta}{\int_0^1 (1-\theta)\,\theta\; d\theta} = \frac{(1-\theta)\,\theta}{\int_0^1 (\theta - \theta^2)\; d\theta} = \frac{(1-\theta)\,\theta}{\left[\frac{\theta^2}{2}\right]_0^1 - \left[\frac{\theta^3}{3}\right]_0^1} \\ &= \frac{(1-\theta)\,\theta}{\frac{1}{2} - \frac{1}{3}} = \boxed{6(1-\theta)\,\theta} \end{aligned}$$

For $N=6$

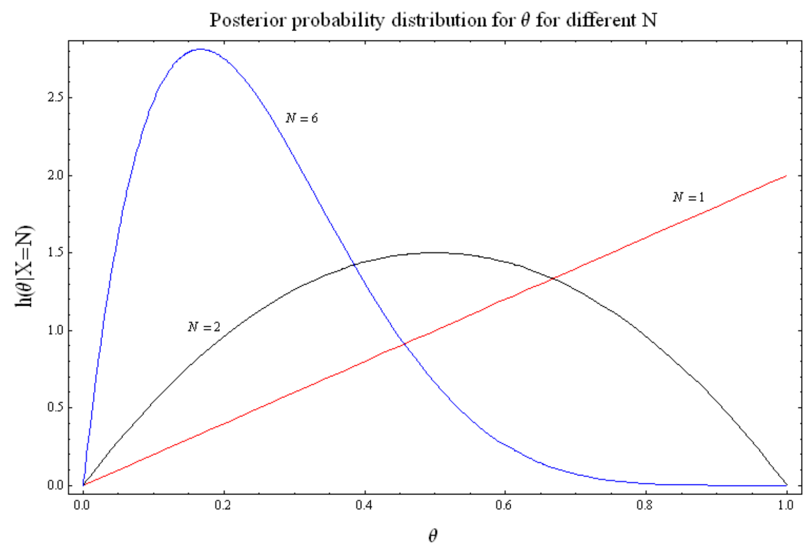
$$h(\Theta = \theta | X = 6) = \frac{(1-\theta)^{6-1}\,\theta}{\int_0^1 (1-\theta)^{6-1}\,\theta\; d\theta} = \frac{(1-\theta)^5\,\theta}{\int_0^1 (1-\theta)^5\,\theta\; d\theta}$$

We can use integration by parts for the denominator, where $u=\theta, dv = (1-\theta)^5$, when we do this we obtain

$$h(\Theta = \theta | X = 6) = \boxed{42(1-\theta)^5\,\theta}$$

Now we plot the above 3 cases on the same plot:

```
In[93]:= f[n_, x_] := Which[n == 1, 2 x, n == 2, 6 (1 - x) x, n == 6, 42 (1 - x)^5 x]
 Plot[{f[1, x], f[2, x], f[6, x]}, {x, 0, 1},
  Frame -> True, PlotStyle -> {Red, Black, Blue},
  FrameLabel -> {Style["θ", 14], Style["h(θ|X=N)", 16],
    Style["Posterior probability distribution for θ for different N", 16]},
  ImageSize -> 600]
```



What the above plot is saying is the following:

If it takes 'longer' to see a head comes up ($N = 6$), then the coin is taken as biased towards a tail, and the probability of getting a head becomes smaller, this is why we see that the most likely probability in this case to be around 0.15 (looking at the N=6 curve). We say that based on the observation of $N = 6$, then the coin has a higher probability of having its probability of getting a head to be about 0.15 than any other value. (The area around $\theta = 1.5$ is larger than any other area for the same $\delta\theta$)

Now, when $N = 2$, i.e. we flipped the coin 2 times, and got a head on the second time, then we see from the $N = 2$ curve that the coin has a most likelihood of having a probability of getting a head to be 0.5.

This is what we would expect, since in an unbiased coin, the probability of getting a head is $\frac{1}{2}$, and hence with a fair coin, we expect to see a head half of the times it is flipped, and since we flipped 2 times, and saw a head the second time, this posterior probability has its most likely value to be around .5 as well.

When $N = 1$, this says that we got a head in the first time we flipped the coin. We see that the posterior probability of getting a head now has it maximum around 1. This means the posterior probability is saying this coin is biased towards a head.

The above is a method to estimate the probability distribution of the probability itself of getting a head based on the observed events and based on the prior known probability of getting a head. Hence the events observed allow us to estimate the probability of getting a head. Hence the posterior probability is conditioned on each event as in this problem.

4.3.3 long version

QUIZ 3 MATH 502AB Fall 2007

Name (please print) NASSER ABBASI

1. Suppose that two components have independent exponentially distributed lifetimes T_1 and T_2, with parameters α and β, respectively. Find (a) $P(T_1 > T_2)$, (b) identify the distribution of $W = 2T_2$, and (c) use the results in parts (a) and (b) to obtain $P(T_1 > 2T_2)$.

Figure 4.30: Problem 1

(a)

Problem review:

T_1 is a random variable and T_2 is a random variable, where $T_1 = \alpha e^{-\alpha t_1}$ and $T_2 = \beta e^{-\beta t_2}$

α and β can be thought of as the failure rate for each respective component. T_i is the lifetime of component i. Hence $P(T_1 = t_1)$ means to ask for the probability of the first component to have a lifetime of t_1 given that the failure rate of this kind of components is α.

solution:

Now we know that

$$P(T_1 > T_2) = \int \int f_{T_1,T_2}(t_1,t_2)\, dt_2 dt_1$$

Looking at the following diagram to help determine the region to integrate:

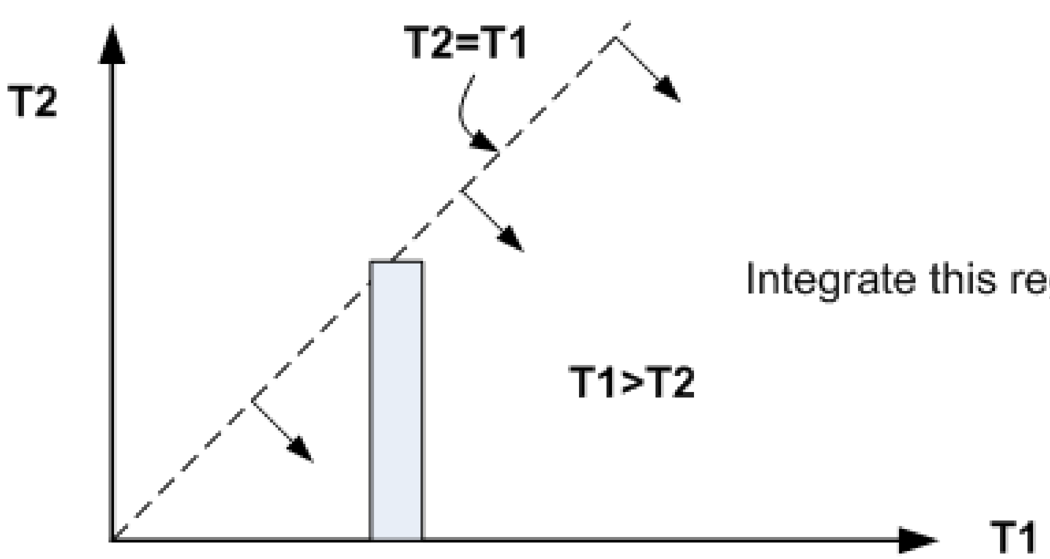


Figure 4.31: determine the region to integrate

Hence

$$P(T_1 > T_2) = \int_{t_1=0}^{t_1=\infty} \int_{t_2=0}^{t_2=t_1} f_{T_1,T_2}(t_1,t_2)\, dt_2\ dt_1$$

But since $T_1 \perp T_2$, then the joint density is the product of the marginal densites.

Hence

$$\begin{aligned} f_{T_1,T_2}(t_1,t_2) &= f_{T_1}(t_1)\, f_{T_2}(t_2) \\ &= \alpha e^{-\alpha t_1} \beta e^{-\beta t_2} \end{aligned}$$

Therefore

$$\begin{aligned} P(T_1 > T_2) &= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \beta e^{-\beta t_2}\ dt_2\ dt_1 \\ &= \beta\alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta t_2}\ dt_2 \right)\ dt_1 \\ &= \beta\alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{1}{\beta} \left[e^{-\beta t_2} \right]_{t_2=0}^{t_2=t_1} \right)\ dt_1 \\ &= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta t_1} - 1 \right]\ dt_1 \\ &= -\alpha \int_0^\infty e^{t_1(-\alpha-\beta)} - e^{-\alpha t_1}\ dt_1 \\ &= -\alpha \left(\left[\frac{1}{(-\alpha-\beta)} e^{t_1(-\alpha-\beta)} \right]_0^\infty - \frac{1}{-\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \end{aligned}$$

We take $\alpha, \beta \geq 0$ since we expect the lifetime to go to zero eventually. Also this is a requirment for the integrals to not diverge.

Hence the above becomes

$$\begin{aligned} P(T_1 > T_2) &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{t_1(-\alpha-\beta)} \right]_0^\infty + \frac{1}{\alpha} \left[e^{-\alpha t_1} \right]_0^\infty \right) \\ &= -\alpha \left(\frac{1}{(-\alpha-\beta)} \left[e^{-\infty} - 1 \right] + \frac{1}{\alpha} \left[e^{-\infty} - 1 \right] \right) \\ &= -\alpha \left(\frac{1}{(-\alpha-\beta)} [0-1] + \frac{1}{\alpha} [0-1] \right) \\ &= -\alpha \left(\frac{1}{(\alpha+\beta)} - \frac{1}{\alpha} \right) \\ &= -\alpha \left(\frac{\alpha - (\alpha+\beta)}{\alpha\,(\alpha+\beta)} \right) \\ &= -\left(\frac{\alpha - \alpha - \beta}{(\alpha+\beta)} \right) \end{aligned}$$

Hence

$$\boxed{P(T_1 > T_2) = \frac{\beta}{(\alpha+\beta)}}$$

(b)

$$\begin{aligned} F_W(w) &= P(W \leq w) \\ &= P(2T_2 \leq w) \\ &= P\left(T_2 \leq \frac{w}{2} \right) \\ &= F_{T_2}\left(\frac{w}{2} \right) \end{aligned}$$

Hence

$$f_W(w) = f_{T_2}\left(\frac{w}{2}\right) \times \frac{d}{dw}\left(\frac{w}{2}\right)$$

Hence

$$\boxed{f_W(w) = \tfrac{1}{2} f_{T_2}\left(\tfrac{w}{2}\right)}$$

(c)Need to find $P(T_1 > 2T_2)$ which is the same as $P(T_1 > W)$, hence this is the same as part(a) but replace T_2 by W as show in the following diagram

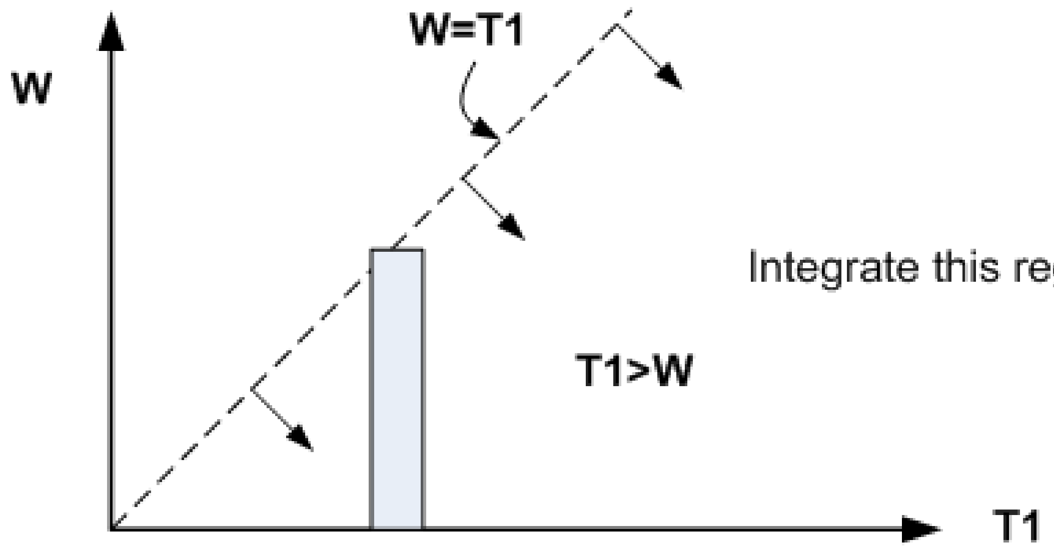


Figure 4.32: diagram

Hence

$$\begin{aligned}
P(T_1 > W) &= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1)\, f_W(w)\ dw\ dt_1 \\
&= \int_0^\infty \int_0^{t_1} f_{T_1}(t_1) \left[\frac{1}{2} f_{T_2}\left(\frac{w}{2}\right)\right]\ dw\ dt_1 \\
&= \int_0^\infty \int_0^{t_1} \alpha e^{-\alpha t_1} \left[\frac{1}{2}\beta e^{-\beta\left(\frac{w}{2}\right)}\right]\ dw\ dt_1 \\
&= \frac{1}{2}\beta\alpha \int_0^\infty e^{-\alpha t_1} \left(\int_0^{t_1} e^{-\beta\left(\frac{w}{2}\right)}\ dw\right)\ dt_1 \\
&= \frac{1}{2}\beta\alpha \int_0^\infty e^{-\alpha t_1} \left(-\frac{2}{\beta}\left[e^{-\beta\left(\frac{w}{2}\right)}\right]_{w=0}^{w=t_1}\right)\ dt_1 \\
&= -\alpha \int_0^\infty e^{-\alpha t_1} \left[e^{-\beta\left(\frac{t_1}{2}\right)} - 1\right]\ dt_1 \\
&= -\alpha \int_0^\infty e^{t_1\left(-\alpha - \frac{\beta}{2}\right)} - e^{-\alpha t_1}\ dt_1 \\
&= -\alpha \int_0^\infty e^{t_1\left(\frac{-2\alpha-\beta}{2}\right)} - e^{-\alpha t_1}\ dt_1 \\
&= -\alpha\left(\left[\frac{2}{(-2\alpha-\beta)} e^{t_1\left(\frac{-2\alpha-\beta}{2}\right)}\right]_0^\infty - \frac{1}{-\alpha}\left[e^{-\alpha t_1}\right]_0^\infty\right) \\
&= -\alpha\left(\frac{2}{(-2\alpha-\beta)}[0-1] + \frac{1}{\alpha}[0-1]\right) \\
&= -\alpha\left(\frac{2}{(2\alpha+\beta)} - \frac{1}{\alpha}\right)
\end{aligned}$$

Hence

$$\begin{aligned} P(T_1 > W) &= -\left(\frac{2\alpha - (2\alpha + \beta)}{(2\alpha + \beta)}\right) \\ &= -\left(\frac{2\alpha - 2\alpha - \beta}{(2\alpha + \beta)}\right) \end{aligned}$$

Then

$$\boxed{P(T_1 > W) = \frac{\beta}{(2\alpha+\beta)}}$$

2. Consider a Poisson process on the real line, and denote by $N(t_1, t_2)$ the number of events in the interval (t_1, t_2). If $t_0 < t_1 < t_2$, find the conditional distribution of $N(t_0, t_1)$ given that $N(t_0, t_2) = n$, and identify the distribution.

Figure 4.33: Problem 2

Problem review: Poisson probability density is a discrete probability function (We normally call it the probability mass function pmf). This means the random variable is a discrete random variable.

The random varible X in this case is the number of success in n trials where the probability of success in each one trial is p and the trials are independent from each others. The difference between Poisson and Binomial is that in Poisson we are looking at the problem as n becomes very large and p becomes very small in such a way that the product np goes to a fixed value which is called λ, the Poisson parameter. And then we write $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$ where $k = 0, 1, 2, \cdots$ The following diagram illustrates this problem, showing the three r.v. we need to analyze and the time line.

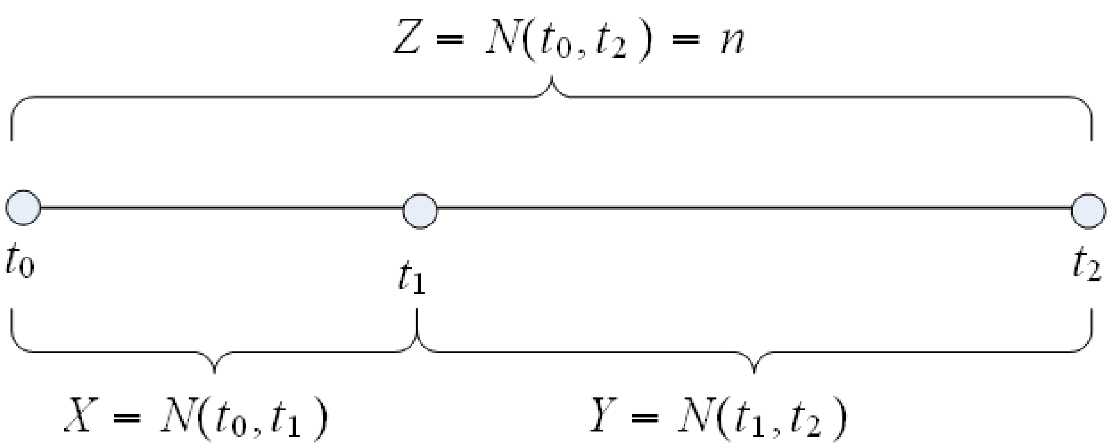


Figure 4.34: illustrates this problem

But what is "trials" in this problem? If we divide the time line itself into very small time intervals δt then the number of time intervals is the number of trials, and we assume that at most one event will occure in this time interval (since it is too small). The probability p of event occuring in this δt is the same in the interval $[t_0, t_1]$ and in the interval $[t_1, t_2]$. Now let us find λ for X and Y and Z based on this. Since $\lambda = np$ where n is the number of trials, then for X we have $\boxed{\lambda_x = n_x p = \frac{(t_1 - t_0)}{\delta t} p}$ where we divided the time interval by the time width δt to obtain the number of time slots for X. We do the same for Y and obtain that $\boxed{\lambda_y = \frac{(t_2 - t_1)}{\delta t} p}$

Similary, $\lambda_Z = \frac{(t_2-t_0)}{\delta t}p = \frac{(t_2-t_1)+(t_1-t_0)}{\delta t}p = \frac{(t_2-t_1)}{\delta t}p + \frac{(t_1-t_0)}{\delta t}P$, hence $\boxed{\lambda_z = \lambda_x + \lambda_y}$

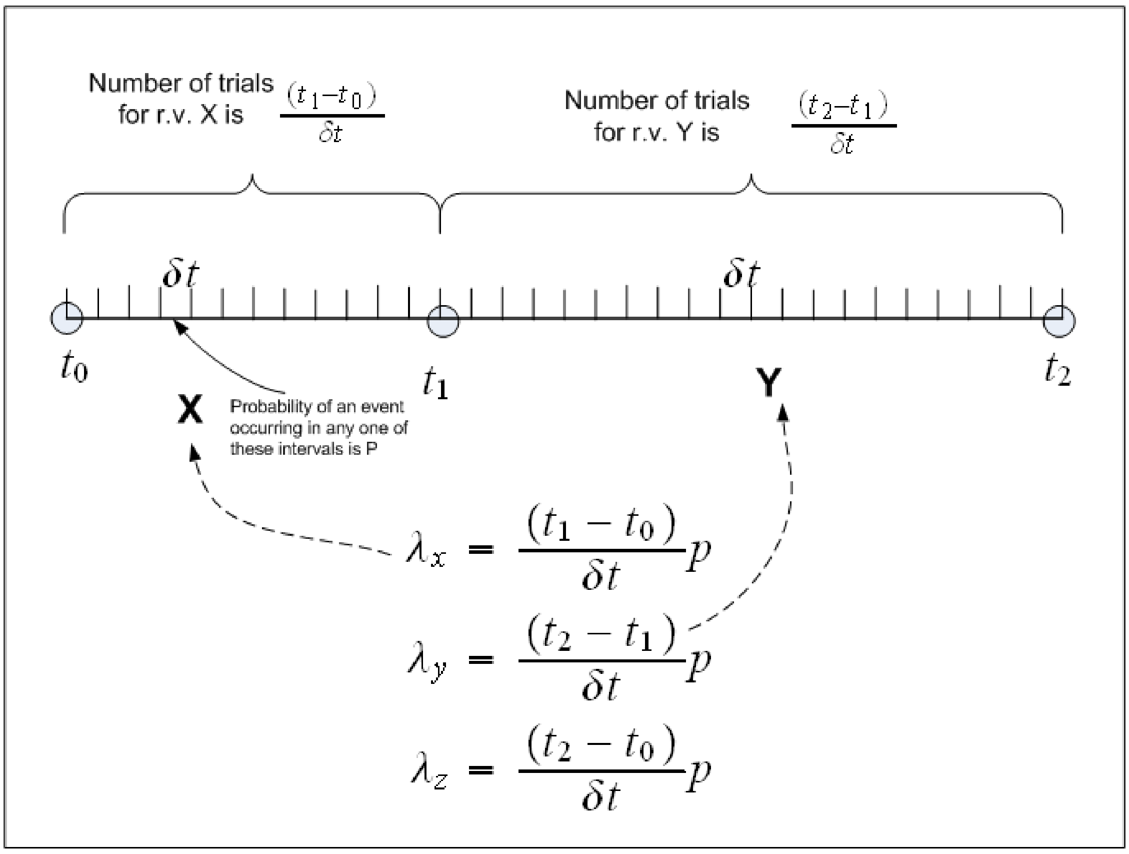


Figure 4.35: delta

$$\lambda_x = \frac{(t_1-t_0)}{\delta t}p$$
$$\lambda_y = \frac{(t_2-t_1)}{\delta t}p$$
$$\lambda_z = \frac{(t_2-t_0)}{\delta t}p$$

Let us refer to the random variable $N(t_1, t_2)$ as Y and the r.v. $N(t_0, t_1)$ as X and the r.v. $N(t_0, t_2)$ as Z

The problem is then asking to find $P(X = x|Z = n)$ and to identify $pmf(X|Z)$

To help in the solution, we first draw a diagram to make it more clear.

We take λ to the same for the 3 random variables X, Y, Z.

$$P(X = x|Z = n) = \frac{P(X = x, Z = n)}{P(Z = n)}$$

But $Z = n$ is the same as $X + Y = n$ hence

$$\begin{aligned} P(X = x|Z = n) &= \frac{P(X = x, (X + Y) = n)}{P(Z = n)} \\ &= \frac{P(X = x, Y = n - x)}{P(Z = n)} \end{aligned}$$

Now r.v. $X \perp Y$, since the number of events in $[t_0, t_1]$ is indepenent from the number of events that could occur in $[t_1, t_2]$.

Given this, we can now write the joint probability of X, Y as the product of the marginal probabilitites. Hence the numerator in the above can be rewritten and we obtain

$$P(X=x|Z=n)=\frac{P(X=x)\,P(Y=n-x)}{P(Z=n)} \tag{1}$$

Now since each of the above is a poisson process, then

$$\begin{aligned}
P(X=x) &= \frac{(\lambda_x)^x}{x!}e^{-\lambda_x}\\
P(Y=n-x) &= \frac{(\lambda_y)^{n-x}}{(n-x)!}e^{-\lambda_y}\\
P(Z=n) &= \frac{(\lambda_z)^n}{n!}e^{-\lambda_z}
\end{aligned}$$

Hence (1) becomes

$$P(X=x|Z=n)=\left(\frac{(\lambda_x)^x}{x!}e^{-\lambda_x}\right)\left(\frac{(\lambda_y)^{n-x}}{(n-x)!}e^{-\lambda_y}\right)\frac{1}{\frac{(\lambda_z)^n}{n!}e^{-\lambda_z}} \tag{2}$$

Now we simplify this further and try to idensity the resulting distribution. First we note

Hence (2) becomes

$$P(X=x|Z=n)=\left(\frac{\left(\frac{(t_1-t_0)}{\delta t}p\right)^x}{x!}e^{-\left(\frac{(t_1-t_0)}{\delta t}p\right)}\right)\left(\frac{\left(\frac{(t_2-t_1)}{\delta t}p\right)^{n-x}}{(n-x)!}e^{-\left(\frac{(t_2-t_1)}{\delta t}p\right)}\right)\frac{1}{\frac{\left(\frac{(t_2-t_0)}{\delta t}p\right)^n}{n!}e^{-\left(\frac{(t_2-t_0)}{\delta t}p\right)}}$$

Let $\frac{p}{\delta t}=\varphi$ then the above becomes

$$\begin{aligned}
P(X=x|Z=n) &= \left(\frac{((t_1-t_0)\,\varphi)^x}{x!}e^{-((t_1-t_0)\varphi)}\right)\left(\frac{((t_2-t_1)\,\varphi)^{n-x}}{(n-x)!}e^{-((t_2-t_1)\varphi)}\right)\frac{n!}{((t_2-t_0)\,\varphi)^n\,e^{-((t_2-t_0)\varphi)}}\\
&= \left(\frac{(t_1\varphi-t_0\varphi)^x}{x!}e^{-t_1\varphi+t_0\varphi}\right)\left(\frac{(t_2\varphi-t_1\varphi)^{n-x}}{(n-x)!}e^{-t_2\varphi+t_1\varphi}\right)\frac{n!}{(t_2\varphi-t_0\varphi)^n\,e^{-t_2\varphi+t_0\varphi}}\\
&= \left(\frac{(t_1\varphi-t_0\varphi)^x}{x!}e^{-t_1\varphi+t_0\varphi}\right)\left(\frac{(t_2\varphi-t_1\varphi)^{n-x}}{(n-x)!}e^{-t_2\varphi+t_1\varphi}\right)\frac{n!}{(t_2\varphi-t_0\varphi)^n}e^{t_2\varphi-t_0\varphi}\\
&= \left(\frac{(t_1\varphi-t_0\varphi)^x}{x!}\right)\left(\frac{(t_2\varphi-t_1\varphi)^{n-x}}{(n-x)!}\right)\frac{n!}{(t_2\varphi-t_0\varphi)^n}e^{(t_2\varphi-t_0\varphi-t_1\varphi+t_0\varphi-t_2\varphi+t_1\varphi)}\\
&= \left(\frac{(t_1\varphi-t_0\varphi)^x}{x!}\right)\left(\frac{(t_2\varphi-t_1\varphi)^{n-x}}{(n-x)!}\right)\frac{n!}{(t_2\varphi-t_0\varphi)^n}e^{(0)}\\
&= \left(\frac{(t_1\varphi-t_0\varphi)^x}{x!}\right)\left(\frac{(t_2\varphi-t_1\varphi)^{n-x}}{(n-x)!}\right)\frac{n!}{(t_2\varphi-t_0\varphi)^n}\\
&= \frac{n!}{x!\,(n-x)!}\frac{(t_1\varphi-t_0\varphi)^x\,(t_2\varphi-t_1\varphi)^{n-x}}{(t_2\varphi-t_0\varphi)^n}
\end{aligned}$$

We see that the parameter φ will occure in the numerator and denomerator with the same powers, hence we can factor it out and cancel it. Hence we obtain

$$P(X=x|Z=n)=\frac{n!}{x!\,(n-x)!}\frac{(t_1-t_0)^x\,(t_2-t_1)^{n-x}}{(t_2-t_0)^n}$$

Hence

$$\boxed{P(X=x|Z=n)=\binom{n}{x}\frac{(t_1-t_0)^x(t_2-t_1)^{n-x}}{(t_2-t_0)^n}}$$

$$P(X=x|Z=n)=\frac{n!}{x!\,(n-x)!}\left((\lambda_x)^x e^{-\lambda_x}\right)\left((\lambda_y)^{n-x}e^{-\lambda_y}\right)\frac{e^{\lambda_z}}{(\lambda_z)^n}$$

But we found that $\lambda_z=\lambda_x+\lambda_y$, hence the exponential term above vanish and we get

$$\begin{aligned}
P(X=x|Z=n) &= \frac{n!}{x!\,(n-x)!}\frac{(\lambda_x)^x(\lambda_y)^{n-x}}{(\lambda_z)^n}\\
&=\binom{n}{x}\frac{(\lambda_x)^x(\lambda_y)^{n-x}}{(\lambda_z)^n}\\
&=\binom{n}{x}\frac{(\lambda_x)^x(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^n}\\
&=\binom{n}{x}\frac{(\lambda_x)^x(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^x(\lambda_x+\lambda_y)^{n-x}}\\
&=\binom{n}{x}\frac{(\lambda_x)^x}{(\lambda_x+\lambda_y)^x}\frac{(\lambda_y)^{n-x}}{(\lambda_x+\lambda_y)^{n-x}}\\
&=\binom{n}{x}\left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x\left(\frac{\lambda_y}{\lambda_x+\lambda_y}\right)^{n-x}
\end{aligned}$$

Let $k=\frac{\lambda_x}{\lambda_x+\lambda_y}$, then $1-k=1-\frac{\lambda_x}{\lambda_x+\lambda_y}=\frac{\lambda_x+\lambda_y-\lambda_x}{\lambda_x+\lambda_y}=\frac{\lambda_y}{\lambda_x+\lambda_y}$ hence the last line above can be written as

$$\begin{aligned}
P(X=x|Z=n) &= \binom{n}{x}\left(\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^x\left(1-\frac{\lambda_x}{\lambda_x+\lambda_y}\right)^{n-x}\\
&=\binom{n}{x}(k)^x(1-k)^{n-x}
\end{aligned}$$

But this is a Binomial with parameters n,k, hence

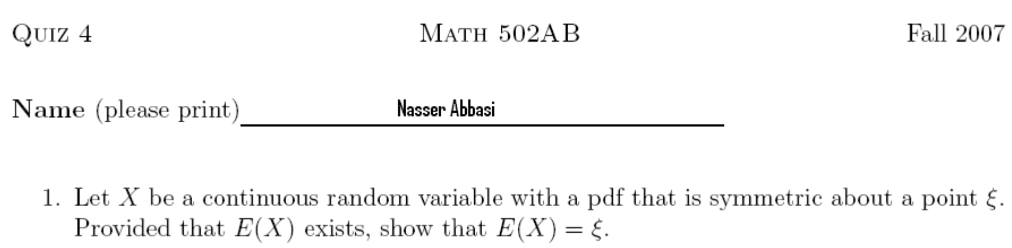
$$\boxed{P(X=x|Z=n)=Binomial\left(n,\frac{\lambda_x}{\lambda_x+\lambda_y}\right)}$$

3. Suppose that the probability Θ of getting heads for a coin is unknown, and let the prior opinion about Θ be represented by the uniform distribution on [0,1]. You spin the coin repeatedly and record the number of times N until a heads comes up. (a) Find the posterior density of Θ given N. (b) Use Matlab or any other software to plot the posterior for cases where $N=1$, $N=2$, and $N=6$. Using your plots, explain what you infer about the probability of heads in each circumstance.

Figure 4.36: Problem 3

4.4 Quiz 4

Local contents



QUIZ 4 MATH 502AB Fall 2007

Name (please print) Nasser Abbasi

1. Let X be a continuous random variable with a pdf that is symmetric about a point ξ. Provided that $E(X)$ exists, show that $E(X) = \xi$.

Figure 4.37: Problem 1

Let $f(x)$ be the pdf of X, hence from definition of expected value of a random variable we write

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

Now break the integral into the sum of integrals as follows

$$E(X) = \cdots + \int_{\xi-2\delta}^{\xi-\delta} x f(x)\, dx + \int_{\xi-\delta}^{\xi} x f(x)\, dx + \int_{\xi}^{\xi+\delta} x f(x)\, dx + \int_{\xi+\delta}^{\xi+2\delta} x f(x)\, dx + \cdots$$

In the limit, as δ is made very small, the above can be written as Riemann sums of areas each of width $dx \to \delta$ as follows

$$\begin{aligned}
E(X) &= \cdots + (\xi - 2\delta)\ f(\xi - 2\delta)\,\delta \ + \ (\xi-\delta)\ f(\xi-\delta)\,\delta \ + \ \xi f_\xi \delta + \\
&\quad (\xi+\delta)\, f(\xi+\delta)\,\delta + \ (\xi+2\delta)\ f\ (\xi+2\delta)\,\delta + \cdots \\
&= \delta[\cdots (\xi - 2\delta)\ f(\xi-2\delta) \ + \ (\xi-\delta)\ f(\xi-\delta) \ + \ \xi f_\xi + \\
&\quad (\xi+\delta)\, f(\xi+\delta) + \ (\xi+2\delta)\ f\ (\xi+2\delta) + \cdots] \\
&= \delta[\cdots + (\xi f(\xi-2\delta) - 2\delta f(\xi-2\delta)) \ + \ (\xi f(\xi-\delta) - \delta f(\xi-\delta)) \ + \\
&\quad \xi f_\xi \ + \ (\xi f(\xi+\delta) + \delta f(\xi+\delta)) + \ (\xi f\ (\xi+2\delta) + 2\delta f\ (\xi+2\delta)) \ + \cdots] \qquad (1)
\end{aligned}$$

But due to symmetry around ξ then

$$f(\xi - i\delta) = f(\xi + i\delta)$$

for any integer i in the above Riemann sum. This causes terms to cancel in the equation (1) above.

For example the term $-\delta f(\xi-\delta)$ onthe left of the ξf_ξ will cancel with the term $+\delta f(\xi-\delta)$ on the right of ξf_ξ, and so on. Then we obtain the following sum

$$E(X) = \delta[\cdots + \xi f(\xi-2\delta) \ + \ \xi f(\xi-\delta) \ + \ \xi f_\xi \ + \ \xi f(\xi+\delta) + \ \xi f\ (\xi+2\delta) \ + \cdots]$$

Take ξ as common factor

$$E(X) = \xi\delta[\cdots + \xi f(\xi - 2\delta) \ + \ \xi f(\xi - \delta) \ + \ \xi f_\xi \ + \ \xi f(\xi + \delta) + \ \xi f \ (\xi + 2\delta) \ + \cdots] \ (2)$$

But

$$\delta[\cdots + \xi f(\xi - 2\delta) \ + \ \xi f(\xi - \delta) \ + \ \xi f_\xi \ + \ \xi f(\xi + \delta) + \ \xi f \ (\xi + 2\delta) \ + \cdots]$$

is just the total area under $f(x)$ in the Riemann sum sense i.e. $\int_{-\infty}^{\infty} f(x)\,dx$.

Hence (2) becomes

$$E(X) = \xi \int_{-\infty}^{\infty} f(x)\,dx$$

But since $f(x)$ is a density, this area is one. Hence

$$\boxed{E(X) = \xi}$$

2. Let X be an exponential random variable with parameter λ. Find

$$P\left[|X - E(X)| > \frac{2}{\lambda}\right]$$

and compare your result to the Chebyshev's bound.

Figure 4.38: Problem 2

The density function of an exponential distribution with parameter $\lambda$is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

First find the expected values of an exponential random variable X. From definition of expected value:

$$\begin{aligned} E(X) &= \int_0^\infty x f(x)\,dx \\ &= \lambda \int_0^\infty x e^{-\lambda x} dx \end{aligned}$$

integrate by parts gives

$$\begin{aligned} E(X) &= \lambda\left(\left[\frac{x e^{-\lambda x}}{-\lambda}\right]_0^\infty + \frac{1}{\lambda}\int_0^\infty e^{-\lambda x} dx\right) \\ &= -\left[x e^{-\lambda x}\right]_0^\infty + \int_0^\infty e^{-\lambda x} dx \\ &= 0 - \frac{1}{\lambda}\left[e^{-\lambda x}\right]_0^\infty \\ &= \frac{1}{\lambda} \end{aligned}$$

Hence $\boxed{E(X)=\frac{1}{\lambda}}$, Hence we need to find $\Delta = P\left(\left|X-\frac{1}{\lambda}\right|>\frac{2}{\lambda}\right)$, But this is the same as finding

$$\begin{aligned}
\Delta &= 1-P\left(\left|X-\frac{1}{\lambda}\right|\leq\frac{2}{\lambda}\right)\\
&= 1-P\left(X<\frac{3}{\lambda}\right)\\
&= 1-\int_0^{\frac{3}{\lambda}} f(x)\,dx\\
&= 1-\int_0^{\frac{3}{\lambda}} \lambda e^{-\lambda x}dx\\
&= 1-\left(1-\frac{1}{e^3}\right)=\frac{1}{e^3}=\boxed{0.049787}
\end{aligned}$$

Now compare to Chebyshev bound. Chebyshev bound says that

$$P(|X-E(X)|\geq t)\leq\frac{Var(X)}{t^2} \tag{1}$$

Hence the upper bound by Chebyshev is $\frac{Var(X)}{\left(\frac{2}{\lambda}\right)^2}$. We now need to find $Var(X)$ and this is given by

$$Var(X)=E\left(X^2\right)-[E(X)]^2$$

But

$$\begin{aligned}
E\left(X^2\right) &= \int_0^\infty x^2 f(x)\,dx=\int_0^\infty x^2\lambda e^{-\lambda x}dx\\
&= \lambda\left[\frac{-1}{\lambda}\left[x^2e^{-\lambda x}\right]_0^\infty+\frac{2}{\lambda}\int_0^\infty xe^{-\lambda x}dx\right]\\
&= \left[-1[0]+2\int_0^\infty xe^{-\lambda x}dx\right]\\
&= 2\int_0^\infty xe^{-\lambda x}dx\\
&= 2\left[\frac{-1}{\lambda}\left[xe^{-\lambda x}\right]_0^\infty+\frac{1}{\lambda}\int_0^\infty e^{-\lambda x}dx\right]\\
&= 2\left[0+\frac{1}{\lambda}\left[\frac{e^{-\lambda x}}{-\lambda}\right]_0^\infty\right]\\
&= 2\left[-\frac{1}{\lambda^2}\left[e^{-\lambda\infty}-e^0\right]\right]\\
&= -\frac{2}{\lambda^2}[0-1]\\
&= \frac{2}{\lambda^2}
\end{aligned}$$

so

$$\begin{aligned}
Var(X) &= \frac{2}{\lambda^2}-\left[\frac{1}{\lambda}\right]^2\\
&= \boxed{\frac{1}{\lambda^2}}
\end{aligned}$$

Hence (1) becomes

$$P\left(|X-E(X)|\geq\frac{2}{\lambda}\right)\leq\frac{\frac{1}{\lambda^2}}{\frac{4}{\lambda^2}}=\boxed{0.25}$$

Hence an upper bound for the probability by Chebyshev is 0.25, and the actual probability found was 0.04 978 7 which is well within this bound.

3. If X is a discrete random variable, taking values on the positive integers, then show that $E(X) = \sum_{k=1}^{\infty} P(X \geq k)$.

Figure 4.39: Problem 3

Let $\Delta = \sum_{k=1}^{\infty} P(X \geq K)$, we need to show that this equals $E(X)$

$$\begin{aligned}\Delta &= \sum_{k=1}^{\infty} P(X \geq K)\\ &= P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \cdots\end{aligned}$$

But

$$P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + \cdots$$

and

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + \cdots$$

and

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + \cdots$$

and so on. Hence adding all the above we obtain repeated terms, which comes out as follows

$$\begin{aligned}\Delta &= P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \cdots\\ &= [P(X = 1) + P(X = 2) + P(X = 3) + \cdots]\\ &\quad + [P(X = 2) + P(X = 3) + P(X = 4) + \cdots]\\ &\quad + [P(X = 3) + P(X = 4) + P(X = 5) + \cdots]\\ &\quad + \cdots\\ &= P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + \cdots\\ &= \sum_{k=1}^{\infty} k\ P(X = k)\end{aligned}$$

But this is the definition of $E(X)$, hence $\Delta = E(X)$

4. Find the mean of a negative binomial random variable X with parameters r and p, by expressing X as sum of indicator variables.

Figure 4.40: Problem 4

X is Number of trials needed to obtain r successes, Each trial has p chance of success.

Let Y_1 be a random variable which represents the number of trials to obtain a success (counting the success trial) (This will be the first success).

Let Y_2 be a random variable which represents the number of trials to obtain a success (this will be the second success so far)

Let Y_3 be a random variable which represents the number of trials to obtain a success (this will be the third success so far)

and so on. Hence

Let Y_i be a random variable which represents the number of trials to obtain the i^{th} success.

Therefore

$$\begin{aligned} X &= Y_1 + Y_2 + \cdots + Y_r \\ &= \sum_{k=1}^{r} Y_r \end{aligned}$$

Hence

$$\begin{aligned} E(X) &= E\left(\sum_{k=1}^{r} Y_r\right) \\ &= \sum_{k=1}^{r} E(Y_r) \end{aligned} \tag{1}$$

But a Geometric r.v. represents the number of trials needed to obtain a success (counting the success trial), with each trial having p chance of success. So we need to find $E(Y)$ where Y is a Geometric r.v. with parameters p

$$E(Y) = \sum_{k=1}^{\infty} kP(X = K)$$

But

$$P(Y = K) = p(1-p)^k$$

Hence

$$\begin{aligned} E(Y) &= \sum_{k=1}^{\infty} kp(1-p)^k = p\sum_{k=1}^{\infty} k(1-p)^k \\ &= p\left(\frac{1-p}{p^2}\right) \\ &= \boxed{\frac{1-p}{p}} \end{aligned} \tag{2}$$

Substitute (2) into (1)

$$\begin{aligned} E(X) &= \sum_{k=1}^{r} \frac{1-p}{p} \\ &= \frac{1-p}{p} \sum_{k=1}^{r} 1 \\ &= \boxed{r\left(\frac{1-p}{p}\right)} \end{aligned}$$

5. If $U = a + bX$ and $V = c + dY$, show that $|\rho_{UV}| = |\rho_{XY}|$.

$$\rho_{U,V} = \frac{Cov(U,V)}{\sqrt{Var\,(U)\,Var\,(V)}} \tag{1}$$

But

$$Cov(U,V) = E(UV) - E(U)\,E(V)$$

and

$$\begin{aligned} E(U) &= E(a+bX) = E(a) + E(bX) \\ &= a + bE(X) \end{aligned}$$

and

$$\begin{aligned} E(V) &= E(c+dY) = E(c) + E(dY) \\ &= c + dE(Y) \end{aligned}$$

so

$$Cov(U,V) = E[(a+bX)\,(c+dY)] - [a+bE(X)]\,[c+dE(Y)] \tag{2}$$

and

$$Var(U) = Var(a+bX) = b^2 Var(X) \tag{3}$$

and

$$Var(V) = Var(c+dY) = d^2 Var(Y) \tag{4}$$

Substitute (2),(3),(4) into (1) we obtain

$$\begin{aligned} \rho_{U,V} &= \frac{E[(a+bX)\,(c+dY)] - [a+bE(X)]\,[c+dE(Y)]}{\sqrt{b^2 Var\,(X)\,d^2 Var\,(Y)}} \\ &= \frac{E[ac + adY + cbX + bXdY] - (ac + adE(Y) + cbE(X) + bdE(X)\,E(Y))}{|bd|\,\sqrt{Var\,(X)\,Var\,(Y)}} \\ &= \frac{ac + adE(Y) + cbE(X) + bdE(XY) - ac - adE(Y) - cbE(X) - bdE(X)\,E(Y)}{|bd|\,\sqrt{Var\,(X)\,Var\,(Y)}} \\ &= \frac{bdE(XY) - bdE(X)\,E(Y)}{|bd|\,\sqrt{Var\,(X)\,Var\,(Y)}} \\ &= \frac{bd[E(XY) - E(X)\,E(Y)]}{|bd|\,\sqrt{Var\,(X)\,Var\,(Y)}} \end{aligned}$$

Now cancel bd term. So depending if $bd < 0$ or $bd > 0$ we obtain $-\rho_{X,Y}$ or $+\rho_{X,Y}$

Hence if we consider absolute sign of bd we write

$$|\rho_{U,V}| = |\rho_{X,Y}|$$

4.4.1 Graded

19/20

19/20

QUIZ 4 MATH 502AB Fall 2007

Name (please print) Nasser Abbasi

1. Let X be a continuous random variable with a pdf that is symmetric about a point ξ. Provided that $E(X)$ exists, show that $E(X) = \xi$.

Let $f(x)$ be the pdf of X, hence from definition of expected value of a random variable we write

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

Now break the integral into the sum of integrals as follows

$$E(X) = \cdots + \int_{\xi-2\delta}^{\xi-\delta} x f(x)\, dx + \int_{\xi-\delta}^{\xi} x f(x)\, dx + \int_{\xi}^{\xi+\delta} x f(x)\, dx + \int_{\xi+\delta}^{\xi+2\delta} x f(x)\, dx + \cdots$$

In the limit, as δ is made very small, the above can be written as Riemann sums of areas each of width $dx \to \delta$ as follows

$$\begin{aligned} E(X) &= \cdots + (\xi - 2\delta)\ f(\xi - 2\delta)\,\delta\ +\ (\xi-\delta)\ f(\xi-\delta)\,\delta\ +\ \xi f_\xi \delta\ + \\ &\quad (\xi+\delta)\, f(\xi+\delta)\,\delta + \ (\xi+2\delta)\ f\ (\xi+2\delta)\,\delta + \cdots \\ &= \delta\left[\cdots(\xi-2\delta)\ f(\xi-2\delta)\ +\ (\xi-\delta)\ f(\xi-\delta)\ +\ \xi f_\xi\ + \right. \\ &\quad \left.(\xi+\delta)\, f(\xi+\delta) +\ (\xi+2\delta)\ f\ (\xi+2\delta)+\cdots\right] \\ &= \delta\left[\cdots + (\xi f(\xi-2\delta) - 2\delta f(\xi-2\delta))\ +\ (\xi f(\xi-\delta) - \delta f(\xi-\delta))\ + \right. \\ &\quad \left.\xi f_\xi\ +\ (\xi f(\xi+\delta) + \delta f(\xi+\delta)) +\ (\xi f\ (\xi+2\delta) + 2\delta f\ (\xi+2\delta))\ + \cdots\right] \qquad (1) \end{aligned}$$

But due to symmetry around ξ then

$$f(\xi - i\delta) = f(\xi + i\delta)$$

for any integer i in the above Riemann sum. This causes terms to cancel in the equation (1) above.

For example the term $-\delta f(\xi-\delta)$ onthe left of the ξf_ξ will cancel with the term $+\delta f(\xi-\delta)$ on the right of ξf_ξ, and so on. Then we obtain the following sum

$$E(X) = \delta\left[\cdots + \xi f(\xi-2\delta)\ +\ \xi f(\xi-\delta)\ +\ \xi f_\xi\ +\ \xi f(\xi+\delta) +\ \xi f\ (\xi+2\delta)\ + \cdots\right]$$

Take ξ as common factor

1

$$E(X) = \xi\delta\left[\cdots + \xi f(\xi - 2\delta) \;+ \xi f(\xi - \delta) \;+\; \xi f_{\xi} \;+ \xi f(\xi + \delta) + \xi f\;(\xi + 2\delta) \;+ \cdots\right] \quad (2)$$

But

$$\delta\left[\cdots + \xi f(\xi - 2\delta) \;+\; \xi f(\xi - \delta) \;+\; \xi f_{\xi} \;+\; \xi f(\xi + \delta) + \;\xi f\;(\xi + 2\delta) \;+ \cdots\right]$$

is just the total area under $f(x)$ in the Riemann sum sense i.e. $\int_{-\infty}^{\infty} f(x)\,dx$.

Hence (2) becomes

$$E(X) = \xi \int_{-\infty}^{\infty} f(x)\,dx$$

δ needs to go to zero

But since $f(x)$ is a density, this area is one. Hence

$$\boxed{E(X) = \xi}$$

I Consider this to be a very poor proof, even from practical point of view

2. Let X be an exponential random variable with parameter λ. Find

$$P\left[|X - E(X)| > \frac{2}{\lambda}\right]$$

and compare your result to the Chebyshev's bound.

The density function of an exponential distribution with parameter $\lambda$is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

First find the expected values of an exponential random variable X. From definition of expected value:

$$\begin{aligned} E(X) &= \int_0^\infty x f(x)\, dx \\ &= \lambda \int_0^\infty x e^{-\lambda x} dx \end{aligned}$$

integrate by parts gives

$$\begin{aligned} E(X) &= \lambda \left(\left[\frac{x e^{-\lambda x}}{-\lambda} \right]_0^\infty + \frac{1}{\lambda} \int_0^\infty e^{-\lambda x} dx \right) \\ &= -\left[x e^{-\lambda x} \right]_0^\infty + \int_0^\infty e^{-\lambda x} dx \\ &= 0 - \frac{1}{\lambda} \left[e^{-\lambda x} \right]_0^\infty \\ &= \frac{1}{\lambda} \end{aligned}$$

Hence $\boxed{E(X) = \frac{1}{\lambda}}$, Hence we need to find $\Delta = P\left(\left|X - \frac{1}{\lambda}\right| > \frac{2}{\lambda}\right)$, But this is the same as finding

$$\begin{aligned} \Delta &= 1 - P\left(\left|X - \frac{1}{\lambda}\right| \leq \frac{2}{\lambda}\right) \\ &= 1 - P\left(X < \frac{3}{\lambda}\right) \\ &= 1 - \int_0^{\frac{3}{\lambda}} f(x)\, dx \\ &= 1 - \int_0^{\frac{3}{\lambda}} \lambda e^{-\lambda x} dx \\ &= 1 - \left(1 - \frac{1}{e^3}\right) = \frac{1}{e^3} = \boxed{0.04\,978\,7} \end{aligned}$$

Now compare to Chebyshev bound. Chebyshev bound says that

$$P\left(|X-E(X)|\geq t\right)\leq\frac{Var\left(X\right)}{t^{2}} \tag{1}$$

Hence the upper bound by Chebyshev is $\frac{Var(X)}{\left(\frac{2}{\lambda}\right)^2}$. We now need to find $Var\left(X\right)$ and this is given by

$$Var\left(X\right)=E\left(X^{2}\right)-\left[E\left(X\right)\right]^{2}$$

But

$$\begin{aligned}
E\left(X^{2}\right) &= \int_{0}^{\infty}x^{2}f\left(x\right)dx=\int_{0}^{\infty}x^{2}\lambda e^{-\lambda x}dx \\
&= \lambda\left[\frac{-1}{\lambda}\left[x^{2}e^{-\lambda x}\right]_{0}^{\infty}+\frac{2}{\lambda}\int_{0}^{\infty}xe^{-\lambda x}dx\right] \\
&= \left[-1\left[0\right]+2\int_{0}^{\infty}xe^{-\lambda x}dx\right] \\
&= 2\int_{0}^{\infty}xe^{-\lambda x}dx \\
&= 2\left[\frac{-1}{\lambda}\left[xe^{-\lambda x}\right]_{0}^{\infty}+\frac{1}{\lambda}\int_{0}^{\infty}e^{-\lambda x}dx\right] \\
&= 2\left[0+\frac{1}{\lambda}\left[\frac{e^{-\lambda x}}{-\lambda}\right]_{0}^{\infty}\right] \\
&= 2\left[-\frac{1}{\lambda^{2}}\left[e^{-\lambda\infty}-e^{0}\right]\right] \\
&= -\frac{2}{\lambda^{2}}\left[0-1\right] \\
&= \frac{2}{\lambda^{2}}
\end{aligned}$$

so

$$\begin{aligned}
Var\left(X\right) &= \frac{2}{\lambda^{2}}-\left[\frac{1}{\lambda}\right]^{2} \\
&= \boxed{\frac{1}{\lambda^{2}}}
\end{aligned}$$

Hence (1) becomes

$$P\left(|X-E\left(X\right)|\geq\frac{2}{\lambda}\right)\leq\frac{\frac{1}{\lambda^{2}}}{\frac{4}{\lambda^{2}}}=\boxed{0.25}$$

Hence an upper bound for the probability by Chebyshev is 0.25, and the actual probability found was 0.04 978 7 which is well within this bound.

4

3. If X is a discrete random variable, taking values on the positive integers, then show that $E(X) = \sum_{k=1}^{\infty} P(X \geq k)$.

Let $\Delta = \sum_{k=1}^{\infty} P(X \geq K)$, we need to show that this equals $E(X)$

$$\begin{aligned}\Delta &= \sum_{k=1}^{\infty} P(X \geq K)\\ &= P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \cdots\end{aligned}$$

But

$$P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + \cdots$$

and

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + \cdots$$

and

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) + \cdots$$

and so on. Hence adding all the above we obtain repeated terms, which comes out as follows

$$\begin{aligned}\Delta &= P(X \geq 1) + P(X \geq 2) + P(X \geq 3) + \cdots\\ &= [P(X = 1) + P(X = 2) + P(X = 3) + \cdots]\\ &+ [P(X = 2) + P(X = 3) + P(X = 4) + \cdots]\\ &+ [P(X = 3) + P(X = 4) + P(X = 5) + \cdots]\\ &+ \cdots\\ &= P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + \cdots\\ &= \sum_{k=1}^{\infty} k\, P(X = k)\end{aligned}$$

But this is the definition of $E(X)$, hence $\Delta = E(X)$

5

4. Find the mean of a negative binomial random variable X with parameters r and p, by expressing X as sum of indicator variables.

X is Number of trials needed to obtain r successes, Each trial has p chance of success.

Let Y_1 be a random variable which represents the number of trials to obtain a success (counting the success trial) (This will be the first success).

Let Y_2 be a random variable which represents the number of trials to obtain a success (this will be the second success so far)

Let Y_3 be a random variable which represents the number of trials to obtain a success (this will be the third success so far)

and so on. Hence

Let Y_i be a random variable which represents the number of trials to obtain the i^{th} success.

Therefore

$$\begin{aligned} X &= Y_1 + Y_2 + \cdots + Y_r \\ &= \sum_{k=1}^{r} Y_r \end{aligned}$$

Hence

$$\begin{aligned} E(X) &= E\left(\sum_{k=1}^{r} Y_r\right) \\ &= \sum_{k=1}^{r} E(Y_r) \end{aligned} \tag{1}$$

But a Geometric r.v. represents the number of trials needed to obtain a success (counting the success trial), with each trial having p chance of success. So we need to find $E(Y)$ where Y is a Geometric r.v. with parameters p

$$E(Y) = \sum_{k=1}^{\infty} kP(X = K)$$

But

$$P(Y = K) = p(1-p)^k$$

Hence

$$\begin{aligned} E(Y) &= \sum_{k=1}^{\infty} kp(1-p)^k = p\sum_{k=1}^{\infty} k(1-p)^k \\ &= p\left(\frac{1-p}{p^2}\right) \\ &= \boxed{\frac{1-p}{p}} \end{aligned} \tag{2}$$

NO

6

Substitute (2) into (1)

$$
\begin{aligned}
E(X) &= \sum_{k=1}^{r} \frac{1-p}{p} \\
&= \frac{1-p}{p} \sum_{k=1}^{r} 1 \\
&= \boxed{r\left(\frac{1-p}{p}\right)}
\end{aligned}
$$

5. If $U = a + bX$ and $V = c + dY$, show that $|\rho_{UV}| = |\rho_{XY}|$.

$$\rho_{U,V} = \frac{Cov\left(U,V\right)}{\sqrt{Var\left(U\right)Var\left(V\right)}} \tag{1}$$

But

$$Cov\left(U,V\right) = E\left(UV\right) - E\left(U\right)E\left(V\right)$$

and

$$\begin{aligned} E\left(U\right) &= E\left(a+bX\right) = E\left(a\right) + E\left(bX\right) \\ &= a + bE\left(X\right) \end{aligned}$$

and

$$\begin{aligned} E\left(V\right) &= E\left(c+dY\right) = E\left(c\right) + E\left(dY\right) \\ &= c + dE\left(Y\right) \end{aligned}$$

so

$$Cov\left(U,V\right) = E\left[\left(a+bX\right)\left(c+dY\right)\right] - \left[a+bE\left(X\right)\right]\left[c+dE\left(Y\right)\right] \tag{2}$$

and

$$Var\left(U\right) = Var\left(a+bX\right) = b^2 Var\left(X\right) \tag{3}$$

and

$$Var\left(V\right) = Var\left(c+dY\right) = d^2 Var\left(Y\right) \tag{4}$$

Substitute (2),(3),(4) into (1) we obtain

$$\begin{aligned} \rho_{U,V} &= \frac{E\left[\left(a+bX\right)\left(c+dY\right)\right] - \left[a+bE\left(X\right)\right]\left[c+dE\left(Y\right)\right]}{\sqrt{b^2 Var\left(X\right)d^2 Var\left(Y\right)}} \\ &= \frac{E\left[ac+adY+cbX+bXdY\right] - \left(ac+adE\left(Y\right)+cbE\left(X\right)+bdE\left(X\right)E\left(Y\right)\right)}{|bd|\sqrt{Var\left(X\right)Var\left(Y\right)}} \\ &= \frac{ac+adE\left(Y\right)+cbE\left(X\right)+bdE\left(XY\right)-ac-adE\left(Y\right)-cbE\left(X\right)-bdE\left(X\right)E\left(Y\right)}{|bd|\sqrt{Var\left(X\right)Var\left(Y\right)}} \\ &= \frac{bdE\left(XY\right)-bdE\left(X\right)E\left(Y\right)}{|bd|\sqrt{Var\left(X\right)Var\left(Y\right)}} \\ &= \frac{bd\left[E\left(XY\right)-E\left(X\right)E\left(Y\right)\right]}{|bd|\sqrt{Var\left(X\right)Var\left(Y\right)}} \end{aligned}$$

Working too hard

Now cancel bd term. So depending if $bd < 0$ or $bd > 0$ we obtain $-\rho_{X,Y}$ or $+\rho_{X,Y}$
Hence if we consider absolute sign of bd we write

$$|\rho_{U,V}| = |\rho_{X,Y}|$$

8

4.5 Quiz 5

Local contents

QUIZ 5 MATH 502AB October 20, 2006

Name (please print) NASSER ABBASI

1. The moment generating function for a random variable X having a χ^2 distribution with degrees of freedom $n \geq 1$ is $M_X(t) = (1-2t)^{-n/2}$. Let W have a χ^2 distribution with degrees of freedom $n > 1$, and let V have a χ^2 distribution with degrees of freedom 1. (a). If $W = U + V$, and U and V are independent, determine the distribution of U? (b). What are the mean and variance of W?

Figure 4.41: Problem 1

(a) Consider

$$
\begin{aligned}
M_W(t) &= E\left(e^{wt}\right) \\
&= E\left(e^{(u+v)t}\right) = E\left(e^{ut}e^{vt}\right)
\end{aligned}
$$

But since $U \perp V$ then the above reduces to

$$
\begin{aligned}
M_W(t) &= E\left(e^{ut}\right) E\left(e^{vt}\right) \\
&= M_U(t)\, M_V(t)
\end{aligned}
$$

Hence

$$
\begin{aligned}
M_U(t) &= \frac{M_W(t)}{M_U(t)} = \frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}} \\
&= (1-2t)^{-\frac{n}{2}+\frac{1}{2}} \\
&= (1-2t)^{-\frac{(n-1)}{2}}
\end{aligned}
$$

Hence

$$
U\chi^2\text{with}(n-1)\text{ degrees of freedom}
$$

(b)

$$
\begin{aligned}
E(W) &= E(U+V) \\
&= E(U) + E(V)
\end{aligned}
$$

Now use the moment generation function to find the expectations of U and V.

Need to find $M_X'(t)$ where $M_X(t) = (1-2t)^{-\frac{m}{2}}$ where m is the degree of freedom

$$
\begin{aligned}
\frac{d}{dt}M_X(t) &= \frac{d}{dt}(1-2t)^{-\frac{m}{2}} \\
&= -\frac{m}{2}(1-2t)^{-\frac{m}{2}-1}(-2)
\end{aligned}
$$

Hence

$$M'(t) = m(1-2t)^{-\frac{m}{2}-1} \tag{1}$$

at $t=0$ the above becomes

$$M'_X(0) = m$$

For U, we found that $m=(n-1)$, hence

$$\boxed{E(U) = (n-1)}$$

and for V we are told its degree of freedom is $m=1$ hence

$$E(V) = 1$$

Therefore

$$E(W) = (n-1)+1$$

Hence

$$E(W) = n$$

Now

$$\begin{aligned}
Var(W) &= E(W^2) - [E(W)]^2 \\
&= E(W^2) - n^2 \\
&= E((U+V)^2) - n^2 \\
&= E((U^2+V^2+2UV)) - n^2 \\
&= E(U^2) + E(V^2) + 2E(UV) - n^2
\end{aligned}$$

But $U \perp V$ so the above becomes

$$Var(W) = E(U^2) + E(V^2) + 2E(U)\,E(V) - n^2$$

Lets find $E(Z^2)$ for a Z chi square random variable of degree of freedom m. We already found $M'(t)$ above in (1)

$$\begin{aligned}
E(Z^2) &= M''_Z(t)\Big|_{t=0} \\
&= \frac{d}{dt}(M'_Z(t)) \\
&= \frac{d}{dt}\left(m(1-2t)^{-\frac{m}{2}-1}\right) \\
&= m\left(\left(-\frac{m}{2}-1\right)(1-2t)^{-\frac{m}{2}-2}(-2)\right)
\end{aligned}$$

At $t=0$

$$E(Z^2) = -2m\left(-\frac{m}{2}-1\right)$$

Hence

$$E\left(Z^2\right) = m(m+2) \tag{2}$$

Hence using (2) above, we now can find $E(U^2)$ and $E(V^2)$

For U it has degree of freedom $m = (n-1)$, hence

$$\begin{aligned} E(U^2) &= (n-1)\left((n-1)+2\right) \\ &= n^2 - 1 \end{aligned}$$

For V it has degree of freedom $m = 1$, hence

$$\begin{aligned} E(V^2) &= 1 \times (1+2) \\ &= 3 \end{aligned}$$

Hence

$$\begin{aligned} Var(W) &= \left(n^2 - 1\right) + 3 + 2(n-1) \times 1 - n^2 \\ &= n^2 - 1 + 3 + 2n - 2 - n^2 \end{aligned}$$

Hence

$$Var(W) = 2n$$

2. Find the approximate variance of $Y = \sqrt{X}$, where X is a Poisson random variable with parameter λ.

Figure 4.42: Problem 2

$$X(x) \frac{\lambda^x e^{-\lambda}}{x!}$$

Moment generating function for a Poisson r.v. or parameter λ is (from page 144)

$$\boxed{M_X(t) = e^{-\lambda} e^{e^t \lambda}}$$

Now

$$\begin{aligned} M_Y(t) &= E\left(e^{yt}\right) \\ &= E\left(e^{\sqrt{x}t}\right) \end{aligned}$$

Hence

$$M_Y'(t) = E\left(\sqrt{x} e^{\sqrt{x}t}\right)$$

and

$$\begin{aligned} M_Y''(t) &= E\left(\sqrt{x}\sqrt{x}e^{\sqrt{x}t}\right) \\ &= E\left(xe^{\sqrt{x}t}\right) \end{aligned}$$

Therefore

$$M_Y''(0) = E(x)$$

But

$$\begin{aligned} E(x) &= M_X'(t)|_{t=0} \\ &= \lambda \end{aligned}$$

Hence

$$M_Y''(0) = \lambda$$

But

$$M_Y''(0) = E\left(Y^2\right)$$

then

$$\boxed{E(Y^2) = \lambda}$$

Now to find $Var(Y)$

$$Var(Y) = E\left(Y^2\right) - [E(Y)]^2$$

Where

$$\begin{aligned} E(Y) &= M_Y'(0) \\ &= E\left(\sqrt{x}\right) \end{aligned}$$

So we need to find $E\left(\sqrt{x}\right)$ to complete the solution.

$$\begin{aligned}
E(\sqrt{x}) &= \sum_{x=0}^{\infty}\sqrt{x}\frac{\lambda^x e^{-\lambda}}{x!}\\
&= 0+\frac{\lambda e^{-\lambda}}{1!}+\sqrt{2}\frac{\lambda^2 e^{-\lambda}}{2!}+\sqrt{3}\frac{\lambda^3 e^{-\lambda}}{3!}+\sqrt{4}\frac{\lambda^4 e^{-\lambda}}{4!}+\cdots\\
&= e^{-\lambda}\left(\lambda+\sqrt{2}\frac{\lambda^2}{2!}+\sqrt{3}\frac{\lambda^3}{3!}+\sqrt{4}\frac{\lambda^4}{4!}+\cdots\right)\\
&= \lambda e^{-\lambda}\left(1+\frac{1}{\sqrt{2}}\lambda+\frac{1}{\sqrt{3}}\frac{\lambda^2}{2!}+\frac{1}{\sqrt{4}}\frac{\lambda^3}{3!}+\frac{1}{\sqrt{5}}\frac{\lambda^4}{4!}+\cdots\right)\\
&= \lambda e^{-\lambda}\left(1+\frac{\lambda}{\sqrt{2}}+\frac{2}{\sqrt{3}}\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^2}{2!}+\frac{2\sqrt{2}}{\sqrt{4}}\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^3}{3!}+\frac{4}{\sqrt{5}}\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^4}{4!}+\cdots\right)\\
&= \lambda e^{-\lambda}\left(1+\frac{\lambda}{\sqrt{2}}+1.1547\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^2}{2!}+1.4142\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^3}{3!}+1.7889\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^4}{4!}+\cdots\right)\\
&\simeq \lambda e^{-\lambda}\left(e^{\frac{\lambda}{\sqrt{2}}}\right)\\
&= \lambda e^{\frac{\lambda}{\sqrt{2}}-\lambda} = \lambda e^{\frac{\lambda(1-\sqrt{2})}{\sqrt{2}}}
\end{aligned}$$

Hence

$$\begin{aligned}
Var(Y) &= \lambda-\left[\lambda e^{\frac{\lambda(1-\sqrt{2})}{\sqrt{2}}}\right]^2\\
&= \lambda-\lambda^2 e^{\frac{2\lambda(1-\sqrt{2})}{\sqrt{2}}}\\
&= \lambda-\lambda^2 e^{\sqrt{2}\lambda(1-\sqrt{2})}
\end{aligned}$$

Hence

$$Var(Y) \simeq \lambda\left(1-\lambda e^{-0.58578\lambda}\right)$$

3. The random variable Y has a Gamma distribution with parameters α and λ. Furthermore, assume that X given Y has a Poisson distribution with parameter Y^2. (a) Obtain $E(X)$. (b) Obtain $Var(X)$.

Figure 4.43: Problem 3

(a)

$$\begin{aligned}
f_Y(y) &= \frac{\lambda^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\lambda y} && y\geq 0\\
f_{(X|Y=y)}(x|y) &= \frac{\left(y^2\right)^x e^{-y^2}}{y!} && x=0,1,2,\cdots
\end{aligned}$$

Now

$$E(X)=E(E(X|Y))$$

But $E(X|Y)$ is expectation of a Poisson r.v. with parameter Y^2. But we know that mean of a poisson r.v. with parameter λ is λ. Hence $E(X|Y)=Y^2$ since we are told Y^2 is the parameter.

Hence

$$E(X) = E(Y^2)$$

But the moment generating function for Gamma is $M_Y(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha$ (book page 145 second edition).

Hence $E(Y^2) = M_Y''(0) = \frac{\alpha(\alpha+1)}{\lambda^2}$ (page 145)

Hence

$$\boxed{E(X) = \frac{\alpha(\alpha+1)}{\lambda^2}}$$

(b)

$$Var(X) = E(X^2) - [E(X)]^2 \tag{1}$$

But

$$E(X^2) = E(E(X^2|Y))$$

But $E(X^2|Y)$ is $E(X^2)$ of a poisson r.v. with parameter Y^2. But we know that $E(X^2)$ of a poisson r.v. with parameter λ is $\lambda^2 + \lambda$ (book page 144 example A). Hence since we are told Y^2 is the parameter, then

$$\begin{aligned} E(X^2|Y) &= (Y^2)^2 + Y^2 \\ &= (Y^4 + Y^2) \end{aligned}$$

Hence

$$\begin{aligned} E(X^2) &= E(Y^4 + Y^2) \\ &= E(Y^4) + E(Y^2) \end{aligned}$$

But using mgf for Gamma distribution we can find $E(Y^4)$.

$$\begin{aligned} M_Y''''(t) &= \frac{d^4}{dt^4}\left(\frac{\lambda}{\lambda - t}\right)^\alpha \\ &= \frac{\alpha(6 + 11\alpha + 6\alpha^2 + \alpha^3)}{(t-\lambda)^4}\left(\frac{\lambda}{\lambda - t}\right)^\alpha \end{aligned}$$

Then

$$M_Y''''(0) = \frac{\alpha(6 + 11\alpha + 6\alpha^2 + \alpha^3)}{\lambda^4}$$

Therefore

$$E(X^2) = \frac{\alpha(6 + 11\alpha + 6\alpha^2 + \alpha^3)}{\lambda^4} + \frac{\alpha(\alpha+1)}{\lambda^2}$$

Then (1) becomes

$$\begin{aligned} Var(X) &= \left(\frac{\alpha(6 + 11\alpha + 6\alpha^2 + \alpha^3)}{\lambda^4} + \frac{\alpha(\alpha+1)}{\lambda^2}\right) - \left(\frac{\alpha(\alpha+1)}{\lambda^2}\right)^2 \\ &= \frac{\alpha(6 + 11\alpha + 6\alpha^2 + \alpha^3)}{\lambda^4} + \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2(\alpha+1)^2}{\lambda^4} \end{aligned}$$

Then

$$\boxed{Var(X) = \frac{\alpha}{\lambda^4}(\alpha+1)\left(\lambda^2+4\alpha+6\right)}$$

4.5.1 Graded

20/20

Nasser, I am going to start penalizing you for material that are not relevant or don't need explanation. Be brief and to the point. Thanks

20/20

QUIZ 5 MATH 502AB October 20, 2006

Name (please print) NASSER ABBASI

1. The moment generating function for a random variable X having a χ^2 distribution with degrees of freedom $n \geq 1$ is $M_X(t) = (1-2t)^{-n/2}$. Let W have a χ^2 distribution with degrees of freedom $n > 1$, and let V have a χ^2 distribution with degrees of freedom 1. (a). If $W = U + V$, and U and V are independent, determine the distribution of U? (b). What are the mean and variance of W?

(a) Consider

$$\begin{aligned} M_W(t) &= E\left(e^{wt}\right) \\ &= E\left(e^{(u+v)t}\right) = E\left(e^{ut}e^{vt}\right) \end{aligned}$$

But since $U \perp V$ then the above reduces to

$$\begin{aligned} M_W(t) &= E\left(e^{ut}\right)E\left(e^{vt}\right) \\ &= M_U(t)\,M_V(t) \end{aligned}$$

Hence

$$\begin{aligned} M_U(t) &= \frac{M_W(t)}{M_V(t)} = \frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}} \\ &= (1-2t)^{-\frac{n}{2}+\frac{1}{2}} \\ &= (1-2t)^{-\frac{(n-1)}{2}} \end{aligned}$$

Hence

$$\boxed{U \sim \chi^2 \text{ with } (n-1) \text{ degrees of freedom}}$$

(b)

$$\begin{aligned} E(W) &= E(U+V) \\ &= E(U) + E(V) \end{aligned}$$

Now use the moment generation function to find the expectations of U and V.
Need to find $M_X'(t)$ where $M_X(t) = (1-2t)^{-\frac{m}{2}}$ where m is the degree of freedom

$$\begin{aligned} \frac{d}{dt}M_X(t) &= \frac{d}{dt}(1-2t)^{-\frac{m}{2}} \\ &= -\frac{m}{2}(1-2t)^{-\frac{m}{2}-1}(-2) \end{aligned}$$

Hence

1

$$\boxed{M'(t) = m(1-2t)^{-\frac{m}{2}-1}} \qquad (1)$$

at $t=0$ the above becomes

$$\boxed{M'_X(0) = m}$$

For U, we found that $m = (n-1)$, hence

$$\boxed{E(U) = (n-1)}$$

and for V we are told its degree of freedom is $m=1$ hence

$$\boxed{E(V) = 1}$$

Therefore

$$E(W) = (n-1) + 1$$

Hence

$$\boxed{E(W) = n}$$

Now

Working too hard Again!

$$\begin{aligned}
Var(W) &= E\left(W^2\right) - [E(W)]^2 \\
&= E\left(W^2\right) - n^2 \\
&= E\left((U+V)^2\right) - n^2 \\
&= E\left(\left(U^2+V^2+2UV\right)\right) - n^2 \\
&= E\left(U^2\right) + E\left(V^2\right) + 2E(UV) - n^2
\end{aligned}$$

But $U \perp V$ so the above becomes

$$\boxed{Var(W) = E(U^2) + E(V^2) + 2E(U)E(V) - n^2}$$

Lets find $E(Z^2)$ for a Z chi square random variable of degree of freedom m. We already found $M'(t)$ above in (1)

$$\begin{aligned}
E\left(Z^2\right) &= M''_Z(t)\Big|_{t=0} \\
&= \frac{d}{dt}\left(M'_Z(t)\right) \\
&= \frac{d}{dt}\left(m(1-2t)^{-\frac{m}{2}-1}\right) \\
&= m\left(\left(-\frac{m}{2}-1\right)(1-2t)^{-\frac{m}{2}-2}(-2)\right)
\end{aligned}$$

At $t=0$

$$E\left(Z^2\right) = -2m\left(-\frac{m}{2}-1\right)$$

Hence

$$\boxed{E\left(Z^2\right) = m\left(m+2\right)} \tag{2}$$

Hence using (2) above, we now can find $E\left(U^2\right)$ and $E\left(V^2\right)$
For U it has degree of freedom $m = (n-1)$, hence

$$\begin{aligned} E(U^2) &= (n-1)\left((n-1)+2\right) \\ &= n^2 - 1 \end{aligned}$$

For V it has degree of freedom $m = 1$, hence

$$\begin{aligned} E(V^2) &= 1 \times (1+2) \\ &= 3 \end{aligned}$$

Hence

$$\begin{aligned} Var(W) &= \left(n^2 - 1\right) + 3 + 2\left(n-1\right) \times 1 - n^2 \\ &= n^2 - 1 + 3 + 2n - 2 - n^2 \end{aligned}$$

Hence

$$\boxed{Var(W) = 2n}$$

Why did I do it so long? just use M(t)!

2. Find the approximate variance of $Y = \sqrt{X}$, where X is a Poisson random variable with parameter λ.

$$X(x) \sim \frac{\lambda^x e^{-\lambda}}{x!}$$

Moment generating function for a Poisson r.v. or parameter λ is (from page 144)

$$\boxed{M_X(t) = e^{-\lambda} e^{e^t \lambda}}$$

Now

$$\begin{aligned} M_Y(t) &= E\left(e^{yt}\right) \\ &= E\left(e^{\sqrt{x}t}\right) \end{aligned}$$

Hence

$$M_Y'(t) = E\left(\sqrt{x} e^{\sqrt{x}t}\right)$$

and

$$\begin{aligned} M_Y''(t) &= E\left(\sqrt{x}\sqrt{x} e^{\sqrt{x}t}\right) \\ &= E\left(x e^{\sqrt{x}t}\right) \end{aligned}$$

Therefore

$$M_Y''(0) = E(x)$$

But

$$\begin{aligned} E(x) &= M_X'(t)\big|_{t=0} \\ &= \lambda \end{aligned}$$

Hence

$$M_Y''(0) = \lambda$$

But

$$M_Y''(0) = E\left(Y^2\right)$$

then

$$\boxed{E\left(Y^2\right) = \lambda}$$

Now to find $Var(Y)$

$$Var(Y) = E\left(Y^2\right) - [E(Y)]^2$$

Where

$E(Y) = M_Y'(0) = E(\sqrt{X})$.

so need to find $E(\sqrt{X})$ to complete solution.

$\boxed{Y = \sqrt{X} = g(X)}$

expand $g(X)$ around μ_X using taylor series

so $g(X)$ near μ_X can be approximated as

$$g(X) \approx g(\mu_X) + (X-\mu_X)\, g'(\mu_X) + \frac{(X-\mu_X)^2}{2!} g''(\mu_X)$$

Note $\mu_X = \lambda$

so $g(X) \approx g(\lambda) + (X-\lambda) \frac{d}{dx} x^{\frac{1}{2}}\Big|_{x=\lambda} + \frac{(X-\lambda)^2}{2!} \frac{d^2}{dx^2} x^{\frac{1}{2}}\Big|_{x=\lambda}$

$$g(X) = \sqrt{\lambda} + (X-\lambda)\frac{1}{2}\frac{1}{\sqrt{\lambda}} + \frac{X^2+\lambda^2-2X\lambda}{2!}\left(-\frac{1}{4}\frac{1}{\lambda^{3/2}}\right)$$

so $E(Y) = E(g(x))$

$$= E\left[\sqrt{\lambda} + \frac{X}{2\sqrt{\lambda}} - \frac{\sqrt{\lambda}}{2} - \frac{(X^2+\lambda^2-2X\lambda)}{8\lambda^{3/2}}\right]$$

$$= E\left[\sqrt{\lambda} + \frac{X}{2\sqrt{\lambda}} - \frac{\sqrt{\lambda}}{2} - \frac{X^2}{8\lambda^{3/2}} - \frac{\lambda^2}{8\lambda^{3/2}} + \frac{2X\lambda}{8\lambda^{3/2}}\right]$$

$$= E\left[\sqrt{\lambda} + \frac{1}{2\sqrt{\lambda}}X - \frac{\sqrt{\lambda}}{2} - \frac{X^2}{8\sqrt{\lambda^3}} - \frac{\sqrt{\lambda}}{8} + \frac{1}{4\sqrt{\lambda}}X\right]$$

so $E(Y) = \sqrt{\lambda} + \frac{1}{2\sqrt{\lambda}}E(X) - \frac{\sqrt{\lambda}}{2} - \frac{1}{8\sqrt{\lambda^3}}E(X^2) - \frac{\sqrt{\lambda}}{8} + \frac{1}{4\sqrt{\lambda}}E(X)$.

But $\boxed{E(X) = \lambda,\ E(X^2) = \lambda^2+\lambda}$

so $E(Y) = \sqrt{\lambda} + \frac{1}{2\sqrt{\lambda}}\lambda - \frac{\sqrt{\lambda}}{2} - \frac{1}{8\sqrt{\lambda^3}}(\lambda^2+\lambda) - \frac{\sqrt{\lambda}}{8} + \frac{1}{4\sqrt{\lambda}}\lambda$

$$= \sqrt{\lambda} + \frac{\sqrt{\lambda}}{2} - \frac{\sqrt{\lambda}}{2} - \frac{\lambda^2}{8\lambda^{3/2}} - \frac{\lambda}{8\lambda^{3/2}} - \frac{\sqrt{\lambda}}{8} - \frac{\sqrt{\lambda}}{4}$$

$$= \sqrt{\lambda} - \frac{\sqrt{\lambda}}{8} - \frac{1}{8\sqrt{\lambda}} - \frac{\sqrt{\lambda}}{8} - \frac{\sqrt{\lambda}}{4} = \sqrt{\lambda} - \frac{1}{4}\sqrt{\lambda} - \frac{1}{4}\sqrt{\lambda} - \frac{1}{8\sqrt{\lambda}} = \sqrt{\lambda} - \frac{1}{2}\sqrt{\lambda} - \frac{1}{8\sqrt{\lambda}}$$

$$= \frac{8\lambda - 4\lambda - 1}{8\sqrt{\lambda}} = \boxed{\frac{4\lambda-1}{8\sqrt{\lambda}}} = \boxed{E(\sqrt{X})} = \boxed{E(Y)}$$

so $Var(Y) = E(Y^2) - [E(Y)]^2 = \lambda - \left(\frac{4\lambda-1}{8\sqrt{\lambda}}\right)^2 = \boxed{\frac{48\lambda^2+8\lambda-1}{64\lambda}}$

approximate $Var(Y)$ using ~~taylor~~ approximation

Just use the result. No need to derive formulas that we have derived before!

$$
\begin{aligned}
M_Y''''(t) &= \frac{d^4}{dt^4}\left(\frac{\lambda}{\lambda-t}\right)^{\alpha} \\
&= \frac{\alpha\left(6+11\alpha+6\alpha^2+\alpha^3\right)}{(t-\lambda)^4}\left(\frac{\lambda}{\lambda-t}\right)^{\alpha}
\end{aligned}
$$

Then

$$
M_Y''''(0) = \frac{\alpha\left(6+11\alpha+6\alpha^2+\alpha^3\right)}{\lambda^4}
$$

Therefore

$$
E\left(X^2\right) = \frac{\alpha\left(6+11\alpha+6\alpha^2+\alpha^3\right)}{\lambda^4} + \frac{\alpha(\alpha+1)}{\lambda^2}
$$

Then (1) becomes

$$
\begin{aligned}
Var(X) &= \left(\frac{\alpha\left(6+11\alpha+6\alpha^2+\alpha^3\right)}{\lambda^4} + \frac{\alpha(\alpha+1)}{\lambda^2}\right) - \left(\frac{\alpha(\alpha+1)}{\lambda^2}\right)^2 \\
&= \frac{\alpha\left(6+11\alpha+6\alpha^2+\alpha^3\right)}{\lambda^4} + \frac{\alpha(\alpha+1)}{\lambda^2} - \frac{\alpha^2(\alpha+1)^2}{\lambda^4}
\end{aligned}
$$

Then

$$
\boxed{Var(X) = \tfrac{\alpha}{\lambda^4}(\alpha+1)\left(\lambda^2+4\alpha+6\right)}
$$

$$\begin{aligned} E(Y) &= M_Y'(0) \\ &= E(\sqrt{x}) \end{aligned}$$

So we need to find $E(\sqrt{x})$ to complete the solution.

$$\begin{aligned} E(\sqrt{x}) &= \sum_{x=0}^{\infty} \sqrt{x}\frac{\lambda^x e^{-\lambda}}{x!} \\ &= 0 + \frac{\lambda e^{-\lambda}}{1!} + \sqrt{2}\frac{\lambda^2 e^{-\lambda}}{2!} + \sqrt{3}\frac{\lambda^3 e^{-\lambda}}{3!} + \sqrt{4}\frac{\lambda^4 e^{-\lambda}}{4!} + \cdots \\ &= e^{-\lambda}\left(\lambda + \sqrt{2}\frac{\lambda^2}{2!} + \sqrt{3}\frac{\lambda^3}{3!} + \sqrt{4}\frac{\lambda^4}{4!} + \cdots\right) \\ &= \lambda e^{-\lambda}\left(1 + \frac{1}{\sqrt{2}}\lambda + \frac{1}{\sqrt{3}}\frac{\lambda^2}{2!} + \frac{1}{\sqrt{4}}\frac{\lambda^3}{3!} + \frac{1}{\sqrt{5}}\frac{\lambda^4}{4!} + \cdots\right) \\ &= \lambda e^{-\lambda}\left(1 + \frac{\lambda}{\sqrt{2}} + \frac{2}{\sqrt{3}}\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^2}{2!} + \frac{2\sqrt{2}}{\sqrt{4}}\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^3}{3!} + \frac{4}{\sqrt{5}}\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^4}{4!} + \cdots\right) \\ &= \lambda e^{-\lambda}\left(1 + \frac{\lambda}{\sqrt{2}} + 1.1547\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^2}{2!} + 1.4142\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^3}{3!} + 1.7889\frac{\left(\frac{\lambda}{\sqrt{2}}\right)^4}{4!} + \cdots\right) \\ &\simeq \lambda e^{-\lambda}\left(e^{\frac{\lambda}{\sqrt{2}}}\right) \\ &= \lambda e^{\frac{\lambda}{\sqrt{2}}-\lambda} = \lambda e^{\frac{\lambda(1-\sqrt{2})}{\sqrt{2}}} \end{aligned}$$

Hence

$$\begin{aligned} Var(Y) &= \lambda - \left[\lambda e^{\frac{\lambda(1-\sqrt{2})}{\sqrt{2}}}\right]^2 \\ &= \lambda - \lambda^2 e^{\frac{2\lambda(1-\sqrt{2})}{\sqrt{2}}} \\ &= \lambda - \lambda^2 e^{\sqrt{2}\lambda(1-\sqrt{2})} \end{aligned}$$

Hence

$$\boxed{Var(Y) \simeq \lambda\left(1 - \lambda e^{-0.58578\lambda}\right)}$$

3. The random variable Y has a Gamma distribution with parameters α and λ. Furthermore, assume that X given Y has a Poisson distribution with parameter Y^2. (a) Obtain $E(X)$. (b) Obtain $Var(X)$.

(a)

$$f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} \qquad y \geq 0$$

$$f_{(X|Y=y)}(x|y) = \frac{(y^2)^x e^{-y^2}}{y!} \qquad x = 0, 1, 2, \cdots$$

Now

$$E(X) = E(E(X|Y))$$

But $E(X|Y)$ is expectation of a Poisson r.v. with parameter Y^2. But we know that mean of a poisson r.v. with parameter λ is λ. Hence $E(X|Y) = Y^2$ since we are told Y^2 is the parameter.
Hence

$$E(X) = E\left(Y^2\right)$$

But the moment generating function for Gamma is $M_Y(t) = \left(\frac{\lambda}{\lambda-t}\right)^\alpha$ (book page 145 second edition).
Hence $E(Y^2) = M_Y''(0) = \frac{\alpha(\alpha+1)}{\lambda^2}$ (page 145)
Hence

$$\boxed{E(X) = \frac{\alpha(\alpha+1)}{\lambda^2}}$$

(b)

$$Var(X) = E\left(X^2\right) - [E(X)]^2 \tag{1}$$

But

$$E\left(X^2\right) = E\left(E\left(X^2|Y\right)\right)$$

But $E(X^2|Y)$ is $E(X^2)$ of a poisson r.v. with parameter Y^2. But we know that $E(X^2)$ of a poisson r.v. with parameter λ is $\lambda^2 + \lambda$ (book page 144 example A). Hence since we are told Y^2 is the parameter, then

$$\begin{aligned} E\left(X^2|Y\right) &= \left(Y^2\right)^2 + Y^2 \\ &= \left(Y^4 + Y^2\right) \end{aligned}$$

Hence

$$\begin{aligned} E\left(X^2\right) &= E\left(Y^4 + Y^2\right) \\ &= E\left(Y^4\right) + E\left(Y^2\right) \end{aligned}$$

But using mgf for Gamma distribution we can find $E(Y^4)$.

4.6 Quiz 6

Local contents

QUIZ 6 MATH 502AB October 30, 2006

Name (please print) NASSER ABBASI

In solving the problems below, you can use all the results that we have derived in class. You do not need to re-derive results. Make sure to cite the results that you use.

1. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(\mu, \sigma^2)$, and S^2 be the sample variance. What is $\mathrm{Var}(S^2)$?

Figure 4.44: Problem 1

By theorem B, chapter 6, "Mathematical Statistics and Data Analysis", 2nd edition, John Rice, page 181, which states that the distribution of $\frac{(n-1)S^2}{\sigma^2}$ is a chi-square distribution with $n-1$ degrees of freedom.

Hence

$$Var\left(\frac{(n-1)\,S^2}{\sigma^2}\right) = Var(t_{n-1})$$

Since $\frac{n-1}{\sigma^2}$ is not random, then applying the property that $Var(c\,X) = c^2 Var(X)$ when c is not random to the above, where in this case $c = \frac{(n-1)}{\sigma^2}$ and rearranging, we obtain

$$Var\left(S^2\right) = \frac{\sigma^4}{(n-1)^2} Var(t_{n-1})$$

However, $Var(t_{n-1}) = 2(n-1)$ [1], hence

$$Var\left(S^2\right) = 2\frac{\sigma^4}{n-1}$$

2. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(0,1)$. Determine the asymptotic distribution of

$$(1/n)\sum_{i=1}^{n} |X_i|.$$

Figure 4.45: Problem 2

Let

$$S_n = \frac{1}{n}\sum_{i=1}^{n} |X_i|$$

[1] I found $Var(t_{n-1})$ from Chi-square moment generation function. Since $M(t_n) = (1-2t)^{\frac{-n}{2}}$, then $M(t_{n-1}) = (1-2t)^{\frac{-n-1}{2}}$ and then $Var(t_{n-1}) = E(t^2) - E(t)^2 = M''(0) - [M'(0)]^2$ which comes out to $2(n-1)$

Find moment generation function

$$
\begin{aligned}
M_{S_n}(t) &= E\left(e^{\frac{t}{n}\sum_{i=1}^{n}|X_i|}\right)\\
&= E\left(\Pi e^{\frac{t}{n}|X_i|}\right)\\
&= \Pi E\left(e^{\frac{t}{n}|X_i|}\right)\\
&= \left[M_{|X|}\left(\frac{t}{n}\right)\right]^n
\end{aligned}
$$

To find $M_{|X|}\left(\frac{t}{n}\right)$, and noting that $\mu = 0$ and $\sigma = 1$ we obtain[2]

$$
M_{|X|}\left(\frac{t}{n}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{t}{n}|x|} e^{\frac{-x^2}{2}}\,dx
$$

Due to symmetry of normal distribution and since $|x|$ is positive always the above can be written as[3]

$$
\begin{aligned}
M_{|X|}\left(\frac{t}{n}\right) &= \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} e^{\frac{tx}{n}-\frac{x^2}{2}}\,dx\\
&= e^{\frac{t^2}{2n^2}}\left(1+\operatorname{erf}\left(\frac{t}{\sqrt{2}n}\right)\right)
\end{aligned}
$$

Hence

$$
M_{S_n}(t) = \left[e^{\frac{t^2}{2n^2}}\left(1+\operatorname{erf}\left(\frac{t}{\sqrt{2}n}\right)\right)\right]^n
$$

The limit of the above as $n \to \infty$ is $e^{\sqrt{\frac{2}{\pi}}t}$. Therefore

$$
M_{S_n}(t) = e^{\sqrt{\frac{2}{\pi}}t}
$$

We see now that $E(S_n) = M'(0) = \sqrt{\frac{2}{\pi}}$ and $E(S_n^2) = M''(0) = \frac{2}{\pi}$, therefore $Var(S_n) = \frac{2}{\pi} - \left(\sqrt{\frac{2}{\pi}}\right)^2 = 0$. (this means all sums add to same value for large n, did I make a mistake? I did not expect this). Hence

$$
S_n \overset{\text{in distribution}}{\rightarrow} N\left(\sqrt{\frac{2}{\pi}}, 0\right)
$$

3.. Let $X_1,\cdots,X_n$ be iid random variables from a $\mathcal{N}(0,\sigma_1^2)$ and $Y_1,\cdots,Y_n$ be iid random variables from a $\mathcal{N}(0,\sigma_2^2)$. Write a 95% confidence interval for σ_1^2/σ_2^2.

Figure 4.46: Problem 3

For pivotal term use $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$, where s^2 is sample variance σ^2 is population variance, and hence we write (following class notes on 10/29/07) the confidence interval as

[2] I started by write $\log\left(M_S\left(\frac{t}{n}\right)\right) = n\log\left(M_{|X|}\left(\frac{t}{n}\right)\right)$ and then expanding $\log\left(M_{|X|}\left(\frac{t}{n}\right)\right)$ around $t = 0$ using taylor series. But due to the absolute x present, I was not sure I was doing it correctly so I changed to using the integral approach.

[3] Please note that I used Mathematica for solving this integral and the limit. I need to learn better how to do this by hand using the Log expansion?

$$P[-z_p < \theta < z_p] = 1 - \alpha$$

Where from table A7, $z_p = 1.96$ for normal r.v. at 95% and Where $\theta = \frac{\frac{(n-1)s_2^2}{\sigma_2^2}}{\frac{(n-1)s_1^2}{\sigma_1^2}}$

Hence the C.I. becomes

$$P\left[-1.96 < \frac{\sigma_1^2 s_2^2}{\sigma_2^2 s_1^2} < 1.96\right] = 1 - \alpha$$
$$P\left[-1.96\frac{s_1^2}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < 1.96\frac{s_1^2}{s_2^2}\right] = 1 - \alpha$$

Where the sample variance $s_2^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}\right)^2$, and $s_1^2 = \frac{1}{n-1}\sum_{i=1}^n \left(Y_i - \bar{Y}\right)^2$

For 95% confidence, $\alpha = 0.05$. Hence the the final answer for the C.I. is

$$P\left[-1.96\frac{s_1^2}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < 1.96\frac{s_1^2}{s_2^2}\right] = 0.95$$

Not sure what more I can do with the above so I think I will stop here.

4. Let $X \sim N(0,2)$ and $Y \sim exponential(1)$. Provided that X is independent of Y, identify the distribution of $X/\sqrt{Y}$.

Figure 4.47: Problem 4

First find the joint density of X,Y. Since X,Y are independent, then the joint density $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$ over $-\infty < x < \infty$ and $y > 0$

But $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ and $f_Y(y) = \lambda e^{-\lambda y}$, hence the joint density is (after substituting for $\mu = 0, \sigma^2 = 2, \lambda = 1$ is

$$\boxed{f_{X,Y}(x,y) = \frac{1}{2\sqrt{\pi}}e^{\frac{-x^2}{4}}\,e^{-y} \qquad -\infty < x < \infty,\ y > 0}$$

Now Let $Z = \frac{X}{\sqrt{Y}}$, and let $U = Y$

Hence

$$f_{Z,U}(z,u) = |J|\,f_{X,Y}(z,u) \tag{1}$$

Where

$$J = \det\begin{bmatrix}\frac{\partial Z}{\partial X} & \frac{\partial Z}{\partial Y}\\ \frac{\partial U}{\partial X} & \frac{\partial U}{\partial Y}\end{bmatrix} = \det\begin{bmatrix}\frac{1}{\sqrt{Y}} & \frac{-X}{2\sqrt{Y}}\\ 0 & 1\end{bmatrix}$$
$$= \sqrt{\frac{1}{Y}}$$

so

$$\left|J^{-1}\right| = \sqrt{Y} = \sqrt{U}$$

Hence, from (1) and substitute $X = Z\sqrt{U}$ and $Y = U$, we obtain

$$f_{Z,U}(z,u) = \sqrt{u}\frac{1}{2\sqrt{\pi}}e^{\frac{-z^2 u}{4}}e^{-u}$$

Hence the marginal density

$$\boxed{f_Z(z) = \int_0^\infty f_{Z,U}(z,u)\ du}$$

Then

$$\begin{aligned} f_Z(z) &= \frac{1}{2\sqrt{\pi}}\int_0^\infty u^{\frac{1}{2}}e^{\frac{-z^2u-4u}{4}}\ du \\ &= \frac{1}{2\sqrt{\pi}}\int_0^\infty u^{\frac{1}{2}}e^{-\left(1+\frac{z^2}{4}\right)u}\ du \end{aligned}$$

Now Gamma distribution is $f(w) = \frac{\lambda^\alpha}{\Gamma(\alpha)}w^{\alpha-1}e^{-\lambda w}$, hence if we replace $\lambda = 1+\frac{z^2}{4}$ and $\alpha = \frac{3}{2}$, then we have

$$\begin{aligned} f_Z(z) &= \frac{1}{2\sqrt{\pi}}\frac{\Gamma(\alpha)}{\lambda^\alpha}\overbrace{\int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)}w^{\alpha-1}e^{-\lambda w}\ dw}^{=1} \\ &= \frac{1}{2\sqrt{\pi}}\frac{\Gamma(\alpha)}{\lambda^\alpha} \end{aligned}$$

To simplify further,

$$f_Z(z) = \frac{1}{2\sqrt{\pi}}\frac{\Gamma\left(\frac{3}{2}\right)}{\left(1+\frac{z^2}{4}\right)^{\frac{3}{2}}}$$

But $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$, hence

$$\boxed{f_Z(z) = \frac{1}{4}\left(1+\frac{z^2}{4}\right)^{-\frac{3}{2}}}$$

Hence the pdf of $\frac{X}{\sqrt{Y}}$ is

$$f_{X,Y}(x,y) = \frac{1}{4}\left(1+\frac{x^2}{4y}\right)^{-\frac{3}{2}}$$

To verify this is a pdf, I integrate it from $-\infty$ to $+\infty$ to see if I get 1:

```
In[16]:= Integrate[1/4 1/(1 + z^2/4)^(3/2), {z, -Infinity, Infinity}]

Out[16]= 1
```

Figure 4.48: verify

Here is a plot of the distribution

```
f = 1/4 (1 + x^2/(4 y))^(-3/2)
Plot3D[f, {x, -20, 20}, {y, 0, 20}, AxesLabel → {"X", "Y", "f(x,y)"}]
```

$$\frac{1}{4\left(1+\frac{x^2}{4\,y}\right)^{3/2}}$$

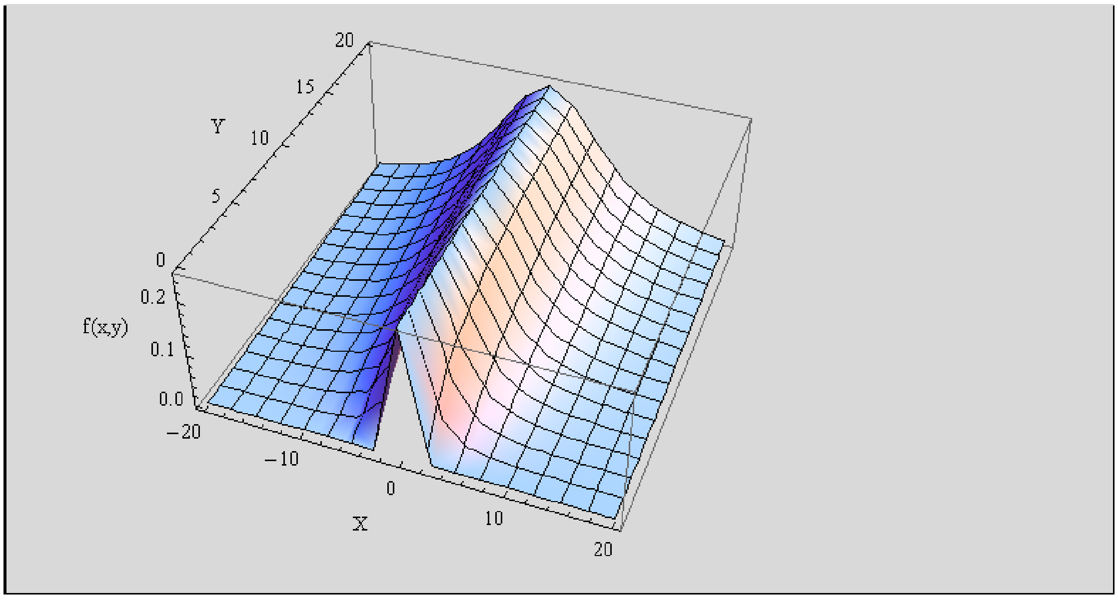


Figure 4.49: plot of the distribution

Another attempt at problem (2)

$$\begin{aligned}\bar{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i\\ \left|\bar{X}_n\right| &= \left|\frac{1}{n}\sum_{i=1}^{n} X_i\right|\\ &\leq \frac{1}{n}\sum_{i=1}^{n} |X_i| \end{aligned} \tag{1}$$

By definition, the CDF of $\left|\bar{X}_n\right|$ is

$$\begin{aligned}F_{|\bar{X}_n|}\left(\left|\bar{X}_n\right| = c\right) &= P\left(\left|\bar{X}_n\right| < c\right)\\ &= P\left(-c < \bar{X}_n < c\right)\\ &= P\left(\frac{-c-\mu}{\sigma/\sqrt{n}} < \frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}} < \frac{c-\mu}{\sigma/\sqrt{n}}\right)\end{aligned}$$

Since $\mu = 0, \sigma = 1$ we obtain

$$F_{|\bar{X}_n|}\left(\left|\bar{X}_n\right| = c\right) = P\left(\frac{-c}{1/\sqrt{n}} < \frac{\bar{X}_n}{1/\sqrt{n}} < \frac{c}{1/\sqrt{n}}\right) \tag{2}$$

Now I need to combine (1) and (2). I am not sure how.

But central limit theorem tells us that as n gets large, the distribution of the sample mean $\bar{X}_n$ approach normal distribution with mean μ and variance $\frac{\sigma^2}{n}$, hence $\bar{X}_n \overset{\text{in distribution}}{\rightarrow} N\left(\mu, \frac{\sigma^2}{n}\right)$, hence the above becomes

$$F_{|\bar{X}_n|}\left(\left|\bar{X}_n\right| = c\right) \simeq \Phi\left(c\sqrt{n}\right) - \Phi\left(-c\sqrt{n}\right)$$

4.6.1 Graded

15/20

QUIZ 6 MATH 502AB 15/20 October 30, 2006

Name (please print) NASSER ABBASI

In solving the problems below, you can use all the results that we have derived in class. You do not need to re-derive results. Make sure to cite the results that you use.

1. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(\mu, \sigma^2)$, and S^2 be the sample variance. What is $\mathrm{Var}(S^2)$?

By theorem B, chapter 6, "Mathematical Statistics and Data Analysis", 2nd edition, John Rice, page 181, which states that the distribution of $\frac{(n-1)S^2}{\sigma^2}$ is a chi-square distribution with $n-1$ degrees of freedom.

Hence

$$Var\left(\frac{(n-1)\,S^2}{\sigma^2}\right) = Var\left(t_{n-1}\right)$$

Since $\frac{n-1}{\sigma^2}$ is not random, then applying the property that $Var\left(c\,X\right) = c^2 Var\left(X\right)$ when c is not random to the above, where in this case $c = \frac{(n-1)}{\sigma^2}$ and rearranging, we obtain

$$Var\left(S^2\right) = \frac{\sigma^4}{\left(n-1\right)^2} Var\left(t_{n-1}\right)$$

However, $Var\left(t_{n-1}\right) = 2(n-1)$[1], hence

$$\boxed{Var\left(S^2\right) = 2\frac{\sigma^4}{n-1}}$$

[1] I found $Var\left(t_{n-1}\right)$ from Chi-square moment generation function. Since $M\left(t_n\right) = (1-2t)^{\frac{-n}{2}}$, then $M\left(t_{n-1}\right) = (1-2t)^{\frac{-n-1}{2}}$ and then $Var\left(t_{n-1}\right) = E\left(t^2\right) - E\left(t\right)^2 = M''\left(0\right) - \left[M'\left(0\right)\right]^2$ which comes out to $2\left(n-1\right)$

1

2. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(0,1)$. Determine the asymptotic distribution of

$$(1/n)\sum_{i=1}^{n}|X_i|.$$

Use CLT theorem

Let

$$S_n = \frac{1}{n}\sum_{i=1}^{n}|X_i|$$

Working too hard again!!

Find moment generation function

$$\begin{aligned} M_{S_n}(t) &= E(e^{\frac{t}{n}\sum_{i=1}^{n}|X_i|}) \\ &= E\left(\Pi e^{\frac{t}{n}|X_i|}\right) \\ &= \Pi E\left(e^{\frac{t}{n}|X_i|}\right) \\ &= \left[M_{|X|}\left(\frac{t}{n}\right)\right]^n \end{aligned}$$

To find $M_{|X|}\left(\frac{t}{n}\right)$, and noting that $\mu = 0$ and $\sigma = 1$ we obtain[2]

$$M_{|X|}\left(\frac{t}{n}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{t}{n}|x|}e^{\frac{-x^2}{2}}dx$$

Due to symmetry of normal distribution and since $|x|$ is positive always the above can be written as[3]

3/

$$\begin{aligned} M_{|X|}\left(\frac{t}{n}\right) &= \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} e^{\frac{tx}{n}-\frac{x^2}{2}}dx \\ &= e^{\frac{t^2}{2n^2}}\left(1+\operatorname{erf}\left(\frac{t}{\sqrt{2}n}\right)\right) \end{aligned}$$

$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$

Hence

$$M_{S_n}(t) = \left[e^{\frac{t^2}{2n^2}}\left(1+\operatorname{erf}\left(\frac{t}{\sqrt{2}n}\right)\right)\right]^n$$

need to show!

The limit of the above as $n\to\infty$ is $e^{\sqrt{\frac{2}{\pi}}t}$. Therefore

$$\boxed{M_{S_n}(t) = e^{\sqrt{\frac{2}{\pi}}t}}$$

[2] I started by write $\log\left(M_S\left(\frac{t}{n}\right)\right) = n\log\left(M_{|X|}\left(\frac{t}{n}\right)\right)$ and then expanding $\log\left(M_{|X|}\left(\frac{t}{n}\right)\right)$ around $t=0$ using taylor series. But due to the absolute x present, I was not sure I was doing it correctly so I changed to using the integral approach.

[3] Please note that I used Mathematica for solving this integral and the limit. I need to learn better how to do this by hand using the Log expansion?

?

We see now that $E(S_n) = M'(0) = \sqrt{\frac{2}{\pi}}$ and $E(S_n^2) = M''(0) = \frac{2}{\pi}$, therefore $Var(S_n) = \frac{2}{\pi} - \left(\sqrt{\frac{2}{\pi}}\right)^2 = 0$. (this means all sums add to same value for large n, did I make a mistake? I did not expect this). Hence

$$\boxed{S_n \overset{\text{in distribution}}{\longrightarrow} N\left(\sqrt{\frac{2}{\pi}}, 0\right)}$$

3.. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(0, \sigma_1^2)$ and $Y_1, \cdots, Y_n$ be iid random variables from a $\mathcal{N}(0, \sigma_2^2)$. Write a 95% confidence interval for σ_1^2/σ_2^2.

For pivotal term use $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$, where s^2 is sample variance σ^2 is population variance, and hence we write (following class notes on 10/29/07) the confidence interval as

$$P\left[-z_p < \theta < z_p\right] = 1 - \alpha$$

Where from table A7, $z_p = 1.96$ for normal r.v. at 95% and Where $\theta = \frac{\frac{(n-1)s_2^2}{\sigma_2^2}}{\frac{(n-1)s_1^2}{\sigma_1^2}}$

Hence the C.I. becomes

$$P\left[-1.96 < \frac{\sigma_1^2 s_2^2}{\sigma_2^2 s_1^2} < 1.96\right] = 1 - \alpha$$

$$P\left[-1.96\frac{s_1^2}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < 1.96\frac{s_1^2}{s_2^2}\right] = 1 - \alpha$$

Where the sample variance $s_2^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}\right)^2$, and $s_1^2 = \frac{1}{n-1}\sum_{i=1}^n \left(Y_i - \bar{Y}\right)^2$
For 95% confidence, $\alpha = 0.05$. Hence the the final answer for the C.I. is

$$\boxed{P\left[-1.96\frac{s_1^2}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < 1.96\frac{s_1^2}{s_2^2}\right] = 0.95}$$

Not sure what more I can do with the above so I think I will stop here.

4. Let $X \sim N(0,2)$ and $Y \sim exponential(1)$. Provided that X is independent of Y, identify the distribution of $X/\sqrt{Y}$.

First find the joint density of X,Y. Since X,Y are independent, then the joint density $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ over $-\infty < x < \infty$ and $y > 0$

But $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$ and $f_Y(y) = \lambda e^{-\lambda y}$, hence the joint density is (after substituting for $\mu = 0, \sigma^2 = 2, \lambda = 1$ is

$$\boxed{f_{X,Y}(x,y) = \frac{1}{2\sqrt{\pi}} e^{\frac{-x^2}{4}} e^{-y} \qquad -\infty < x < \infty,\ y > 0}$$

Now Let $Z = \frac{X}{\sqrt{Y}}$, and let $U = Y$

Hence

$$f_{Z,U}(z,u) = |J|\, f_{X,Y}(z,u) \tag{1}$$

Where

$$J = \det \begin{bmatrix} \frac{\partial Z}{\partial X} & \frac{\partial Z}{\partial Y} \\ \frac{\partial U}{\partial X} & \frac{\partial U}{\partial Y} \end{bmatrix} = \det \begin{bmatrix} \frac{1}{\sqrt{Y}} & \frac{-X}{2\sqrt{Y}} \\ 0 & 1 \end{bmatrix}$$
$$= \sqrt{\frac{1}{Y}}$$

so

$$\left|J^{-1}\right| = \sqrt{Y} = \sqrt{U}$$

Hence, from (1) and substitute $X = Z\sqrt{U}$ and $Y = U$, we obtain

$$f_{Z,U}(z,u) = \sqrt{u} \frac{1}{2\sqrt{\pi}} e^{\frac{-z^2 u}{4}} e^{-u}$$

Hence the marginal density

Working for hand

$$\boxed{f_Z(z) = \int_0^\infty f_{Z,U}(z,u)\ du}$$

Then

$$f_Z(z) = \frac{1}{2\sqrt{\pi}} \int_0^\infty u^{\frac{1}{2}} e^{\frac{-z^2 u - 4u}{4}}\ du$$
$$= \frac{1}{2\sqrt{\pi}} \int_0^\infty u^{\frac{1}{2}} e^{-\left(1+\frac{z^2}{4}\right)u}\ du$$

Now Gamma distribution is $f(w) = \frac{\lambda^\alpha}{\Gamma(\alpha)} w^{\alpha-1} e^{-\lambda w}$, hence if we replace $\lambda = 1 + \frac{z^2}{4}$ and $\alpha = \frac{3}{2}$, then we have

$$f_Z(z) = \frac{1}{2\sqrt{\pi}} \frac{\Gamma(\alpha)}{\lambda^\alpha} \overbrace{\int_0^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} w^{\alpha-1} e^{-\lambda w}\ dw}^{=1}$$
$$= \frac{1}{2\sqrt{\pi}} \frac{\Gamma(\alpha)}{\lambda^\alpha}$$

To simplify further,

$$f_Z(z) = \frac{1}{2\sqrt{\pi}} \frac{\Gamma\left(\frac{3}{2}\right)}{\left(1+\frac{z^2}{4}\right)^{\frac{3}{2}}}$$

But $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$, hence

$$\boxed{f_Z(z) = \frac{1}{4}\left(1+\frac{z^2}{4}\right)^{-\frac{3}{2}}}$$

Hence the pdf of $\frac{X}{\sqrt{Y}}$is

$$f_{X,Y}(x,y) = \frac{1}{4}\left(1+\frac{x^2}{4y}\right)^{-\frac{3}{2}}$$

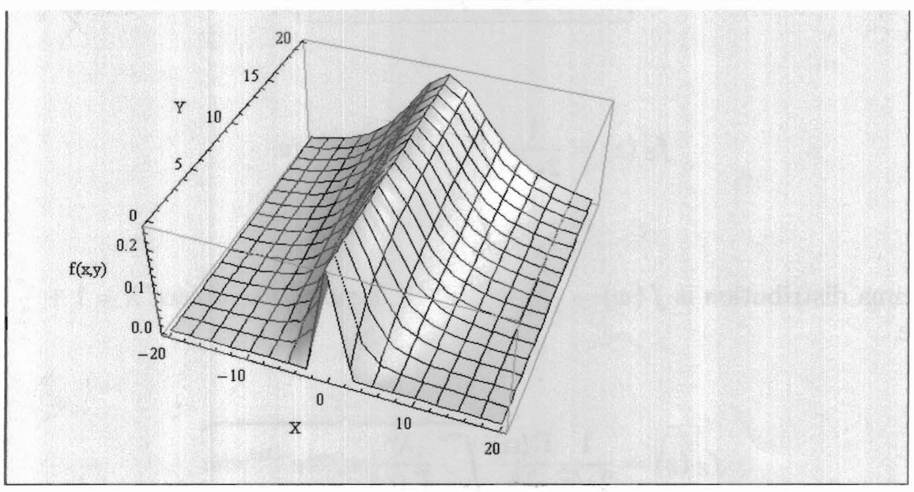
To verify this is a pdf, I integrate it from $-\infty$ to $+\infty$ to see if I get 1:

```
In[16]:= Integrate[1/4 1/(1 + z^2/4)^(3/2), {z, -Infinity, Infinity}]
Out[16]= 1
```

Here is a plot of the distribution

```
f = 1/4 (1 + x^2/(4 y))^(-3/2)
Plot3D[f, {x, -20, 20}, {y, 0, 20}, AxesLabel -> {"X", "Y", "f(x,y)"}]
```

$$\frac{1}{4\left(1+\frac{x^2}{4y}\right)^{3/2}}$$



6

4.7 Quiz 7

Local contents

QUIZ 7 MATH 502AB November 27, 2007

Name (please print) NASSER ABBASI

1. Consider $X_1, \cdots, X_n$, an i.i.d. sample from a random variables with density

$$f(x|\sigma) = \frac{1}{2\sigma}\exp\left(-\frac{|x|}{\sigma}\right).$$

(a) Find the method of moment estimate of σ.

(b) Find MLE of σ.

(c) Find the asymptotic distribution of the MLE.

Figure 4.50: Problem 1

PART(A)

In the method of moments we formulate the moments of the probability law of the distribution in which the random variables belong to and equate these moments to the moments obtained from the sample at hand and solve for the unknown parameters.

$$\begin{aligned}
\mu_1 &= E(X) \\
&= \int_{-\infty}^{\infty} x f(x)\,dx \\
&= \frac{1}{2\sigma}\int_{-\infty}^{\infty} x\ e^{-\frac{|x|}{\sigma}}\,dx
\end{aligned}$$

But $e^{-\frac{|x|}{\sigma}}$ is symmtric around the $x = 0$ due to absolute x in the power of the exp.(This assumes σ positive, which is ofcourse true) but it is multiplied by negative x to the left of y-axis and multiplied by positive x on the right of the y-axis, hence the area of the left of the y-axis will be equal but negative to the area on the right of the y-axis. Hence the above integral is zero. Hence $\mu_1 = 0$

This moment provides no information. Find the second moment.

$$\begin{aligned}
\mu_2 &= E(X^2) \\
&= \frac{1}{2\sigma}\int_{-\infty}^{\infty} x^2 e^{-\frac{|x|}{\sigma}}\,dx
\end{aligned}$$

Due to the symmetry of $e^{-\frac{|x|}{\sigma}}$ and also x^2 is even and symmetrical around $x = 0$, the above integral is then twice the integral from $x = 0 \cdots \infty$ and it becomes

$$\mu_2 = \frac{1}{\sigma}\int_0^{\infty} x^2 e^{-\frac{x}{\sigma}}\,dx$$

Integration by parts gives

$$\mu_2 = 2\sigma^2$$

Hence

$$\sigma = \sqrt{\frac{\mu_2}{2}} \tag{1}$$

Now find μ_2 from the sample itself and substitute for it in the above. From the sample,

$$\begin{aligned}\mu_2 &= Var(sample) + Mean(Sample)\\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 + \bar{X}\end{aligned}$$

Since the mean of the population was found to be zero, we can take the mean of the sample $\bar{X} = 0$

Hence μ_2 from the sample becomes

$$\mu_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Replace the above in (1) we obtain estimate of the population σ as

$$\boxed{\hat{\sigma} = \sqrt{\tfrac{1}{2n}\sum_{i=1}^{n} X_i^2}}$$

PART(B)

The MLE of σ is found as follows. Since i.i.d. random variables we write (Where $L(\sigma)$ mean $lik(\sigma)$ and $l(\sigma)$ means $\log(lik(\sigma))$

$$\begin{aligned}L_n(\sigma) &= \prod_{i=1}^{n} f(X_i|\sigma)\\ l_n(\sigma) &= \sum_{i=1}^{n} \log f(X_i|\sigma)\\ &= \sum_{i=1}^{n} \log\left(\frac{1}{2\sigma}e^{-\frac{|X_i|}{\sigma}}\right)\\ &= \sum_{i=1}^{n}\left(-\log(2\sigma) - \frac{|X_i|}{\sigma}\right)\end{aligned}$$

Therefore

$$\boxed{l_n(\sigma) = -n\log(2\sigma) - \tfrac{1}{\sigma}\sum_{i=1}^{n}|X_i|}$$

Now we find the MLE, which is the value of σ which maximizes the above function.

$$\begin{aligned}
l_n'(\sigma) &= 0 \\
\frac{-n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|X_i| &= 0 \\
-\sigma n + \sum_{i=1}^{n}|X_i| &= 0
\end{aligned}$$

Hence

$$\boxed{\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|X_i|}$$

The above is the MLE estimate of the parameter σ.

PART(C)

The asymptotic distribution of the MLE $\hat{\sigma}$ is normal with mean σ and variance $\frac{1}{nI(\sigma)}$ where

$$I(\sigma) = -E[l_n''(\sigma)]$$

But

$$\begin{aligned}
l_n'(\sigma) &= \frac{\partial}{\partial\sigma}\left[-n\log(2\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n}|X_i|\right] \\
&= \frac{-n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|X_i|
\end{aligned}$$

and

$$l_n''(\sigma) = \frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{n}|X_i|$$

Hence

$$\begin{aligned}
I(\sigma) &= -E\left[\frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{n}|X_i|\right] \\
&= -E\left[\frac{n}{\sigma^2}\right] + \frac{2}{\sigma^3}E\left[\sum_{i=1}^{n}|X_i|\right] \\
&= -\frac{n}{\sigma^2} + \frac{2}{\sigma^3}\sum_{i=1}^{n}E|X_i| \qquad (2)
\end{aligned}$$

Need to find $E|X_i|$, since i.i.d. all random variables has the same expected value as X, hence

$$\begin{aligned}
E|X| &= \frac{1}{2\sigma}\int_{-\infty}^{\infty}|x|\,e^{-\frac{|x|}{\sigma}}\,dx \\
&= \frac{1}{\sigma}\int_{0}^{\infty}xe^{-\frac{x}{\sigma}}\,dx \\
&= \sigma
\end{aligned}$$

Therefore from (2)

$$
\begin{aligned}
I(\sigma) &= -\frac{n}{\sigma^2} + \frac{2}{\sigma^3}\sum_{i=1}^{n}\sigma \\
&= -\frac{n}{\sigma^2} + \frac{2n}{\sigma^2}
\end{aligned}
$$

Hence the Fischer information matrix is

$$I(\sigma) = \frac{n}{\sigma^2}$$

Hence MLE $\hat{\sigma}_n$ has an asymptotic distribution $\sim N\left(\sigma, \frac{1}{nI(\sigma)}\right)$

i.e.

$$E(\hat{\sigma}_n) = \sigma$$

and

$$
\begin{aligned}
Var(\hat{\sigma}_n) &= \frac{1}{nI(\sigma)} \\
&= \frac{\sigma^2}{n^2}
\end{aligned}
$$

2. Suppose that certain electronic components have lifetimes that are exponentially distributed with parameter λ. Five new components are put on test, the first one fails at 100 days, and no further observations are recorded.

   (a) What is the MLE of λ?

   (b) What is the sampling distribution of the MLE?

   (c) What is the standard error of the MLE?

Figure 4.51: Problem 2

PART(A)

The random variable here is the lifetime of a component.

$$X \sim \lambda e^{-\lambda t}$$

In this problem the contribution to the likelihood function of λ comes from only one random variable. Hence we need to find the pdf of this random observation, which is an order statistics. It is the minimum random variable among n random variables where $n = 5$ here.

Since this is an exponential distribution, we know that the distribution of $X_{(1)}$ is given by (from section 3.7, chapter 3, textbook)

$$f_{X_{(1)}}(t) = n\lambda e^{-n\lambda t}$$

Where in the above, the t is the time of the first failures in each sample taken. (sample size is 5 in this problem).

Hence

$$L_n(\lambda) = n\lambda e^{-n\lambda t}$$

so for $n = 5$, the likelihood function is

$$L(\lambda) = 5\lambda e^{-5\lambda t}$$

Hence we need to find the maximum of the above function. Since we have only one r.v., no need to take logs, use standard method:

$$\begin{aligned} L_n'(\lambda) &= ne^{-n\lambda t} - n^2 t\lambda e^{-n\lambda t} \\ &= 0 \end{aligned}$$

Solve for $\hat{\lambda}$

$$\begin{aligned} 1 - nt\hat{\lambda} &= 0 \\ \hat{\lambda} &= \frac{1}{nt} \end{aligned}$$

But here $n = 5$ and time of first failure is $t = 100$ hence the above becomes (write $T = 100$) then we have

$$\hat{\lambda}_{n=5} = \frac{1}{5T} = \frac{1}{500}$$

PART(B)

Since $\hat{\lambda} = \frac{1}{5T}$ where T is a r.v (the first time to fail) which has the distribution $5\lambda e^{-5\lambda t}$, Hence we conclude that the distribution of $\hat{\lambda} \sim \frac{1}{5}\frac{1}{5\lambda e^{-5\lambda t}}$ But an exponential distribution is $\tau e^{-\tau t}$, hence now we see that sampling distribution of $\hat{\lambda}$ ~multiple one over an exponential distribution with parameter $(\tau = 5\lambda)$.

(When asked to find distribution of some r.v., do we always have to express in terms of "known" distributions?)

PART(C)

We need to find the standard deviation of the sampling distribution of $\hat{\lambda}$ found above.

Since we found that $\hat{\lambda} \sim \frac{1}{\text{exponential distribution with parameter}(\tau)}$ and the variance of an exponential with parameter τ is $\frac{1}{\tau^2}$, hence variance of $\hat{\lambda} = \tau^2$

Hence standard error is the square root of this variance. $\boxed{\text{Hence standard error of the MLE } \hat{\lambda} = \tau = 5\lambda}$

3. Do problem 43(a-e) of Chapter 8. To obtain MLE of α, you may use the function "fzero" in MATLAB, or the function "uniroot" in R.

43. The file `gamma-arrivals` contains another set of gamma-ray data, this one consisting of the times between arrivals (interarrival times) of 3,935 photons (units are seconds).

a. Make a histogram of the interarrival times. Does it appear that a gamma distribution would be a plausible model?

b. Fit the parameters by the method of moments and by maximum likelihood. How do the estimates compare?

c. Plot the two fitted gamma densities on top of the histogram. Do the fits look reasonable?

d. For both maximum likelihood and the method of moments, use the bootstrap to estimate the standard errors of the parameter estimates. How do the estimated standard errors of the two methods compare?

e. For both maximum likelihood and the method of moments, use the bootstrap to form approximate confidence intervals for the parameters. How do the confidence intervals for the two methods compare?

f. Is the interarrival time distribution consistent with a Poisson process model for the arrival times?

Figure 4.52: Problem 3

PART (A)

First, I want to say that I am using the following defintion of the Gamma function (using β instead of λ) in the defintion. Since The data given has units of time and are not rate (i.e. 1/time). So I am using this definition of Gamma PDF

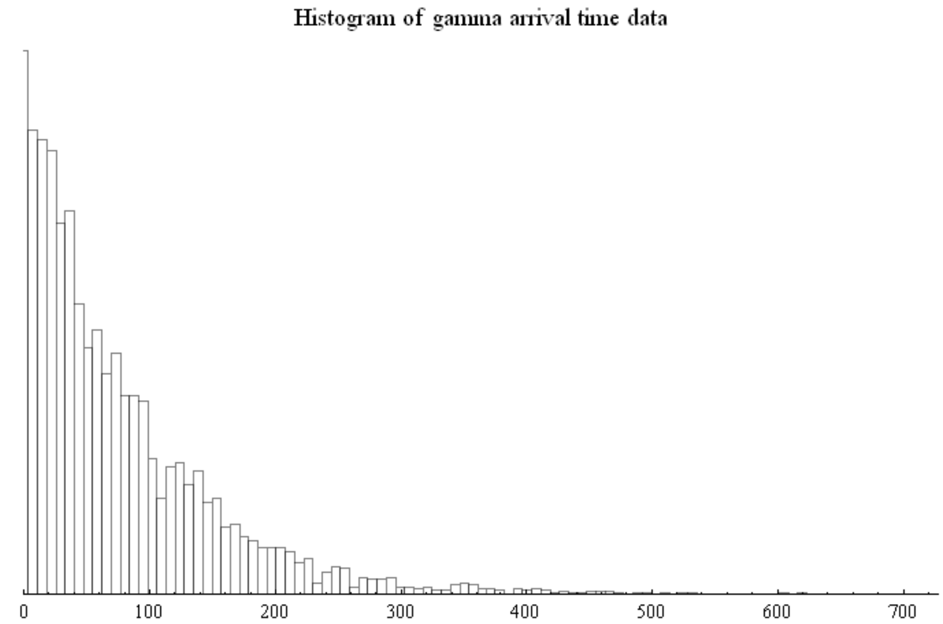
$$f(t) = \frac{1}{\beta}\frac{1}{\Gamma(\alpha)}t^{\alpha-1}e^{-\frac{t}{\beta}}$$

Now to answer part (A).

Yes. The following shows the histogram of the data, and a plot of a Gamma distribution with the shape parameter $\alpha = 1$ and scale parameter β set to the average of the data.

```
In[71]:= nBins = 100;
         gz = nmaMakeDensityHistogram[data, nBins];
         histPlot = GeneralizedBarChart[gz, BarStyle → White,
           ImageSize → 450, PlotRange → {{Min[data], Max[data]}, All},
           PlotLabel → "Histogram of gamma arrival time data"]
```

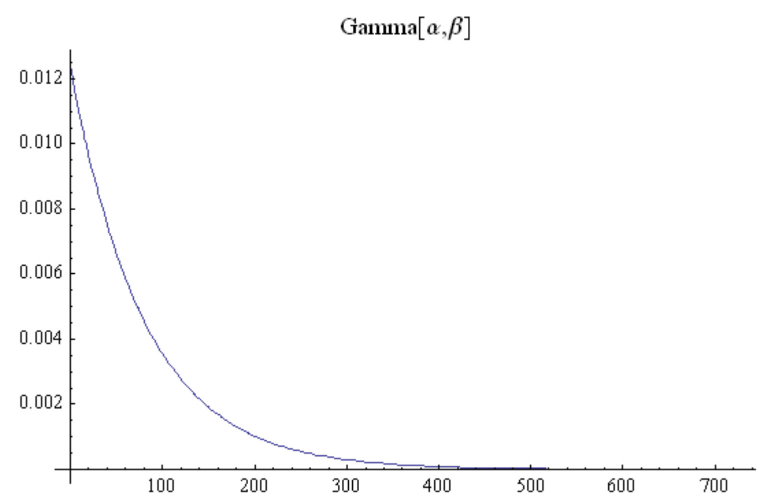
Out[73]=



Let me plot a Gamma distribution with the mean arrival time of the data

```
In[74]:= β = Mean[data]
Out[74]= 79.9352

In[75]:= α = 1;

In[76]:= Plot[PDF[GammaDistribution[α, β], x], {x, 0, Max[data]},
          PlotRange → All, PlotLabel → "Gamma[α,β]"]
```

Out[76]=



From the above we see that a Gamma distribution is possible.

Figure 4.53: histogram of the data

PART(B)

Using method of moments. We need 2 equations since we have to estimate 2 parameters α, λ. For Gamma

$$\mu_1 = \frac{\alpha}{\lambda}$$
$$\begin{aligned}\mu_2 &= Var(X) + (E(X))^2 \\ &= \frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{\lambda}\right)^2 \\ &= \frac{\alpha}{\lambda^2}(\alpha + 1)\end{aligned}$$

Now from the data itself, calculate the First and Second moments and equate to the above and solve for α, λ and these will be our estimate. This little code does the above

Part (B)

Start by finding the first and second moments of the Gamma distribution

```
In[166]:= Clear[x, α, β]
          m1 = ExpectedValue[x, GammaDistribution[α, β], x]
Out[167]= α β

In[168]:= m2 = ExpectedValue[x^2, GammaDistribution[α, β], x]
Out[168]= α (1 + α) β^2
```

Now find the first and second moments of the data itself

```
In[169]:= m1ForData = Mean[data]
Out[169]= 79.9352

In[170]:= m2ForData = Variance[data] + (Mean[data])^2
Out[170]= 12702.9
```

Now solve for α, β

```
In[171]:= sol = First@Solve[{m1 == m1ForData, m2 == m2ForData}, {α, β}];
          αByMoments = α /. sol
Out[172]= 1.01209

In[173]:= βByMoments = β /. sol
Out[173]= 78.98
```

Hence we find λ

```
In[176]:= λ = 1/βByMoments
Out[176]= 0.0126614
```

Figure 4.54: First and Second moments

Now using the MLE method. For α

$$
\begin{aligned}
L(\alpha,\lambda) &= \prod_i^n \frac{\lambda^\alpha}{\Gamma(\alpha)} X_i^{\alpha-1} e^{-\lambda X_i} \\
l(\alpha,\lambda) &= \sum_i^n \log \frac{\lambda^\alpha}{\Gamma(\alpha)} X_i^{\alpha-1} e^{-\lambda X_i} \\
&= \sum_i^n \log \frac{\lambda^\alpha}{\Gamma(\alpha)} + (\alpha-1)\sum_i^n \log X_i - \lambda \sum_i^n X_i \\
&= n\alpha \log\lambda - n\log\Gamma(\alpha) + (\alpha-1)\sum_i^n \log X_i - \lambda\sum_i^n X_i
\end{aligned}
$$

Hence we obtain the 2 equations

$$
\begin{aligned}
\frac{\partial l(\alpha,\lambda)}{\partial \alpha} &= n\log\lambda - n\ polyGamma(0,\alpha) + \sum_i^n \log X_i \\
\frac{\partial l(\alpha,\lambda)}{\partial \lambda} &= n\alpha\frac{1}{\lambda} - \sum_i^n X_i
\end{aligned}
$$

From the second equation, set it to zero we obtain

$$
\hat{\lambda} = \frac{n\hat{\alpha}}{\sum_i^n X_i} = \frac{\hat{\alpha}}{\bar{X}}
$$

Substitute the above in the first equation and set to zero we obtain

$$
\begin{aligned}
0 &= n\log\left(\frac{\hat{\alpha}}{\bar{X}}\right) - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_i^n \log X_i \\
0 &= n\log(\hat{\alpha}) - n\log\bar{X} - n\ polyGamma(0,\alpha) + \sum_i^n \log X_i
\end{aligned}
$$

And solve for $\hat{\alpha}$. Once we find $\hat{\alpha}$ we then find also $\hat{\lambda} = \frac{\hat{\alpha}}{\bar{X}}$

Now find estimates of the parameters by MLE method and do the same as above

In[299]:=
```
Clear[α, λ]
n = Length[data]
```
Out[300]=
```
3935
```
In[301]:=
```
xBar = Mean[data]
```
Out[301]=
```
79.9352
```
In[302]:=

$$eq1 = n\,Log[\alpha] - n\,Log[xBar] - n\,PolyGamma[0, \alpha] + \sum_{i=1}^{n} Log[data[[i]]]$$

Out[302]=
```
-2206.38 + 3935 Log[α] - 3935 PolyGamma[0, α]
```
Solve the above numerically using FindRoot

In[303]:=
```
FindRoot[eq1, {α, 1}];
αMLE = α /. %
```
Out[304]=
```
1.02633
```
Now that we found MLE for alpha, use it to find MLE for λ

In[305]:=

$$\lambda MLE = \frac{\alpha MLE}{xBar}$$

Out[305]=
```
0.0128395
```
Make a table to compare the α, λ found by the above 2 methods

In[306]:=
```
TableForm[{{αByMoments, λByMoments}, {αMLE, λMLE}},
 TableHeadings → {{"Method of moments", "MLE"}, {"α", "λ"}}]
```
Out[306]//TableForm=

	α	λ
Method of moments	1.01209	0.0126614
MLE	1.02633	0.0128395

Figure 4.55: result

Part(C)

Now Fit this model again, and compare the MLE fitting to the method of moments fitting

```
In[350]:= β = 1/λByMoments; (*to Make Mathematica convention of the Gamma function*)
pdfByMomentMethod = Plot[PDF[GammaDistribution[αByMoments, β], x],
  {x, 0, Max[data]}, PlotRange → All, PlotLabel → "Gamma[α,λ]",
  PlotStyle → {Red, Thick}];
p1 = Show[{histPlot, pdfByMomentMethod},
   PlotLabel → "Fitting Gamma by method of moments on data", ImageSize → 300];
```

```
In[352]:= pdfByMLE = Plot[PDF[GammaDistribution[αMLE, 1 / λMLE], x], {x, 0, Max[data]},
   PlotRange → All, PlotLabel → "Gamma[α,λ]", PlotStyle → {Red, Thick}];
p2 = Show[{histPlot, pdfByMLE}, PlotLabel → "Fitting Gamma by method of MLE",
   ImageSize → 300];
Grid[ {{p2, p1}}]
```

Out[354]=

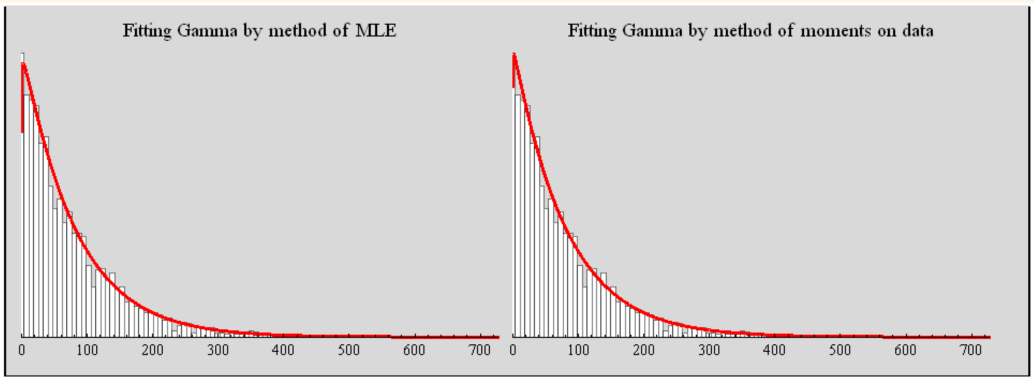


Figure 4.56: MLE fitting to the method of moments fitting

This plot shows more closely the fitting on top of each others. They are very close so hard to see the difference other than near the high frequency part.

```
In[554]:= pdfByMomentMethod = Plot[PDF[GammaDistribution[αByMoments, β], x],
   {x, 0, Max[data]}, PlotRange → {{0, 300}, All}, PlotLabel → "Gamma[α,λ]",
   PlotStyle → {Dashed, Thick}];
pdfByMLE = Plot[PDF[GammaDistribution[αMLE, 1 / λMLE], x], {x, 0, Max[data]},
   PlotRange → {{0, 300}, All}, PlotLabel → "Gamma[α,λ]", PlotStyle → {Red, Thick}];
p2 = Show[{histPlot, pdfByMomentMethod, pdfByMLE},
  PlotLabel → "Fitting Gamma by method of MLE and Moments", ImageSize → 300,
  PlotRange → {{0, 150}, All}]
```

Out[556]=

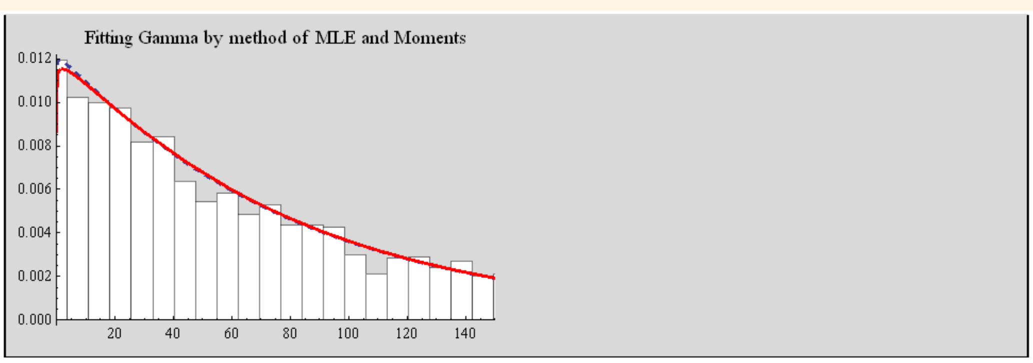


Figure 4.57: the fitting on top of each others

The fits above both look reasonable.

Part(d)

Use bootstrap method.

For the method of moments.

Try for $n = 500$ be the same size. Use the method of moments parameters to generate an n random variables from Gamma distribution. First time use the parameters estimated from the data as shown above.

Now, use the sample generated above to estimate the parameters from it again using also the method of moments. Use these parameters to generate another n random variables. repeat this process for say $N = 5000$ and find the variances of the parameters α, λ, and hence we find the standard error which is the square root of these variances.

Here is the code to do the above and the result

(Last minute update), I am getting large result for standard error from the bootstrap method. I think I have something wrong. Here is the result I get and the code

For Method of moments, I get standard error for alpha=918 and for lambda=18

Part (d)

In[1]:=
```
Remove["Global`*"]

SeedRandom[01010101];
m1 = ExpectedValue[x, GammaDistribution[α, β], x];
m2 = ExpectedValue[x^2, GammaDistribution[α, β], x];

getMethodOfMomentsParameters[data_] := Module[{sol, αByMoments, βByMoments},
  m1ForData = Mean[data];
  m2ForData = Variance[data] + (m1ForData)^2;
  sol = First@Solve[{m1 == m1ForData, m2 == m2ForData}, {α, β}];
  αByMoments = α /. sol;
  βByMoments = β /. sol;

  {αByMoments, βByMoments}
 ]
```

In[6]:=
```
file =
  "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";
data = Flatten[Import[file, "Table"]];

n = 500; (*sample size*)
nIter = 5000; (*number of iterations*)
alpha = Table[0, {i, nIter}];
beta = Table[0, {i, nIter}];

{alpha[[1]], beta[[1]]} = getMethodOfMomentsParameters[data];

For[i = 2, i ≤ nIter, i ++,
 {
  sample = RandomReal[GammaDistribution[alpha[[i - 1]], beta[[i - 1]]], n];
  {alpha[[i]], beta[[i]]} = getMethodOfMomentsParameters[sample];
 }
]
Print["Standard error for alpha=", Sqrt[Variance[alpha]]]
Print["Standard error for lambda=", Sqrt[Variance[1 / beta]]]
```
```
Standard error for alpha=918.308
Standard error for lambda=18.7966
```

Figure 4.58: result

For MLE I get

Standard error for alpha=1.68697*10^8

Standard error for lambda=60.2585

Now do the same for MLE method

```
In[1]:= Remove["Global`*"]

SeedRandom[01010101];

getMLEParameters[data_] := Module[{sol, xBar, αMLE, λMLE, eq, α, n},
  xBar = Mean[data];
  n = Length[data];
  eq = n Log[α] - n Log[xBar] - n PolyGamma[0, α] + Sum[Log[data[[i]]], {i, 1, n}];
  sol = FindRoot[eq, {α, 1}];
  αMLE = α /. sol;
  λMLE = αMLE/xBar;

  {αMLE, λMLE}
 ]
```

Remove::rmnsm : There are no symbols matching "Global`*". ≫

```
In[4]:= file =
  "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";
data = Flatten[Import[file, "Table"]];

n = 500; (*sample size*)
nIter = 5000; (*number of iterations*)
alpha = Table[0, {i, nIter}];
lambda = Table[0, {i, nIter}];

{alpha[[1]], lambda[[1]]} = getMLEParameters[data];

For[i = 2, i ≤ nIter, i++,
 {
  sample = RandomReal[GammaDistribution[alpha[[i - 1]], lambda[[i - 1]]], n];
  {alpha[[i]], lambda[[i]]} = getMLEParameters[sample];
 }
]
Print["Standard error for alpha=", Sqrt[Variance[alpha]]]
Print["Standard error for lambda=", Sqrt[Variance[lambda]]]
```

Figure 4.59: result

Part (e) and (f)

Run out of time.

4.7.1 corrected problem 3

Problem 3 Quiz 7 part(a)
by Nasser Abbasi

3. Do problem 43(a-e) of Chapter 8. To obtain MLE of α, you may use the functic
"fzero" in MATLAB, or the function "uniroot" in R.

43. The file `gamma-arrivals` contains another set of gamma-ray data, this one consisting of the times between arrivals (interarrival times) of 3,935 photons (units are seconds).

a. Make a histogram of the interarrival times. Does it appear that a gamma distribution would be a plausible model?
b. Fit the parameters by the method of moments and by maximum likelihood. How do the estimates compare?
c. Plot the two fitted gamma densities on top of the histogram. Do the fits look reasonable?
d. For both maximum likelihood and the method of moments, use the bootstrap to estimate the standard errors of the parameter estimates. How do the estimated standard errors of the two methods compare?
e. For both maximum likelihood and the method of moments, use the bootstrap to form approximate confidence intervals for the parameters. How do the confidence intervals for the two methods compare?
f. Is the interarrival time distribution consistent with a Poisson process model for the arrival times?

Part A

PART (A)

First, I want to say that I am using the following defintion of the Gamma function (using β instead of λ) in the defintion. Si
given has units of time and are not rate (i.e. 1/time). So I am using this definition of Gamma PDF

$$\boxed{f(t) = \frac{1}{\beta}\frac{1}{\Gamma(\alpha)}t^{\alpha-1}e^{-\frac{t}{\beta}}}$$

Now to answer part (A).

Yes. The following shows the histogram of the data, and a plot of a Gamma distribution with the shape parameter $\alpha = 1$
parameter β set to the average of the data.

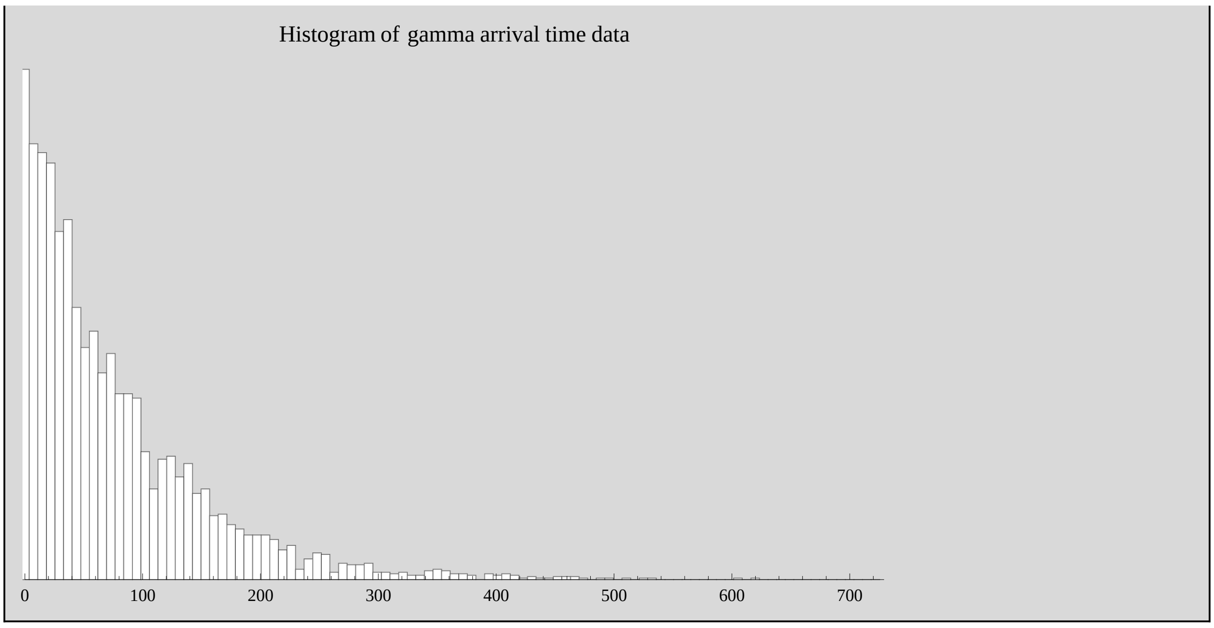
Open the gamma file and load the data

```
Remove["Global`*"]
```

Decide on number of bins and make a histogram (see appendix for function I wrote to make probability histogram)

In[5]:=
```
nBins = 100;
gz = nmaMakeDensityHistogram[data, nBins];
histPlot = GeneralizedBarChart[gz, BarStyle → White,
  ImageSize → 450, PlotRange → {{Min[data], Max[data]}, All},
  PlotLabel → "Histogram of gamma arrival time data"]
```

Out[7]=



Let me plot a Gamma distribution with the mean arrival time of the data

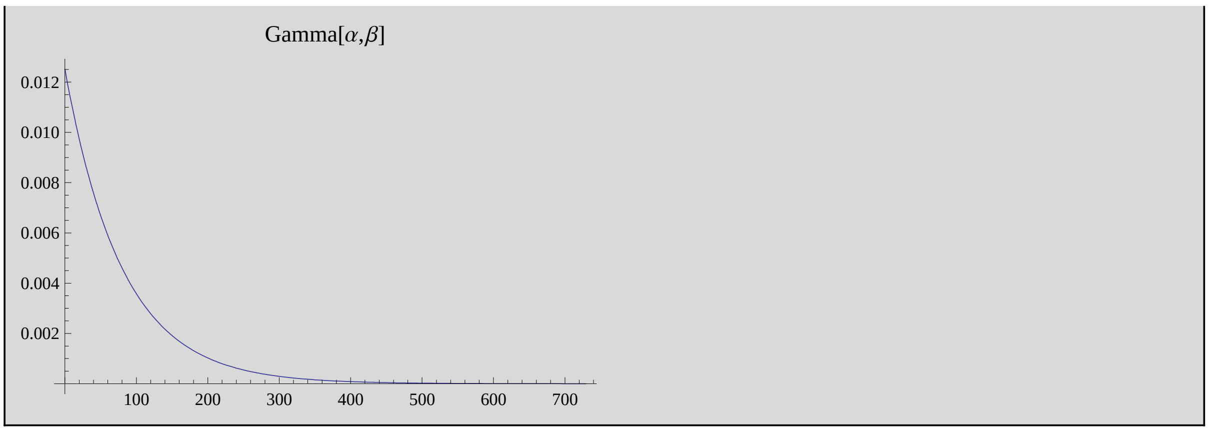
In[8]:=
```
β = Mean[data]
```

Out[8]=
```
79.9352
```

In[9]:=
```
α = 1;
```

In[10]:=
```
Plot[PDF[GammaDistribution[α, β], x],
 {x, 0, Max[data]}, PlotRange → All, PlotLabel → "Gamma[α,β]"]
```

Out[10]=



From the above we see that a Gamma distribution is possible.

Part B

PART(B)

Using method of moments. We need 2 equations since we have to estimate 2 parameters α, λ. For Gamma

$$\begin{aligned}\mu_1 &= \frac{\alpha}{\lambda}\\ \mu_2 &= Var(X) + (E(X))^2\\ &= \frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{\lambda}\right)^2\\ &= \frac{\alpha}{\lambda^2}(\alpha+1)\end{aligned}$$

Now from the data itself, calculate the First and Second moments and equate to the above and solve for α, λ and these estimate. This little code does the above

Start by finding the first and second moments of the Gamma distribution

In[11]:=
```
Clear[x, α, β]
m1 = ExpectedValue[x, GammaDistribution[α, β], x]
```

Out[12]= $\alpha\,\beta$

In[13]:=
```
m2 = ExpectedValue[x^2, GammaDistribution[α, β], x]
```

Out[13]= $\alpha\,(1+\alpha)\,\beta^2$

Now find the first and second moments of the data itself

In[14]:=
```
m1ForData = Mean[data]
```
Out[14]=
```
79.9352
```
In[15]:=
```
m2ForData = Variance[data] + (Mean[data])^2
```
Out[15]=
```
12702.9
```
Now solve for α, β

In[16]:=
```
sol = First@Solve[{m1 == m1ForData, m2 == m2ForData}, {α, β}];
αByMoments = α /. sol
```
Out[17]=
```
1.01209
```
In[18]:=
```
βByMoments = β /. sol
```
Out[18]=
```
78.98
```
Hence we find λ

Now find estimates of the parameters by MLE method and do the same as above

Now using the MLE method. For α :

$$L(\alpha,\lambda) = \prod_i^n \frac{\lambda^\alpha}{\Gamma(\alpha)} X_i^{\alpha-1} e^{-\lambda X_i}$$

$$l(\alpha,\lambda) = \sum_i^n \log \frac{\lambda^\alpha}{\Gamma(\alpha)} X_i^{\alpha-1} e^{-\lambda X_i}$$

$$= \sum_i^n \log \frac{\lambda^\alpha}{\Gamma(\alpha)} + (\alpha-1)\sum_i^n \log X_i - \lambda \sum_i^n X_i$$

$$= n\alpha \log\lambda - n\log\Gamma(\alpha) + (\alpha-1)\sum_i^n \log X_i - \lambda \sum_i^n X_i$$

Hence we obtain the 2 equations

$$\frac{\partial l(\alpha,\lambda)}{\partial \alpha} = n\log\lambda - n\,polyGamma(0,\alpha) + \sum_i^n \log X_i$$

$$\frac{\partial l(\alpha,\lambda)}{\partial \lambda} = n\alpha\frac{1}{\lambda} - \sum_i^n X_i$$

From the second equation, set it to zero we obtain

$$\hat{\lambda} = \frac{n\hat{\alpha}}{\sum_i^n X_i} = \frac{\hat{\alpha}}{\bar{X}}$$

Substitute the above in the first equation and set to zero we obtain

$$0 = n\log\left(\frac{\hat{\alpha}}{\bar{X}}\right) - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_i^n \log X_i$$

$$0 = n\log(\hat{\alpha}) - n\log\bar{X} - n\,polyGamma(0,\alpha) + \sum_i^n \log X_i$$

And solve for $\hat{\alpha}$. Once we find $\hat{\alpha}$ we then find also $\hat{\lambda} = \frac{\hat{\alpha}}{\bar{X}}$

In[19]:=
```
Clear[α, β]
n = Length[data]
```

Out[20]= 3935

In[21]:=
```
xBar = Mean[data]
```

Out[21]= 79.9352

In[22]:=
```
eq1 = n Log[α] - n Log[xBar] - n PolyGamma[0, α] + Sum[Log[data[[i]]], {i, 1, n}]
```

Out[22]=
```
-2206.38 + 3935 Log[α] - 3935 PolyGamma[0, α]
```

Solve the above numerically using FindRoot

In[23]:=
```
FindRoot[eq1, {α, 1}];
αMLE = α /. %
```

Out[24]=
```
1.02633
```

Now that we found MLE for alpha, use it to find MLE for λ

In[25]:=
```
λMLE = αMLE / xBar
```

Out[25]=
```
0.0128395
```

In[26]:=
```
βMLE = 1 / λMLE
```

Out[26]=
```
77.8844
```

Conclusion for part B

Make a table to compare the α, λ found by the above 2 methods

In[27]:=
```
TableForm[{{αByMoments, "\t", βByMoments}, {αMLE, "\t", βMLE}},
 TableHeadings → {{"Method of moments", "MLE"}, {"α", "", "β"}}]
```

Out[27]//TableForm=

	α	β
Method of moments	1.01209	78.98
MLE	1.02633	77.8844

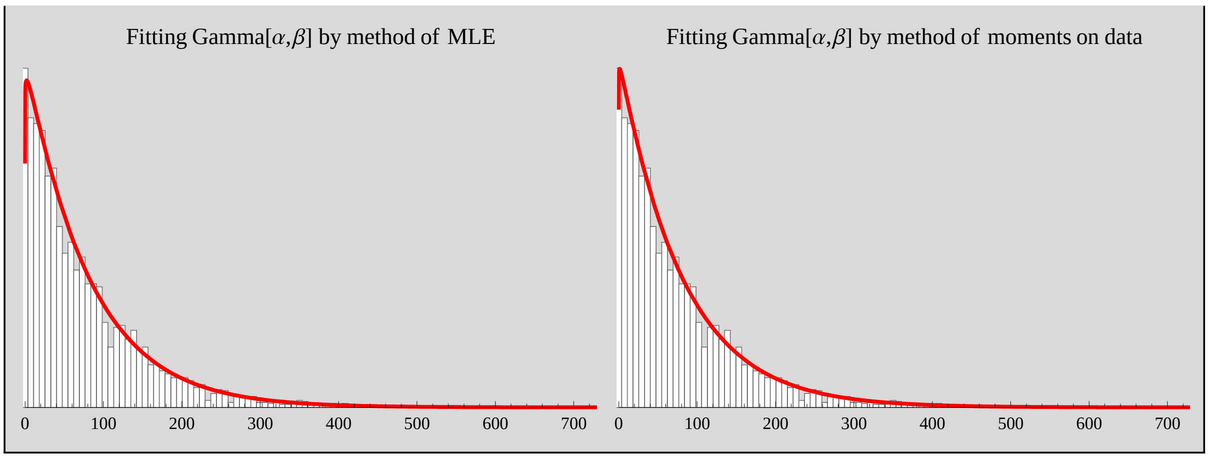
Part C

Put the 2 fitting on top of each others to compare

In[28]:=
```
pdfByMomentMethod = Plot[PDF[GammaDistribution[αByMoments, βByMoments], x],
   {x, 0, Max[data]}, PlotRange → All, PlotLabel → "Gamma[α,β]", PlotStyle → {Red, Thick}];
p1 = Show[{histPlot, pdfByMomentMethod},
   PlotLabel → "Fitting Gamma[α,β] by method of moments on data", ImageSize → 300];
```

In[30]:=
```
pdfByMLE = Plot[PDF[GammaDistribution[αMLE, βMLE], x], {x, 0, Max[data]},
   PlotRange → All, PlotLabel → "Gamma[α,β]", PlotStyle → {Red, Thick}];
p2 = Show[{histPlot, pdfByMLE}, PlotLabel → "Fitting Gamma[α,β] by method of MLE",
   ImageSize → 300];
Grid[ {{p2, p1}}]
```

Out[32]=

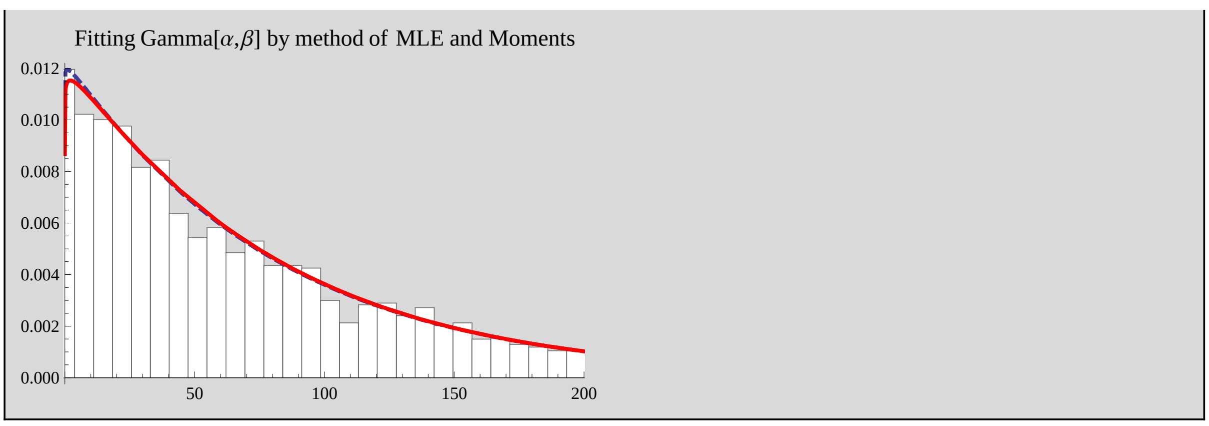


Conclusion for part (C)

This plot shows more closely the fitting on top of each others.They are very close so hard to see the difference other than near the high frequency part.

In[33]:=
```
pdfByMomentMethod = Plot[PDF[GammaDistribution[αByMoments, βByMoments], x], {x, 0, Max[data]},
   PlotRange → {{0, 300}, All}, PlotLabel → "Gamma[α,β]", PlotStyle → {Dashed, Thick}];
pdfByMLE = Plot[PDF[GammaDistribution[αMLE, 1 / λMLE], x], {x, 0, Max[data]},
   PlotRange → {{0, 300}, All}, PlotLabel → "Gamma[α,λ]", PlotStyle → {Red, Thick}];
p2 = Show[{histPlot, pdfByMomentMethod, pdfByMLE},
  PlotLabel → "Fitting Gamma[α,β] by method of MLE and Moments",
  ImageSize → 300, PlotRange → {{0, 200}, All}]
```

Out[35]=



Part (d)

I use bootstrap to compare the standard error of parameters estimated using both the MLE and Method of Moments.
In Bootstrap, we find the estimate of the parameters from the data itself. Then we generate random data from the same distribution of the same size as the data itself using these original estimates, then we estimate the parameters again from the generated random data. We repeat this process keeping track each time of the parameters estimated from each sample. At the end we find the variance and the mean of these parameters and also make histogram to compare the standard errors from the MLE and the moments method. The method which has smaller standard error will be better.

Write 2 functions which find the parameters estimates using the MLE and the moments method

```
Remove["Global`*"]
```

In[36]:=
```
m1 = ExpectedValue[x, GammaDistribution[α, β], x];
m2 = ExpectedValue[x^2, GammaDistribution[α, β], x];

getMethodOfMomentsParameters[data_] := Module[{sol, αByMoments, βByMoments},
  m1ForData = Mean[data];
  m2ForData = Variance[data] + (m1ForData)^2;
  sol = First@Solve[{m1 == m1ForData, m2 == m2ForData}, {α, β}];
  αByMoments = α /. sol;
  βByMoments = β /. sol;

  {αByMoments, βByMoments}
 ]
```

In[39]:=
```
getMLEParameters[data_] := Module[{sol, xBar, αMLE, λMLE, eq, α, n},
  xBar = Mean[data];
  n = Length[data];
  eq = n Log[α] - n Log[xBar] - n PolyGamma[0, α] + Sum[Log[data[[i]]], {i, 1, n}];
  sol = FindRoot[eq, {α, 1}];
  αMLE = α /. sol;
  λMLE = αMLE / xBar;

  {αMLE, 1 / λMLE}
 ]
```

In[40]:=
```
getStandardError[numberOfIteration_, data_, sizeOfData_, f_] := Module[{},
  alpha = Table[0, {i, numberOfIteration}];
  beta = Table[0, {i, numberOfIteration}];

  {dataAlpha, dataBeta} = f[data];

  For[i = 1, i ≤ numberOfIteration, i++,
   {
    sample = RandomReal[GammaDistribution[dataAlpha, dataBeta], sizeOfData];
    {alpha⟦i⟧, beta⟦i⟧} = f[sample];
   }
  ];
  {Sqrt[Variance[alpha]], Mean[alpha], Sqrt[Variance[beta]], Mean[beta]}
 ]
```

Now call the above functions to find the standard errors

In[41]:=
```
SeedRandom[01 010 101];

file = "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";

data = Flatten[Import[file, "Table"]];
sizeOfData = Length[data]; (*data size*)
numberOfIterations = 5000;
```

In[46]:=
```
resultMoments =
 getStandardError[numberOfIterations, data, sizeOfData, getMethodOfMomentsParameters]
```

Out[46]=
```
{0.0319471, 1.01321, 2.76292, 78.9663}
```

Now do the same for MLE method

In[47]:=
```
resultMLE = getStandardError[numberOfIterations, data, sizeOfData, getMLEParameters];
```

Summary table for part D

In[48]:=
```
TableForm[{ {resultMoments[[1]], " ",
   resultMoments[[2]], " ", resultMoments[[3]], " ", resultMoments[[4]]},
  {resultMLE[[1]], " ", resultMLE[[2]], " ", resultMLE[[3]], " ", resultMLE[[4]]}},
 TableHeadings →
  {{"Method of moments", "MLE"}, {"std(α)", " ", "Mean(α)", " ", "std(β)", " ", "Mean(β)"}}]
```

Out[48]//TableForm=

	std(α)	Mean(α)	std(β)	Mean(β)
Method of moments	0.0319471	1.01321	2.76292	78.9663
MLE	0.0203902	1.02741	1.97021	77.8397

Conclusion for part D

We see from the table above that standard error for the MLE method for the parameter α is a little smaller than that for the moments method. Also for the β parameter, the MLE method gives a smaller standard error than the moments method.

We conclude that the MLE is better than the moments method since its standard error is smaller. The difference is not too large here in standard error, I was expecting a larger error than this. May be if we have larger data size we can see this.

Part E

Now we need to find confidence intervals for both α and β based on the MLE and the moments methods estimates. Hence we will need to generate 4 C.I's since we have 2 parameters and 2 methods. I will use the 95% C.I. in all cases. This C.I. is then a random variable which is said to contain the parameter being estimated with probability .95
There are 3 methods to determine the CI for the method of moments and the MLE. Exact methods, approximation based on large sample size, and bootstrap methods. We are asked to use the bootstrap method here.

Algorithm

We start first by writing down the algorithm for finding CI (95 %) for MLE or Moments methods. This algorithm can ofcourse be used for other CI other than 95 % by changing the quantile value, but I used 95 % here for illustration only

1) Load the data from file
2) use the method of MLE to estimate the parameters. Let the estimated parameters be $\hat{\alpha}$ and $\hat{\beta}$
3) Now, using $\hat{\alpha}$ and $\hat{\beta}$ generate 1000 random samples each of size equal to the original data. Use the distribution Gamma.
4) From each sample generated in step (3) determine using MLE an estimate of the sample parameters α,β call these $\alpha_j{}^*$ and $\beta_j{}^*$ where j is the sample number
5) Sort the parameters sequence α^* and the parameters sequence β^* from small to large
6) Find the parameter α^* at position 25 and at position 975. Call these $\alpha^*{}_{25}$ and $\alpha^*{}_{975}$ and do the same for β^* and Call these $\beta^*{}_{25}$ and $\beta^*{}_{975}$
7)Let $\underline{\delta} = \alpha^*{}_{25}$-$\hat{\alpha}$ and let $\bar{\delta} = \alpha^*{}_{975}$-$\hat{\alpha}$, Hence the 95% CI for $\hat{\alpha}$=($\hat{\alpha}$-$\underline{\delta}$,$\hat{\alpha}$-$\bar{\delta}$)
8)Let $\underline{\delta} = \beta^*{}_{25}$-$\hat{\alpha}$ and let $\bar{\delta} = (\beta^*)_{975}$-$\hat{\alpha}$, Hence the 95% CI for $\hat{\beta}$=($\hat{\beta}$-$\underline{\delta}$,$\hat{\beta}$-$\bar{\delta}$)

Code Implementation

In[49]:=
```
SeedRandom[01 010 101];

file = "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";

data = Flatten[Import[file, "Table"]];
sizeOfData = Length[data]; (*data size*)
numberOfIterations = 1000;

getCI[data_, f_] := Module[{},
  alphaStar = Table[0, {i, numberOfIterations}];
  betaStar = Table[0, {i, numberOfIterations}];

  {alphaHat, betaHat} = f[data];

  For[i = 1, i ≤ numberOfIterations, i++,
   {
    sample = RandomReal[GammaDistribution[alphaHat, betaHat], sizeOfData];
    {alphaStar⟦i⟧, betaStar⟦i⟧} = f[sample];
   }
  ];
  alphaStar = Sort[alphaStar];
  betaStar = Sort[betaStar];
  lowerQAlpha = alphaStar⟦25⟧;
  upperQAlpha = alphaStar⟦975⟧;
  lowerQBeta = betaStar⟦25⟧;
  upperQBeta = betaStar⟦975⟧;

  {alphaHat - (upperQAlpha - alphaHat), alphaHat - (lowerQAlpha - alphaHat),
   betaHat - (upperQBeta - betaHat), betaHat - (lowerQBeta - betaHat)}
 ]

{alphaLow, alphaHigh, betaLow, betaHigh} = getCI[data, getMethodOfMomentsParameters];

Print["95% C.I. for α using Method of moments is (" <>
   ToString[alphaLow] <> "," <> ToString[alphaHigh] <> ")"];
Print["95% C.I. for β using Method of moments is (" <>
   ToString[betaLow] <> "," <> ToString[betaHigh] <> ")"];
```

```
95% C.I. for α using Method of moments is (0.948008,1.07588)
95% C.I. for β using Method of moments is (73.1629,83.9168)
```

Now do the same for the MLE

In[58]:=
```
{alphaLow, alphaHigh, betaLow, betaHigh} = getCI[data, getMLEParameters];

Print[
  "95% C.I. for α using MLE is (" <> ToString[alphaLow] <> "," <> ToString[alphaHigh] <> ")"];
Print["95% C.I. for β using MLE is (" <> ToString[betaLow] <> "," <> ToString[betaHigh] <> ")"];
```

```
95% C.I. for α using MLE is (0.985055,1.06597)
95% C.I. for β using MLE is (73.9381,81.7684)
```

Conclusion for part E

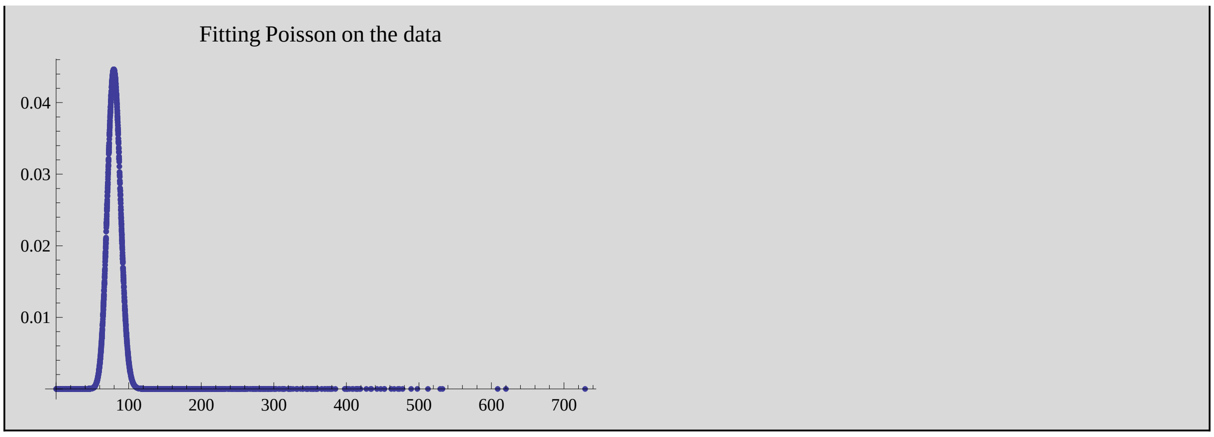
The confidence interval using MLE for the same percentage is smaller than that for the method of moments. This is for both the α and β.
This is good. The smaller the CI is for the same confidence the better it is. This shows that the MLE method of estimation is better than the method of moments (for large sample size as the case here).

Part f

To answer this part, we use MLE to try to fit Poisson distribution on the data. We use MLE to estimate the λ parameter for the poisson distribution and plot the pdf over the data histogram to see if the fit is good
We know that the MLE estimate of λ for a Poisson is $\overline{X}$

```
Remove["Global`*"]
```

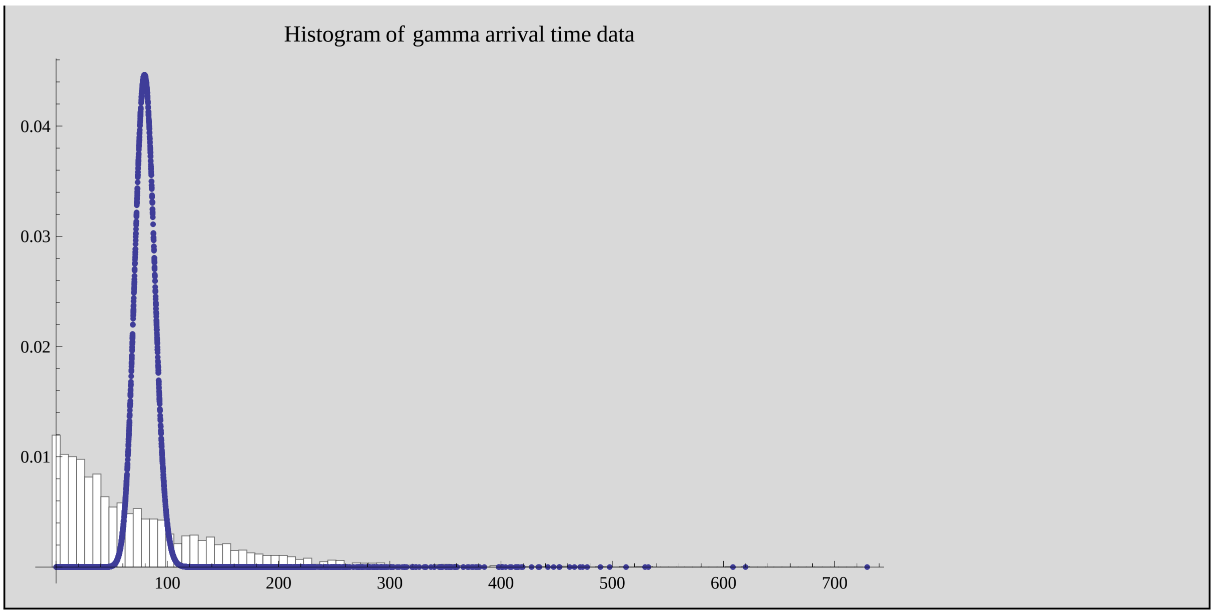
```
lambdaHat = Mean[data];
dist = PDF[PoissonDistribution[lambdaHat], t];
poissonData = Table[{data[[i]], dist /. t -> data[[i]]}, {i, 1, Length[data]}];
p1 = ListPlot[poissonData, Joined → False,
  PlotRange → All, PlotLabel → "Fitting Poisson on the data"]
```



Lets superimpose the histogram

```
nBins = 100;
gz = nmaMakeDensityHistogram[data, nBins];
histPlot = GeneralizedBarChart[gz, BarStyle → White, ImageSize → 450, PlotRange →
    {{Min[data], Max[data]}, All}, PlotLabel → "Histogram of gamma arrival time data"];
```

```
Show[{histPlot, p1}, PlotRange → All]
```



Conclusion part f

From the above we see that a Poisson distribution is not a good probability law to describe the data. Poisson fits a random variable which represents the number of times an event occur in some fixed time. The parameter λ represents the average rate at which the event occur per unit time.
The interarrivals times (from the histogram of the data itself) shows that more events occur (particles arrive) which has the time between them small than when the time is large between them. But in Poisson, the rate should be fixed and not changing. This explains why Poisson does not give a good fit.
(I really would like to know the correct answer for this one if mine is not correct. Thank you)

Appendix

Contain functions needed

In[1]:=
```
file = "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";
data = Flatten[Import[file, "Table"]];

<< "BarCharts`";
nmaMakeDensityHistogram[originalData_, nBins_] :=
  Module[{freq, binSize, from, to, scaleFactor, j, a, currentArea},
   to = Max[originalData]; from = Min[originalData];
   binSize = (to - from) / nBins; freq = BinCounts[originalData, binSize];
   currentArea = Sum[binSize * freq[[i]], {i, nBins}];
   freq = freq / currentArea; a = from;
   Table[{a + (j - 1) * binSize, freq[[j]], binSize}, {j, 1, nBins}]];
```


Mathemtica notebook for corrected version

Text file of gamma arrivals data

4.7.2 Graded

16/20

14/20 +2 16/20

QUIZ 7 MATH 502AB November 27, 2007

Name (please print) NASSER ABBASI

1. Consider $X_1, \cdots, X_n$, an i.i.d. sample from a random variables with density

$$f(x|\sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right).$$

(a) Find the method of moment estimate of σ.
(b) Find MLE of σ.
(c) Find the asymptotic distribution of the MLE.

PART(A)

In the method of moments we formulate the moments of the probability law of the distribution in which the random variables belong to and equate these moments to the moments obtained from the sample at hand and solve for the unknown parameters.

$$\begin{aligned}
\mu_1 &= E(X) \\
&= \int_{-\infty}^{\infty} x f(x)\, dx \\
&= \frac{1}{2\sigma} \int_{-\infty}^{\infty} x\, e^{-\frac{|x|}{\sigma}}\, dx
\end{aligned}$$

But $e^{-\frac{|x|}{\sigma}}$ is symmtric around the $x = 0$ due to absolute x in the power of the exp.(This assumes σ positive, which is ofcourse true) but it is multiplied by negative x to the left of y-axis and multiplied by positive x on the right of the y-axis, hence the area of the left of the y-axis will be equal but negative to the area on the right of the y-axis. Hence the above integral is zero. Hence $\boxed{\mu_1 = 0}$

This moment provides no information. Find the second moment.

$$\begin{aligned}
\mu_2 &= E\left(X^2\right) \\
&= \frac{1}{2\sigma} \int_{-\infty}^{\infty} x^2 e^{-\frac{|x|}{\sigma}}\, dx
\end{aligned}$$

Due to the symmetry of $e^{-\frac{|x|}{\sigma}}$ and also x^2 is even and symmetrical around $x = 0$, the above integral is then twice the integral from $x = 0 \cdots \infty$ and it becomes

$$\mu_2 = \frac{1}{\sigma} \int_0^{\infty} x^2 e^{-\frac{x}{\sigma}} dx$$

Integration by parts gives

$$\boxed{\mu_2 = 2\sigma^2}$$

1

Hence

$$\boxed{\sigma = \sqrt{\frac{\mu_2}{2}}} \tag{1}$$

Now find μ_2 from the sample itself and substitute for it in the above. From the sample,

$$\begin{aligned}\mu_2 &= Var(sample) + Mean(Sample)\\ &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 + \bar{X}\end{aligned}$$

Since the mean of the population was found to be zero, we can take the mean of the sample $\boxed{\bar{X} = 0}$

Hence μ_2 from the sample becomes

$$\mu_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Replace the above in (1) we obtain estimate of the population σ as

$$\boxed{\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}}$$

PART(B)

The MLE of σ is found as follows. Since i.i.d. random variables we write (Where $L(\sigma)$ mean $lik(\sigma)$ and $l(\sigma)$ means $\log(lik(\sigma))$

$$\begin{aligned}L_n(\sigma) &= \prod_{i=1}^{n} f(X_i|\sigma)\\ l_n(\sigma) &= \sum_{i=1}^{n} \log f(X_i|\sigma)\\ &= \sum_{i=1}^{n} \log\left(\frac{1}{2\sigma}e^{-\frac{|X_i|}{\sigma}}\right)\\ &= \sum_{i=1}^{n}\left(-\log(2\sigma) - \frac{|X_i|}{\sigma}\right)\end{aligned}$$

Therefore

$$\boxed{l_n(\sigma) = -n\log(2\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n}|X_i|}$$

Now we find the MLE, which is the value of σ which maximizes the above function.

2

$$l_n'(\sigma) = 0$$
$$\frac{-n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|X_i| = 0$$
$$-\sigma n + \sum_{i=1}^{n}|X_i| = 0$$

Hence

$$\boxed{\hat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}|X_i|}$$

The above is the MLE estimate of the parameter σ.

PART(C)

The asymptotic distribution of the MLE $\hat{\sigma}$ is normal with mean σ and variance $\frac{1}{nI(\sigma)}$where

$$I(\sigma) = -E\left[l_n''(\sigma)\right]$$

But

$$\begin{aligned} l_n'(\sigma) &= \frac{\partial}{\partial\sigma}\left[-n\log(2\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n}|X_i|\right] \\ &= \frac{-n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}|X_i| \end{aligned}$$

and

$$l_n''(\sigma) = \frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{n}|X_i|$$

Hence

$$\begin{aligned} I(\sigma) &= -E\left[\frac{n}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{n}|X_i|\right] \\ &= -E\left[\frac{n}{\sigma^2}\right] + \frac{2}{\sigma^3}E\left[\sum_{i=1}^{n}|X_i|\right] \\ &= -\frac{n}{\sigma^2} + \frac{2}{\sigma^3}\sum_{i=1}^{n}E|X_i| \qquad (2) \end{aligned}$$

Need to find $E|X_i|$, since i.i.d. all random variables has the same expected value as X, hence

$$\begin{aligned} E|X| &= \frac{1}{2\sigma}\int_{-\infty}^{\infty}|x|\,e^{-\frac{|x|}{\sigma}}dx \\ &= \frac{1}{\sigma}\int_{0}^{\infty}xe^{-\frac{x}{\sigma}}dx \\ &= \sigma \end{aligned}$$

Therefore from (2)

$$I(\sigma) = -\frac{n}{\sigma^2} + \frac{2}{\sigma^3}\sum_{i=1}^{n}\sigma$$
$$= -\frac{n}{\sigma^2} + \frac{2n}{\sigma^2}$$

Hence the Fischer information matrix is

$$\boxed{I(\sigma) = \frac{n}{\sigma^2}}$$

Hence MLE $\hat{\sigma}_n$ has an asymptotic distribution ~ $N\left(\sigma, \frac{1}{nI(\sigma)}\right)$ i.e.

$$E(\hat{\sigma}_n) = \sigma$$

and

$$Var(\hat{\sigma}_n) = \frac{1}{nI(\sigma)}$$
$$= \frac{\sigma^2}{n^2}$$

-1/2

2. Suppose that certain electronic components have lifetimes that are exponentially distributed with parameter λ. Five new components are put on test, the first one fails at 100 days, and no further observations are recorded.

   (a) What is the MLE of λ?

   (b) What is the sampling distribution of the MLE?

   (c) What is the standard error of the MLE?

PART(A)

The random variable here is the lifetime of a component.

$$X\tilde{}\lambda e^{-\lambda t}$$

In this problem the contribution to the likelihood function of λ comes from only one random variable. Hence we need to find the pdf of this random observation, which is an order statistics. It is the minimum random variable among n random variables where $n=5$ here.

Since this is an exponential distribution, we know that the distribution of $X_{(1)}$ is given by (from section 3.7, chapter 3, textbook)

$$f_{X_{(1)}}(t) = n\lambda e^{-n\lambda t}$$

Where in the above, the t is the time of the first failures in each sample taken. (sample size is 5 in this problem).

Hence

$$L_n(\lambda) = n\lambda e^{-n\lambda t}$$

so for $n=5$, the likelihood function is

$$\boxed{L(\lambda) = 5\lambda e^{-5\lambda t}}$$

Hence we need to find the maximum of the above function. Since we have only one r.v., no need to take logs, use standard method:

$$\begin{aligned} L_n'(\lambda) &= ne^{-n\lambda t} - n^2 t\lambda e^{-n\lambda t} \\ &= 0 \end{aligned}$$

Solve for $\hat{\lambda}$

$$\begin{aligned} 1 - nt\hat{\lambda} &= 0 \\ \hat{\lambda} &= \frac{1}{nt} \end{aligned}$$

But here $n=5$ and time of first failure is $t=100$ hence the above becomes (write $T=100$) then we have

$$\boxed{\hat{\lambda}_{n=5} = \frac{1}{5T} = \frac{1}{500}}$$

5

What does this mean?

PART(B)

Since $\hat{\lambda} = \frac{1}{5T}$ where T is a r.v ~~(the first time~~ to fail) which has the distribution $5\lambda e^{-5\lambda t}$, Hence we conclude that the distribution of $\hat{\lambda}\tilde{\ }\frac{1}{5}\frac{1}{5\lambda e^{-5\lambda t}}$ But an exponential distribution is $\tau e^{-\tau t}$, hence now we see that sampling distribution of $\hat{\lambda}\tilde{\ }$multiple one over an exponential distribution with parameter $(\tau = 5\lambda)$.

(When asked to find distribution of some r.v., do we always have to express in terms of "known" distributions?) Yes, ~~to find the density~~

PART(C)

We need to find the standard deviation of the sampling distribution of $\hat{\lambda}$ found above.

Since we found that $\hat{\lambda}\tilde{\ }\frac{1}{\text{exponential distribution with parameter}(\tau)}$ and the variance of an exponential with parameter τ is $\frac{1}{\tau^2}$, hence variance of $\hat{\lambda} = \tau^2$

Hence standard error is the square root of this variance. $\boxed{\text{Hence standard error of the MLE } \hat{\lambda} = \tau = 5\lambda}$

3. Do problem 43(a-e) of Chapter 8. To obtain MLE of α, you may use the function "fzero" in MATLAB, or the function "uniroot" in R.

43. The file `gamma-arrivals` contains another set of gamma-ray data, this one consisting of the times between arrivals (interarrival times) of 3,935 photons (units are seconds).

a. Make a histogram of the interarrival times. Does it appear that a gamma distribution would be a plausible model?

b. Fit the parameters by the method of moments and by maximum likelihood. How do the estimates compare?

c. Plot the two fitted gamma densities on top of the histogram. Do the fits look reasonable?

d. For both maximum likelihood and the method of moments, use the bootstrap to estimate the standard errors of the parameter estimates. How do the estimated standard errors of the two methods compare?

e. For both maximum likelihood and the method of moments, use the bootstrap to form approximate confidence intervals for the parameters. How do the confidence intervals for the two methods compare?

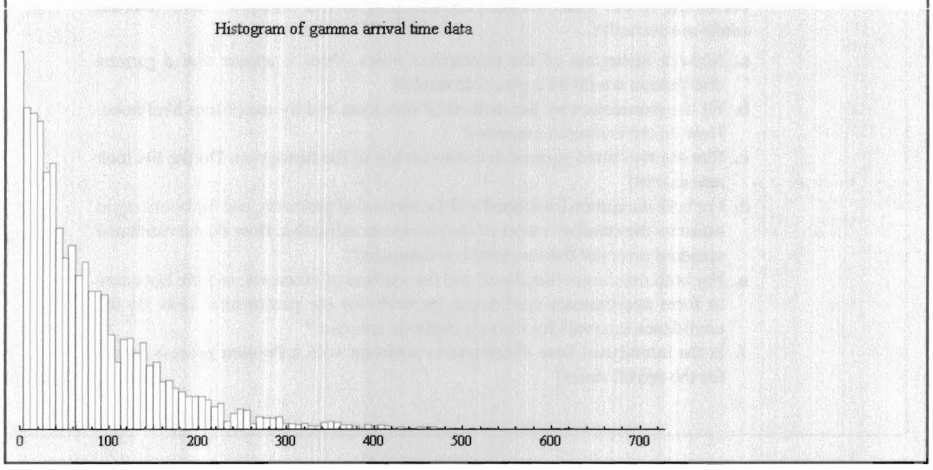
f. Is the interarrival time distribution consistent with a Poisson process model for the arrival times?

PART (A)

Yes. The following shows the histogram of the data, and a plot of a Gamma distribution with the shape parameter $\alpha = 1$ and scale parameter β set to the average of the data.

In[224]:=
```
nBins = 100;
gz = nmaMakeDensityHistogram[data, nBins];
histPlot = GeneralizedBarChart[gz, BarStyle → White, ImageSize → 450,
  PlotRange → {{Min[data], Max[data]}, All},
  PlotLabel → "Histogram of gamma arrival time data"]
```

Out[226]=



Let me plot a Gamma distribution with the mean arrival time of the data

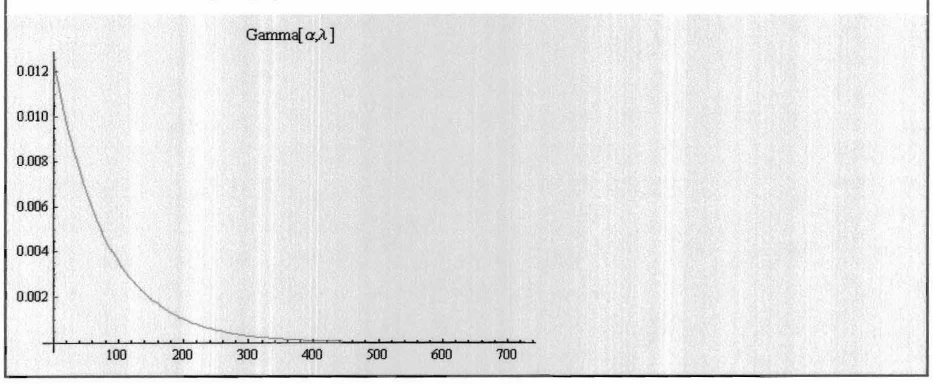
In[227]:=
```
λ = Mean[data]
```

Out[227]= 79.9352

In[228]:=
```
α = 1;
```

In[229]:=
```
Plot[PDF[GammaDistribution[α, λ], x], {x, 0, Max[data]}, PlotRange → All,
 PlotLabel → "Gamma[α,λ]"]
```

Out[229]=



From the above we see that a Gamma distribution is possible.

Part(b)

Using method of moments. We need 2 equations since we have to estimate 2 parameters α, λ. For Gamma

$$
\begin{aligned}
\mu_1 &= \frac{\alpha}{\lambda}\\
\mu_2 &= Var\left(X\right)+\left(E\left(X\right)\right)^2\\
&= \frac{\alpha}{\lambda^2}+\left(\frac{\alpha}{\lambda}\right)^2\\
&= \frac{\alpha}{\lambda^2}\left(\alpha+1\right)
\end{aligned}
$$

Now from the data itself, calculate the First and Second moments and equate to the above and solve for α, λ and these will be our estimate. This little code does the above

Notice, In *Mathematica*, it uses β as one of the parameters to the Gamma distributions, which is $\frac{1}{\lambda}$ of the definition in our textbook.

Start by finding the first and second moments of the Gamma distribution

In[235]:=
```
Clear[x, α, λ]
m1 = ExpectedValue[x, GammaDistribution[α, 1 / λ], x]
```
Out[236]=

$\frac{\alpha}{\lambda}$

In[237]:=
```
m2 = ExpectedValue[x^2, GammaDistribution[α, 1 / λ], x]
```
Out[237]=

$\frac{\alpha (1+\alpha)}{\lambda^2}$

Now find the first and second moments of the data itself

In[238]:=
```
m1ForData = Mean[data]
```
Out[238]=
```
79.9352
```

In[239]:=
```
m2ForData = Variance[data] + (Mean[data])^2
```
Out[239]=
```
12702.9
```

Now solve for α, λ

In[279]:=
```
sol = First@Solve[{m1 == m1ForData, m2 == m2ForData}, {α, λ}];
αByMoments = α /. sol
```
Out[280]=
```
1.01209
```

In[281]:=
```
λByMoments = λ /. sol
```
Out[281]=
```
0.0126614
```

Now using the MLE method. For α :

$$
\begin{aligned}
L(\alpha,\lambda) &= \prod_i^n \frac{\lambda^\alpha}{\Gamma(\alpha)} X_i^{\alpha-1} e^{-\lambda X_i} \\
l(\alpha,\lambda) &= \sum_i^n \log \frac{\lambda^\alpha}{\Gamma(\alpha)} X_i^{\alpha-1} e^{-\lambda X_i} \\
&= \sum_i^n \log \frac{\lambda^\alpha}{\Gamma(\alpha)} + (\alpha-1)\sum_i^n \log X_i - \lambda \sum_i^n X_i \\
&= n\alpha \log\lambda - n\log\Gamma(\alpha) + (\alpha-1)\sum_i^n \log X_i - \lambda\sum_i^n X_i
\end{aligned}
$$

Hence we obtain the 2 equations

$$
\begin{aligned}
\frac{\partial l(\alpha,\lambda)}{\partial\alpha} &= n\log\lambda - n\ polyGamma(0,\alpha) + \sum_i^n \log X_i \\
\frac{\partial l(\alpha,\lambda)}{\partial\lambda} &= n\alpha\frac{1}{\lambda} - \sum_i^n X_i
\end{aligned}
$$

From the second equation, set it to zero we obtain

$$
\hat{\lambda} = \frac{n\hat{\alpha}}{\sum_i^n X_i} = \frac{\hat{\alpha}}{\bar{X}}
$$

Substitute the above in the first equation and set to zero we obtain

$$
\begin{aligned}
0 &= n\log\left(\frac{\hat{\alpha}}{\bar{X}}\right) - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_i^n \log X_i \\
0 &= n\log(\hat{\alpha}) - n\log\bar{X} - n\ polyGamma(0,\alpha) + \sum_i^n \log X_i
\end{aligned}
$$

And solve for $\hat{\alpha}$. Once we find $\hat{\alpha}$ we then find also $\hat{\lambda} = \frac{\hat{\alpha}}{\bar{X}}$

11

This code below does the above.

Now find estimates of the parameters by MLE method and do the same as above

In[299]:=
```
Clear[α, λ]
n = Length[data]
```
Out[300]=
```
3935
```
In[301]:=
```
xBar = Mean[data]
```
Out[301]=
```
79.9352
```
In[302]:=
```
eq1 = n Log[α] - n Log[xBar] - n PolyGamma[0, α] + Sum[Log[data[[i]]], {i, 1, n}]
```
$$\texttt{eq1} = n\,\texttt{Log}[\alpha] - n\,\texttt{Log}[\texttt{xBar}] - n\,\texttt{PolyGamma}[0,\alpha] + \sum_{i=1}^{n}\texttt{Log}[\texttt{data}[[i]]]$$

Out[302]=
```
-2206.38 + 3935 Log[α] - 3935 PolyGamma[0, α]
```
Solve the above numerically using FindRoot

In[303]:=
```
FindRoot[eq1, {α, 1}];
αMLE = α /. %
```
Out[304]=
```
1.02633
```
Now that we found MLE for alpha, use it to find MLE for λ

In[305]:=
$$\texttt{λMLE} = \frac{\texttt{αMLE}}{\texttt{xBar}}$$

Out[305]=
```
0.0128395
```
Make a table to compare the α, λ found by the above 2 methods

In[306]:=
```
TableForm[{{αByMoments, λByMoments}, {αMLE, λMLE}},
 TableHeadings → {{"Method of moments", "MLE"}, {"α", "λ"}}]
```
Out[306]//TableForm=

	α	λ
Method of moments	1.01209	0.0126614
MLE	1.02633	0.0128395

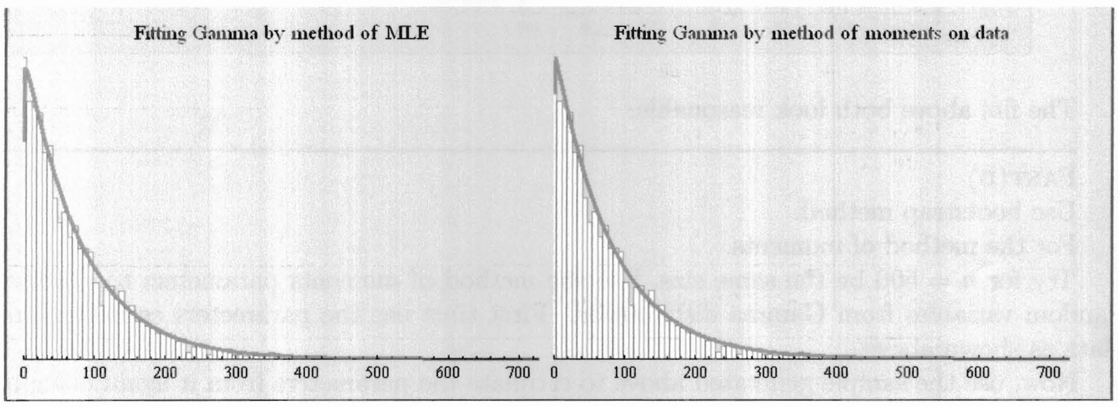
PART(C)

Now Fit this model again, and compare the MLE fitting to the method of moments fitting

In[350]:=
```
β = 1/λByMoments; (*to Make Mathematica convention of the Gamma function*)
pdfByMomentMethod = Plot[PDF[GammaDistribution[αByMoments, β], x],
   {x, 0, Max[data]}, PlotRange → All, PlotLabel → "Gamma[α,λ]",
   PlotStyle → {Red, Thick}];
p1 = Show[{histPlot, pdfByMomentMethod},
    PlotLabel → "Fitting Gamma by method of moments on data", ImageSize → 300];
```

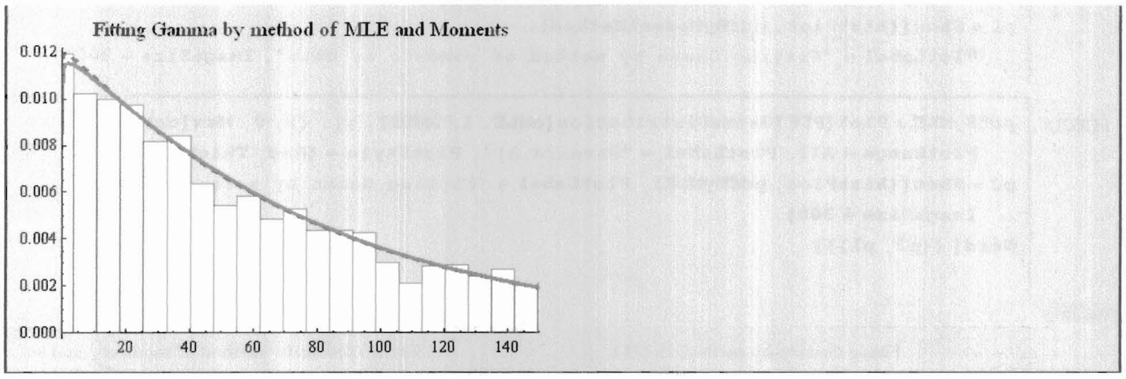
In[352]:=
```
pdfByMLE = Plot[PDF[GammaDistribution[αMLE, 1 / λMLE], x], {x, 0, Max[data]},
    PlotRange → All, PlotLabel → "Gamma[α,λ]", PlotStyle → {Red, Thick}];
p2 = Show[{histPlot, pdfByMLE}, PlotLabel → "Fitting Gamma by method of MLE",
    ImageSize → 300];
Grid[ {{p2, p1}}]
```

Out[354]=



This plot shows more closely the fitting on top of each others. They are very close so hard to see the difference other than near the high frequency part.

In[554]:=
```
pdfByMomentMethod = Plot[PDF[GammaDistribution[αByMoments, β], x],
   {x, 0, Max[data]}, PlotRange → {{0, 300}, All}, PlotLabel → "Gamma[α,λ]",
   PlotStyle → {Dashed, Thick}];
pdfByMLE = Plot[PDF[GammaDistribution[αMLE, 1 / λMLE], x], {x, 0, Max[data]},
   PlotRange → {{0, 300}, All}, PlotLabel → "Gamma[α,λ]", PlotStyle → {Red, Thick}];
p2 = Show[{histPlot, pdfByMomentMethod, pdfByMLE},
  PlotLabel → "Fitting Gamma by method of MLE and Moments", ImageSize → 300,
  PlotRange → {{0, 150}, All}]
```
Out[556]=



The fits above both look reasonable.

PART(D)

Use bootstrap method.

For the method of moments.

Try for $n = 500$ be the same size. Use the method of moments parameters to generate an n random variables from Gamma distribution. First time use the parameters estimated from the data as shown above.

Now, use the sample generated above to estimate the parameters from it again using also the method of moments. Use these parameters to generate another n random variables. repeat this process for say $N = 5000$ and find the variances of the parameters α, λ, and hence we find the standard error which is the square root of these variances.

Here is the code to do the above and the result

(Last minute update), I am getting large result for standard error from the bootstrap method. I think I have something wrong. Here is the result I get and the code

For Method of moments, I get standard error for alpha=918 and for lambda=18

Part (d)

In[1]:=
```
Remove["Global`*"]

SeedRandom[01010101];
m1 = ExpectedValue[x, GammaDistribution[α, β], x];
m2 = ExpectedValue[x^2, GammaDistribution[α, β], x];

getMethodOfMomentsParameters[data_] := Module[{sol, αByMoments, βByMoments},
  m1ForData = Mean[data];
  m2ForData = Variance[data] + (m1ForData)^2;
  sol = First@Solve[{m1 == m1ForData, m2 == m2ForData}, {α, β}];
  αByMoments = α /. sol;
  βByMoments = β /. sol;

  {αByMoments, βByMoments}
 ]
```

In[6]:=
```
file =
  "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";
data = Flatten[Import[file, "Table"]];

n = 500; (*sample size*)
nIter = 5000; (*number of iterations*)
alpha = Table[0, {i, nIter}];
beta = Table[0, {i, nIter}];

{alpha[[1]], beta[[1]]} = getMethodOfMomentsParameters[data];

For[i = 2, i ≤ nIter, i++,
 {
  sample = RandomReal[GammaDistribution[alpha[[i - 1]], beta[[i - 1]]], n];
  {alpha[[i]], beta[[i]]} = getMethodOfMomentsParameters[sample];
 }
]
Print["Standard error for alpha=", Sqrt[Variance[alpha]]]
Print["Standard error for lambda=", Sqrt[Variance[1 / beta]]]
```

```
Standard error for alpha=918.308
Standard error for lambda=18.7966
```

For MLE I get
Standard error for alpha=1.68697*10^8
Standard error for lambda=60.2585

Now do the same for MLE method

In[1]:=
```
Remove["Global`*"]

SeedRandom[01 010 101];

getMLEParameters[data_] := Module[{sol, xBar, αMLE, λMLE, eq, α, n},
  xBar = Mean[data];
  n = Length[data];
  eq = n Log[α] - n Log[xBar] - n PolyGamma[0, α] + Sum[Log[data[[i]]], {i, 1, n}];
  sol = FindRoot[eq, {α, 1}];
  αMLE = α /. sol;
  λMLE = αMLE / xBar;

  {αMLE, λMLE}
 ]
```

Remove::rmnsm : There are no symbols matching "Global`*". ≫

In[4]:=
```
file =
  "E:/nabbasi/data/nabbasi_web_Page/my_courses/FULLERTON_COURSES/Fall_2007/math_502
    _probability_and_statistics/quiz/quiz7/gamma-arrivals.txt";
data = Flatten[Import[file, "Table"]];

n = 500; (*sample size*)
nIter = 5000; (*number of iterations*)
alpha = Table[0, {i, nIter}];
lambda = Table[0, {i, nIter}];

{alpha[[1]], lambda[[1]]} = getMLEParameters[data];

For[i = 2, i ≤ nIter, i ++,
 {
  sample = RandomReal[GammaDistribution[alpha[[i - 1]], lambda[[i - 1]]], n];
  {alpha[[i]], lambda[[i]]} = getMLEParameters[sample];
 }
]
Print["Standard error for alpha=", Sqrt[Variance[alpha]]]
Print["Standard error for lambda=", Sqrt[Variance[lambda]]]
```

Part (e) and (f)
sorry, run out of time.

Nasser I looked at the solⁿ that you e-mailed me, and the CI and the standard errors were not correct!

−4

4.8 Quizes key solution

QUIZ 1 MATH 502AB Fall 2007

Name (please print) KEY

1. Consider a sequence of days, and let R_i denote the event that it rains on day i. Let $P(R_0) = p$ (rain today), $P(R_i|R_{i-1}) = \alpha$, and $P(R_i^c|R_{i-1}^c) = \beta$. Suppose further that only today's weather is relevant to predicting tomorrow's; that is, $P(R_i|R_{i-1} \cap \cdots, \cap R_0) = P(R_i|R_{i-1})$. What is the probability that it rains n days from now? What happens as n approaches infinity?

$$P(R_n) = P(R_n|R_{n-1})P(R_{n-1}) + P(R_n|R_{n-1}^c)P(R_{n-1}^c)$$
$$= \underbrace{\cancel{P(R_n|R_{n-1})}}_{\alpha} P(R_{n-1}) + (1-\beta)[1 - P(R_{n-1})]$$
$$= (\alpha+\beta-1)P(R_{n-1}) + (1-\beta).$$

Let $a = \alpha+\beta-1 \quad b = (1-\beta) \quad p_n = P(R_n)$.

Then
$$p_n = a p_{n-1} + b = a(ap_{n-2}+b)+b$$
$$= a^2 p_{n-2} + b(1+a) = a^2(ap_{n-3}+b) + b(1+a)$$
$$= a^3 p_{n-3} + b(1+a+a^2)$$
$$= \cdots$$
$$= a^n p_0 + b(1+a+a^2+\cdots+a^{n-1})$$
$$= a^n p_0 + b\frac{1-a^n}{1-a}$$

as $n \to \infty \quad a^n \to 0 \Rightarrow p_n \to \dfrac{b}{1-a} = \dfrac{1-\beta}{-\alpha-\beta+2} = \dfrac{P(R_i|R_{i-1}^c)}{P(R_i|R_{i-1}^c) + P(R_i^c|R_{i-1})}$

2. Suppose that a rare disease has an incidence of 1 in 1000. Assuming that members of the population are affected independently, Find the probability that two individuals are affected in a population of 100,000. Give a numerical value.

2 Show that if the conditional probabilities exist, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \overset{P(A_2|A_1)}{\cancel{P(A_2|A_1 \cap A_2)}} \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

~~We have $P(A_1 \cap A_2) = P(A_1|A_2) \cdot P(A_2)$~~

For $n=1$, the equality holds obviously.

Assume true for $n = k-1$, and let $B = A_1 \cap A_2 \cap \cdots \cap A_{k-1}$

Then $P(B) = P(A_1) \overset{P(A_2|A_1)}{\cancel{P(A_2|A_1 \cap A_2)}} \cdots P(A_{k-1}|A_1 \cap \cdots \cap A_{k-2})$

$P(A_1 \cap \cdots \cap A_k) = P(B \cap A_k) = \underset{P(A_k|B) \cdot P(B)}{\cancel{P(B|A_k) \cdot P(A_k)}}$

$= \underbrace{P(A_1) \overset{P(A_2|A_1)}{\cancel{P(A_2|A_1 \cap A_2)}} \cdots P(A_{k-1}|A_1 \cap \cdots \cap A_{k-2})}_{P(B) \text{ by induction}} P(A_k|A_1 \cap \cdots \cap A_{k-1})$

Q.E.D.

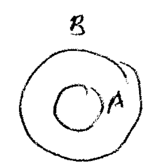
3. Let A and B be arbitrary events. Use the three axioms of probability to show that

$$P(A \cup B) \leq P(A) + P(B).$$

Identify the axiom(s) that you use at each step. You are not allowed to use any theorems.

First we prove that

if $A \subset B$, then $P(A) \leq P(B)$

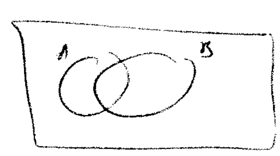


Proof: $B = A \cup [B \cap A^c]$

$\Rightarrow P(B) = P\{A \cup [B \cap A^c]\}$

$= P(A) + P(B \cap A^c)$ By Axiom 3

$\geq P(A)$ by axiom (2) $\boxed{}$



Now $A \cup B = A \cup [B \cap A^c]$

$P(A \cup B) = P(A) + P[B \cap A^c]$ 1st Axiom

$\leq P(A) + P(B)$

$B \cap A^c \subset B$

$P(B \cap A^c) \leq P(B)$

4. Let $X \sim binomial(n, p)$. Derive the mode of the probability mass function of X.

$$p_k = P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\frac{p_k}{p_{k-1}} = \frac{n!}{k!(n-k)!} \cdot \frac{(k-1)!(n-k+1)!}{n!} \cdot \frac{p^k}{p^{k-1}} \cdot \frac{(1-p)^{n-k}}{(1-p)^{n-k+1}}$$

$$= \frac{(n-k+1)}{k} \frac{p}{(1-p)}$$

$$\frac{p_k}{p_{k-1}} \geq 1 \iff np - kp + p \geq k - kp$$

$$\iff k \leq (n+1)p$$

$\therefore$ the mode is at the greatest integer less than or equal to $(n+1)p$

5. Suppose that a rare disease has an incidence of 1 in 1000. Assuming that members of the population are affected independently, Find the probability that two individuals are affected in a population of 100,000. Give a numerical value.

$$P(D) = \frac{1}{1000}$$

$X = \#$ of affected $\sim$ Binomial$(100{,}000, \frac{1}{1000})$

$$P(X=2) = \binom{100{,}000}{100}\left(\frac{1}{1000}\right)^{100}\left(\frac{999}{1000}\right)^{100{,}000-100} = .03988694$$

Poisson approximation: $\lambda = 100{,}000\left(\frac{1}{1000}\right) = 100$

$$P(X=2) = e^{-100}\frac{100^{100}}{100!} = .039861$$

QUIZ 2 MATH 502AB Fall 2007

Name (please print) KEY

1. Use the fact that $\Gamma(1/2) = \sqrt{\pi}$ to show that if n is an odd integer, then

$$\Gamma(n/2) = \frac{\sqrt{\pi}(n-1)!}{2^{n-1}\left(\frac{n-1}{2}\right)!}. \qquad (*)$$

Obviously true for $n=1$.

Suppose $(*)$ is true for $n = 2k-1$ (i.e. $\Gamma\left(\frac{2k-1}{2}\right) = \frac{\sqrt{\pi}(2k-2)!}{2^{2k-2}\left(\frac{2k-2}{2}\right)!}$.

Then

$$\Gamma\left(\frac{2k+1}{2}\right) = \Gamma\left(\frac{2k-1}{2}+1\right) = \left(\frac{2k-1}{2}\right)\Gamma\left(\frac{2k-1}{2}\right) \overset{\text{By induction hyp.}}{=} \left(\frac{2k-1}{2}\right)\frac{\sqrt{\pi}(2k-2)!}{2^{2k-2}(k-1)!}$$

$$= \frac{2k}{2k}\,\frac{(2k-1)(2k-2)!\,\sqrt{\pi}}{2^{2k-1}(k-1)!} = \frac{\sqrt{\pi}(2k)!}{2^{2k}k!}$$

2. If $U \sim Uniform[-1,1]$, find the density of U^2.

$f_U(u) = \frac{1}{2} \quad -1<u<1 \qquad$ Let $Z = U^2$

for $0 \le z \le 1$

$$F_Z(z) = P(Z \le z) = P(U^2 \le z)$$

$$= P(-\sqrt{z} \le U \le \sqrt{z})$$

$$= F_U(\sqrt{z}) - F_U(-\sqrt{z})$$

$$= \int_{-1}^{\sqrt{z}} \frac{1}{2}\,du - \int_{-1}^{-\sqrt{z}} \frac{1}{2}\,du$$

$$= \frac{1}{2}(\sqrt{z}+1) - \frac{1}{2}(-\sqrt{z}+1) = \sqrt{z}$$

$$\therefore\ f_Z(z) = \begin{cases} \frac{1}{2} z^{-1/2} & 0 \le z \le 1 \\ 0 & \text{o.w.} \end{cases}$$

3. The following ~~five~~ four numbers were randomly generated from the uniform random variable on (0,1):

0.0153 0.7468 0.4451 0.9318 ~~0.4959~~

Using these numbers generate five random numbers from the geometric random variable with parameter $p = 1/3$. Very briefly explain how you obtain your solution.

The cdf for the geometric random variable with parameter $\frac{1}{3}$ is given by

$$F(x) = \sum_{k=1}^{x} \left(\frac{2}{3}\right)^{k-1}\left(\frac{1}{3}\right) = \frac{1}{3}\sum_{k=0}^{x-1}\left(\frac{2}{3}\right)^k = \left(\frac{1}{3}\right)\frac{\left(1-\left(\frac{2}{3}\right)^x\right)}{1-\left(\frac{2}{3}\right)} = 1-\left(\frac{2}{3}\right)^x \quad x=1,2,\ldots$$

$$\Rightarrow F^{-1}(x) = \frac{\log(1-x)}{\log(2/3)}.$$

If we apply F^{-1} to the above values we obtain (in order)

0.038, 3.3875, 1.4526, 6.6233

which implies the following generated values:

1, 4, 2, 7

4. Three players play 10 independent rounds of a game, and each player has probability 1/3 of winning each round. **(a)** Find the joint distribution of the numbers of games won by each of the three players. **(b)** Identify the distribution of the number of games won by player one.

(a) N_i = # of games won by player i $\quad i=1,2,3$

$$p(n_1,n_2,n_3) = \binom{10}{n_1,n_2,n_3}\left(\frac{1}{3}\right)^{n_1}\left(\frac{1}{3}\right)^{n_2}\left(\frac{1}{3}\right)^{n_3}$$

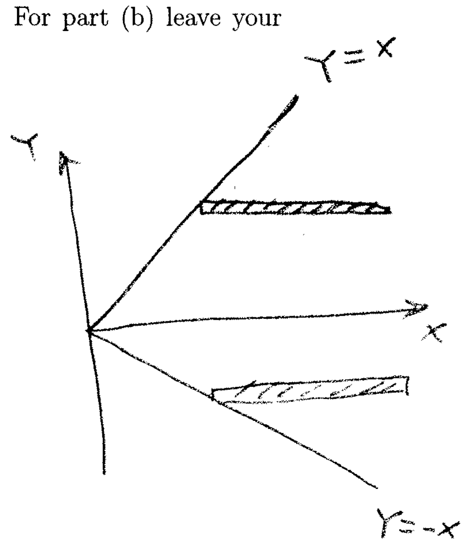
$$n_1+n_2+n_3 = 10$$

(b) The distribution of $N_1 \sim$ Binomial$(10, \frac{1}{3})$.

5. Let (X,Y) be jointly distributed random variables with pdf

$$f(x,y) = \frac{1}{8}(x^2 - y^2)e^{-x} \quad 0 \le x \le \infty \quad -x \le y \le x.$$

(a) Find the marginal density of Y. (b) Find $P(X+Y \le 1)$. For part (b) leave your solution as integrals, and do not calculate the integrals.



a)

If $y>0$

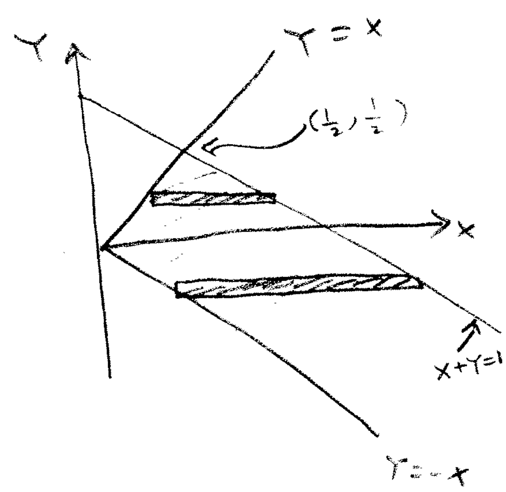
$$f_Y(y) = \int_y^\infty \frac{1}{8}(x^2-y^2)e^{-x}\,dx$$

$$= \frac{1}{4}e^{-y}(y+1)$$

If $y<0$

$$f_Y(y) = \int_{-y}^\infty \frac{1}{8}(x^2-y^2)e^{-x}\,dx$$

$$= \frac{1}{4}e^{y}(1-y) \qquad y<0$$



b)

$$P(X+Y\le 1) = \int_0^{1/2}\int_y^{1-y} \frac{1}{8}(x^2-y^2)e^{-x}\,dx\,dy$$

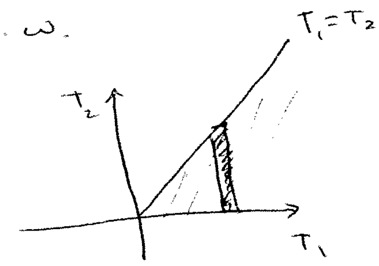
$$+ \int_{-\infty}^0 \int_{-y}^{1-y} \frac{1}{8}(x^2-y^2)e^{-x}\,dx\,dy$$

Quiz 3 Math 502AB Fall 2007

Name (please print) KEY

1. Suppose that two components have independent exponentially distributed lifetimes T_1 and T_2, with parameters α and β, respectively. Find (a) $P(T_1 > T_2)$, (b) determine the distribution of $W = 2T_2$, and (c) use the results in parts (a) and (b) to obtain $P(T_1 > 2T_2)$.

$$f_{T_1,T_2}(t_1,t_2) = \begin{cases} \alpha\beta e^{-\alpha t_1 - \beta t_2} & t_1 \geq 0 \quad t_2 \geq 0 \\ 0 & \text{o.w.} \end{cases}$$



a)

$$P(T_1 > T_2) = \int_0^\infty \int_0^{t_1} \alpha\beta e^{-\alpha t_1 - \beta t_2}\, dt_2\, dt_1$$

$$= \frac{\beta}{\alpha+\beta}$$

b)

$$F_W(w) = P(2T_2 \leq w) = P\left(T_2 \leq \frac{w}{2}\right)$$

$$= \int_0^{\frac{w}{2}} \lambda_2 e^{-\lambda_2 t_2}\, dt_2 = 1 - e^{-\frac{\lambda_2}{2} w}$$

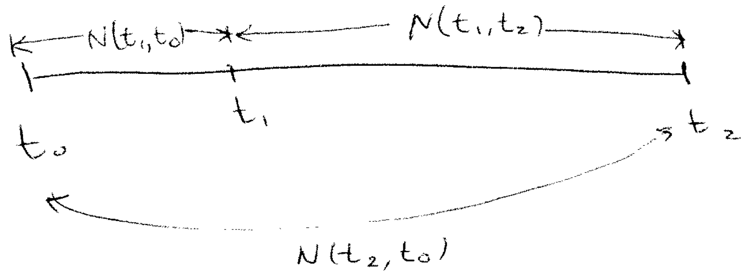
$$\therefore\ W \sim \exp\left(\frac{\lambda_2}{2}\right)$$

c)

$$P(T_1 > T_2) = \frac{\beta}{\alpha+\beta} \quad \text{when } T_1 \sim \exp(\lambda_1) \ \& \ T_2 \sim \exp(\lambda_2)$$

$$P(T_1 > 2T_2) = P(T_1 > W) = \frac{(\lambda_2/2)}{\alpha + (\lambda_2/2)} = \frac{\lambda_2}{2\alpha + \lambda_2}$$

2. Consider a Poisson process on the real line, and denote by $N(t_1, t_2)$ the number of events in the interval (t_1, t_2). If $t_0 < t_1 < t_2$, find the conditional distribution of $N(t_0, t_1)$ given that $N(t_0, t_2) = n$.



$$P(N(t_0,t_1)=k \mid N(t_0,t_2)=n) =$$

$$= \frac{P\{N(t_0,t_1)=k \cap N(t_0,t_2)=n\}}{P\{N(t_0,t_2)=n\}}$$

$$= \frac{P\{N(t_0,t_1)=k \cap N(t_1,t_2)=n-k\}}{P\{N(t_0,t_2)=n\}}$$

$$= \frac{P\{N(t_0,t_1)=k\}\, P\{N(t_1,t_2)=n-k\}}{P\{N(t_0,t_2)=n\}}$$

$$= \frac{\dfrac{[(t_1-t_0)\lambda]^k e^{-(t_1-t_0)\lambda}}{k!} \cdot \dfrac{[(t_2-t_1)\lambda]^{n-k} e^{-(t_2-t_1)\lambda}}{(n-k)!}}{\dfrac{[(t_2-t_0)\lambda]^n e^{-(t_2-t_0)\lambda}}{n!}}$$

$$= \binom{n}{k}\left(\frac{t_1-t_0}{t_2-t_0}\right)^k \left(\frac{t_2-t_1}{t_2-t_0}\right)^{n-k} \quad k=0,1,\dots,n$$

$$\sim \mathrm{Binom}\left(n, \frac{t_1-t_0}{t_2-t_0}\right)$$

3. Suppose that the probability Θ of getting heads for a coin is unknown, and let the prior opinion about Θ be represented by the uniform distribution on [0,1]. You spin the coin repeatedly and record the number of times N until a heads comes up. (a) Find the posterior density of Θ given N. (b) Use Matlab or any other software to plot the posterior for cases where $N = 1$, $N = 2$, and $N = 6$. Using your plots, explain what you infer about the probability of heads in each circumstance.

$\Theta \sim \text{unif}[0,1]$

$N \mid \Theta \sim \text{geometric}(p)$

$$f_{N,\Theta}(n,\theta) = (1-\theta)^{n-1}\theta; \quad n = 1, 2, \cdots, \quad 0 \le \theta \le 1$$

$$f_{\Theta|N}(\theta|n) = \frac{f_{\Theta,N}(\theta,n)}{f_N(n)} = \frac{\theta(1-\theta)^{n-1}}{\int_0^1 \theta(1-\theta)^{n-1}\,d\theta}$$

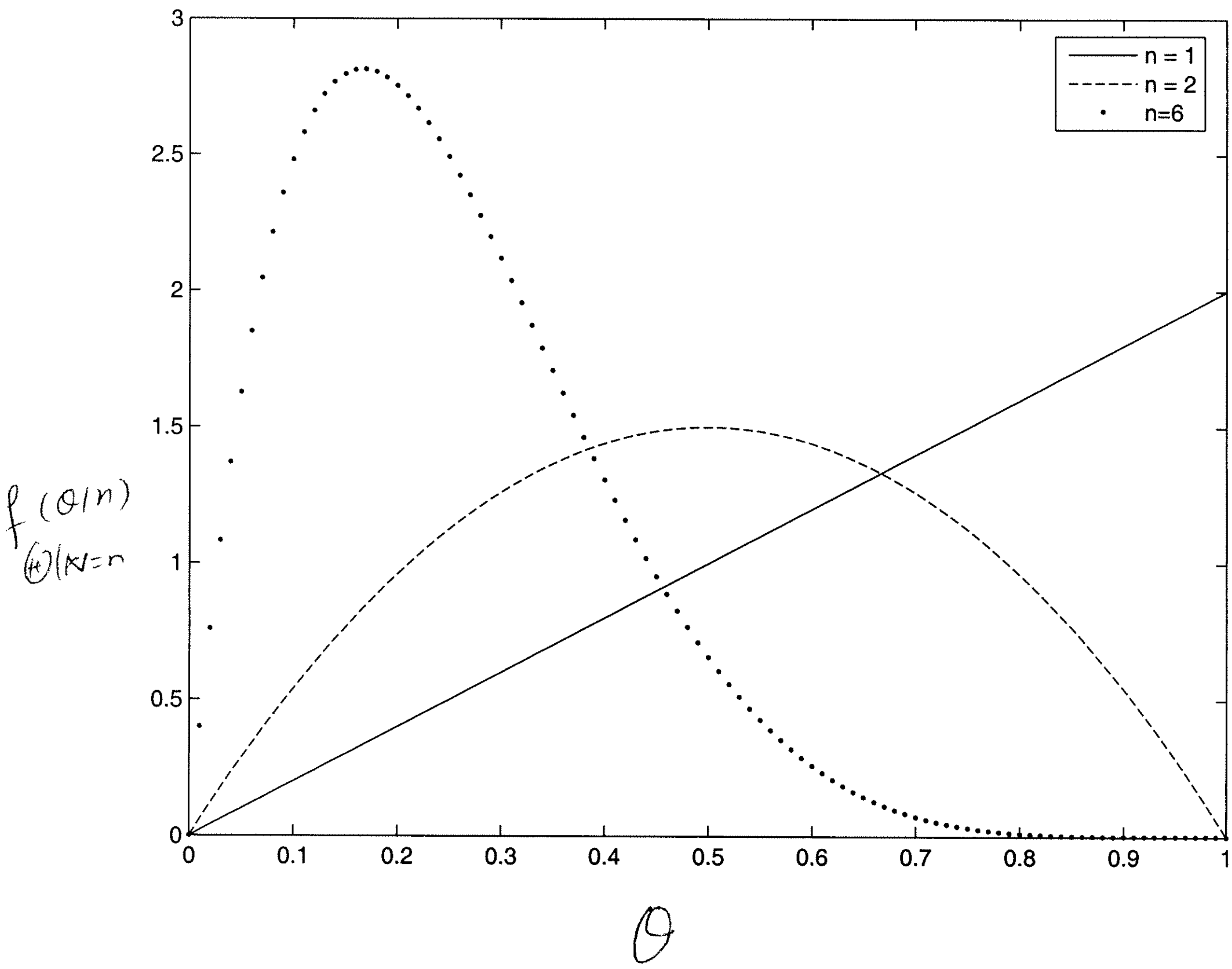
$$= \frac{\theta(1-\theta)^{n-1}}{1/[n(n+1)]} = n(n+1)\,\theta(1-\theta)^{n-1}$$

$\Theta \mid N=n \sim \text{Beta}(a=2, b=n)$

~~The larger n, our prediction of probab~~

Larger n's indicate lower probabilities for getting heads (i.e. smaller θ).

It is clear from the graph (next page) that as n increases, the density gets more concentrated about values closer to zero.


n = 1
n = 2
n=6
0
0.5
1
1.5
2
2.5
3
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
f (θ|n)
Θ|N=n
θ

QUIZ 4 MATH 502AB Fall 2007

Name (please print) Key

1. Let X be a continuous random variable with a pdf that is symmetric about a point ξ. Provided that $E(X)$ exists, show that $E(X) = \xi$.

Since $\xi = \int_{-\infty}^{\infty} f(x)\,dx$, it is sufficient to show that $\int_{-\infty}^{\infty} (x-\xi)\, f(x) = 0$

Let $u = x-\xi$, $x = u+\xi$

$\int_{-\infty}^{\infty} (x-\xi) f(x)\,dx = \int_{-\infty}^{\infty} u f(u+\xi)\,du = \int_{-\infty}^{0} u f(u+\xi)\,du + \int_{0}^{\infty} u f(u+\xi)\,du = \int_{0}^{\infty} u f(u+\xi)\,du - \int_{0}^{\infty} u f(\xi-u)\,du$

But since $f(u+\xi) = f(\xi-u)$ by symmetry, the last expression equals zero.

2. Let X be an exponential random variable with parameter λ. Find

$$P\left[|X - E(X)| > \frac{2}{\lambda}\right]$$

and compare your result to the Chebyshev's bound.

$X \sim \exp(\lambda)$ $E(X) = \frac{1}{\lambda}$

$P(|X-\frac{1}{\lambda}| > \frac{2}{\lambda}) = 1 - P(|X-\frac{1}{\lambda}| \le \frac{2}{\lambda})$

$= 1 - P(-\frac{1}{\lambda} \le X \le \frac{3}{\lambda}) = \int_{0}^{3/\lambda} \lambda e^{-\lambda x}\,dx = e^{-3} = .049787$

$P(|X-E(X)| > \frac{2}{\lambda}) \cong .95$

By C.C. it is at least 0.75

3. If X is a discrete random variable, taking values on the positive integers, then show that $E(X) = \sum_{k=1}^{\infty} P(X \geq k)$.

$P(X=1) + P(X=2) + P(X=3) + \cdots$

$ P(X=2) + P(X=3) + \cdots$

$ + P(X=3) + \cdots$

$P(X=1) + 2P(X=2) + 3P(X=3) + \cdots$

$E(X) = \sum_{\text{all } x} x P(X=x)$

4. Find the mean of a negative binomial random variable X with parameters r and p, by expressing X as sum of indicator variables.

X_i = # of trials until a success after i^{th} success has been attained.

$X = X_1 + X_2 + \cdots + X_r$

$X_i \sim \text{geometric}(p)$

$E(X_i) = \frac{1}{p}$

$E(X) = \frac{r}{p}$

5. If $U = a + bX$ and $V = c + dY$, show that $|\rho_{UV}| = |\rho_{XY}|$.

$$\rho_{UV} = \frac{Cov(U,V)}{\sqrt{Var(U)\ Var(V)}}$$

$$Cov(U,V) = Cov(a+bX, c+dY)$$

$$= bd\ Cov(X,Y)$$

$$Var(U) = b^2\ Var(X)$$

$$Var(V) = d^2\ Var(Y)$$

Plug-in these values and the result follows.

QUIZ 5 MATH 502AB October 20, 2006

Name (please print)______________________________

1. The moment generating function for a random variable X having a χ^2 distribution with degrees of freedom $n \geq 1$ is $M_X(t) = (1-2t)^{-n/2}$. Let W have a χ^2 distribution with degrees of freedom $n > 1$, and let V have a χ^2 distribution with degrees of freedom 1. (a). If $W = U + V$, and U and V are independent, determine the distribution of U? (b). What are the mean and variance of W?

a) $M_W(t) = M_U(t) \cdot M_V(t) \Rightarrow M_U(t) = \dfrac{M_W(t)}{M_V(t)}$

$\Rightarrow M_V(t) = \left(\frac{1}{1-2t}\right)^{n/2} / \left(\frac{1}{1-2t}\right)^{1/2} = \left(\frac{1}{1-2t}\right)^{\frac{n-1}{2}}$

$\Rightarrow V \sim \chi^2_{(n-1)}$

$E(W) = n$
$\mathrm{var}(W) = 2n$ $\Big\}$ Can be obtained from the table, using the $\chi^2_{(n)} \sim \mathrm{gamma}\left(\frac{n}{2}, \frac{1}{2}\right)$

2. Find the approximate variance of $Y = \sqrt{X}$, where X is a Poisson random variable with parameter λ.

$\mathrm{Var}(\cancel{X}) = E(\cancel{X}) - [E(\sqrt{x})]^2 = \lambda - [E(\sqrt{x})]^2$

since $E(x) = \lambda$, and $\mathrm{var}(x) = \lambda$, we have

$E(\sqrt{x}) \approx \sqrt{\lambda} + \frac{1}{2}\lambda g''(\lambda)$,

where $g(x) = \sqrt{x} \Rightarrow g'(x) = \frac{1}{2}x^{-1/2} \Rightarrow g''(x) = -\frac{1}{4}x^{-3/2}$

$\Rightarrow E(\sqrt{x}) \approx \sqrt{\lambda} + \frac{1}{2}\lambda\left[-\frac{1}{4}\lambda^{-3/2}\right] = \lambda^{1/2} - \frac{1}{8}\lambda^{-1/2}$

$\mathrm{Var}(\sqrt{x}) \approx \lambda - \left(\lambda^{1/2} - \frac{1}{8}\lambda^{-1/2}\right)^2$

3. The random variable Y has a Gamma distribution with parameters α and λ. Furthermore, assume that X given Y has a Poisson distribution with parameter Y^2. (a) Obtain $E(X)$. (b) Obtain $Var(X)$.

$$Y \sim \Gamma(\alpha, \lambda)$$

$$X|Y \sim \text{Poisson}(Y^2)$$

$$E(X) = E\{E(X|Y)\}$$

$$= E(Y^2)$$

$$= Var(Y) + (E(Y))^2 = \frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{\lambda}\right)^2$$

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$

$$= E(Y^2) + Var(Y^2)$$

$$= E(Y^2) + E(Y^4) - [E(Y^2)]^2$$

$$E(Y^2) = \frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{\lambda}\right)^2$$

$$E(Y^4) = \int_0^\infty y^4 \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \underbrace{\int_0^\infty y^{\alpha+3} e^{-\lambda y} dy}_{\Gamma(\alpha+4, \lambda)}$$

$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+4)}{\lambda^{\alpha+4}} = \frac{(\alpha+3)(\alpha+2)(\alpha+1)\alpha}{\lambda^4}$$

$$Var(X) = \frac{(\alpha+3)(\alpha+2)(\alpha+1)\alpha}{\lambda^4} + \frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{2}\right)^2$$

QUIZ 6 MATH 502AB October 30, 2006

Name (please print)______________________________

In solving the problems below, you can use all the results that we have derived in class. You do not need to re-derive results. Make sure to cite the results that you use.

1. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(\mu, \sigma^2)$, and S^2 be the sample variance. What is $\mathrm{Var}(S^2)$?

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \Rightarrow Var\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$$

$$\Rightarrow \frac{(n-1)^2}{\sigma^4} Var(S^2) = 2(n-1)$$

$$\Rightarrow Var(S^2) = \frac{2(n-1)\sigma^4}{(n-1)^2} = \frac{2\sigma^4}{(n-1)}$$

2. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(0, 1)$. Determine the asymptotic distribution of

$$(1/n)\sum_{i=1}^{n}|X_i|.$$

$$E(|X_i|) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|x|e^{-\frac{1}{2}x^2}dx = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} x e^{-\frac{1}{2}x^2}dx = \sqrt{\frac{2}{\pi}}$$

$$E(|X_i|^2) = E(X_i^2) = Var(X_i) + (E(X_i))^2 = 1$$

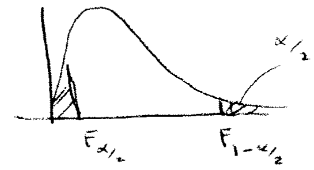
$$Var(|X_i|) = 1 - \frac{2}{\pi} = \frac{\pi-2}{\pi}$$

By Central limit theorem

$$\frac{\frac{1}{n}\sum_{i=1}^{n}|X_i| - \sqrt{2/\pi}}{\sqrt{(\pi-2)/n\pi}} \longrightarrow N(0,1)$$

3.. Let $X_1, \cdots, X_n$ be iid random variables from a $\mathcal{N}(0, \sigma_1^2)$ and $Y_1, \cdots, Y_n$ be iid random variables from a $\mathcal{N}(0, \sigma_2^2)$. Write a 95% confidence interval for σ_1^2/σ_2^2.

$$\frac{(n-1)S_X^2}{\sigma_1^2} \sim \chi^2_{(n-1)} \qquad \frac{(n-1)S_Y^2}{\sigma_2^2} \sim \chi^2_{(n-1)}$$



$$\Rightarrow \frac{\sigma_1^2 S_Y^2}{\sigma_2^2 S_X^2} \sim F_{(n,n)} \Rightarrow P\left[F_{\alpha/2,(n,n)} < \frac{\sigma_1^2 S_Y^2}{\sigma_2^2 S_X^2} < F_{1-\alpha/2\,(n,n)} \right] = 1-\alpha$$

$$\Rightarrow 95\%\ CI: \left[F_{\alpha/2;(n,n)} \cdot \frac{S_X^2}{S_Y^2},\ F_{1-\alpha/2;(n,n)} \frac{S_X^2}{S_Y^2} \right]$$

4. Let $X \sim N(0, 2)$ and $Y \sim exponential(1)$. Provided that X is independent of Y, identify the distribution of $X/\sqrt{Y}$.

$$X/\sqrt{2} \sim N(0,1)$$

$$Y/2 \sim \chi^2(2)$$

$$\frac{(X/\sqrt{2})}{\sqrt{\left(\frac{Y}{2}\right)/2}} \sim t(2)$$

$$\sqrt{2}\,\frac{X}{\sqrt{Y}} \sim t(2)$$

$\therefore \frac{X}{\sqrt{Y}}$ is a multiple of t with 2 degrees of freedom.

Chapter 5

Appendix

Local contents

5.1 My cheat sheet

chiSquare χ^2 distribution with degree $n = X_1^2 + X_2^2 + \cdots + X_n^2$ where X_i is standard $N(0,1)$

Standard Normal $Z = \frac{X-\mu}{\sigma}$ where X is Normal with mean μ and Variance σ^2.

$F(\chi^2_{(n)} < 0) = 0$

$E(\chi^2_{(n)}) = n$. $Var(\chi^2_{(n)}) = 2n$

$$\chi^2_{(n)} = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$$

$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt$

$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, $\Gamma(1) = 1$

For $n=1$, $\chi^2_{(1)} = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}}$; $\chi^2_{(2)} = \frac{1}{2} e^{-\frac{x}{2}}$

For nonnegative integer $n = 0, 1, 2, \ldots$ $\Gamma(n) = n!$

$\Gamma(n+1) = n\Gamma(n)$

Normal distribution $f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ or $= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Standard Normal $N(0,1) = \boxed{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}$.

derive $\chi^2_{(1)}$. let $X = Z^2$ where $Z \sim N(0,1)$.

$F_X(x) = P(X < x) = P(Z^2 < x) = P(-\sqrt{x} < Z < \sqrt{x})$

$F_X(x) = P(-\sqrt{x} < Z) + P(Z < \sqrt{x})$

$-\sqrt{x}$

$\Rightarrow \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$

$\sqrt{x}$

$f_X(x) = f_Z(\sqrt{x}) \cdot \frac{1}{2}\frac{1}{\sqrt{x}} + f_Z(-\sqrt{x}) \frac{1}{2}\frac{1}{\sqrt{x}}$

but $f_Z = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x)^2}$. let $u = \sqrt{x}$, $u = \sqrt{x}$ in above.

Done you'll get $\chi^2_{(1)}$ formula

CLT Sum and mean, Standardized, Converges to $N(0,1)$

$$Z_n = \frac{\sum X_i - n\mu_x}{\sqrt{Var X}\,\sqrt{n}}$$

each X has same Var.

Standardize Sum.

use this.

Standardizing: $Z_n = \frac{X_n - E(X_n)}{\sqrt{Var(X_n)}} \Rightarrow$ Z_n has 0 mean, and Var 1.

apply this to $\sum X_i$ and $\bar{X}_n$

E(X) any X.

$$Z_n = \frac{\bar{X}_n - \mu_x}{\frac{\sqrt{Var(X)}}{\sqrt{n}}}$$

Standardize mean

Chi Square $(\chi^2_{(n)})$.

definition: $\chi^2_{(n)} = \underbrace{Z^2(0,1) + Z^2(0,1) + \cdots + Z^2(0,1)}_{n \text{ times}}$

$$\chi^2_{(n)} = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}. \quad \chi^2_{(1)} = \frac{x^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{x}{2}}. \quad \chi^2_{(2)} = \frac{\frac{1}{2}}{\Gamma(1)_{=1}} x^{\frac{1}{2}} e^{-\frac{x}{2}}$$

$E(\chi^2_{(n)}) = n$, $Var(\chi^2_{(n)}) = 2n$

Some integrals: $\int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-x^2} dx = 1$

Q let $X_1, X_2, \ldots, X_n$ be iid R.V. from $N(0,1)$. Determine the Asymptotic distribution of $\frac{1}{n}\sum |X_i|$.

Answer use CLT. let $S_n = \frac{1}{n}\sum |X_i|$, then $\frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \longrightarrow N(0,1)$

or $\frac{S_n - E(S_n)}{\frac{\sqrt{Var(X)}}{\sqrt{n}}} \longrightarrow N(0,1)$. but $E(S_n) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2}} dx, = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} x e^{-\frac{x^2}{2}} d.$

$= \sqrt{\frac{2}{\pi}}$.

but $Var(|X_i|) = E(|X_i|^2) - (E(|X_i|))^2$ but $E(|X_i|^2) = \int \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$

so $Var(|X_i|) = 1 - \left(\sqrt{\frac{2}{\pi}}\right)^2 = 1 - \frac{2}{\pi}$. Plug in

T-distribution: $\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$

$\frac{\chi^2_{(m)}/m}{\chi^2_{(n)}/n} \sim F_{(m,n)}$

Characterization

$exp(1) \sim \frac{1}{2}\chi^2_{(2)}$

$\frac{N(0,1)}{\sqrt{\frac{\chi^2_{(n)}}{n}}} \sim t_{(n)}$

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

population Variance

$X \sim poisson(\lambda)$

$\frac{X-\lambda}{\sqrt{\lambda}} \xrightarrow{\lambda\to\infty} N(0,1)$

for large λ $\underset{\text{discrete}}{X} \simeq \underset{\text{cont.}}{N(\lambda,\lambda)}$

$P(X=1000) = P(999.5 < X < 1000.5)$

$\frac{N(0,2)}{\sqrt{2}} \sim N(0,1)$

if $X \sim N(\mu,\sigma^2)$ Then $\frac{(X-\mu)^2}{\sigma^2} \sim \chi^2_{(1)}$

Where S^2 is Sample Variance of iid $X_1, \ldots, X_n$
$N(\mu,\sigma^2)$

Ratio of 2 exponentials is F distribution

$\frac{\chi^2_{n-1}}{\chi^2_{m-1}} = F_{n-1,m-1}$

moment generation of $N(P,0) = e^{Pt}$

$exp(\lambda) \longrightarrow M(t) = \left(\frac{\lambda}{\lambda - t}\right)$.

$\Gamma(\alpha = n, \beta = n\lambda) \longrightarrow M(t) = \left(\frac{n\lambda}{n\lambda - t}\right)^n$

*hints when we see sum of R.V. think of Mgf.

Variance and σ. $Var(X) = E[(X - E(X))^2] = E(X^2) - (E(X))^2$

so $Var(X) = \int (x - \mu_x)^2 f(x)\, dx$.

$Y = a + bX \Rightarrow Var(Y) = b^2 Var(X)$.

IF asked to Find Variance of R.V.

IF R.V. Binomial, use indicator variable

$Y = X_1 + X_2 + \cdots + X_n$, where X_i is Bernoulli $\Rightarrow Var(Y) = np(1-p)$.

Covariance $Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y)\, f(x,y)\, dx\, dy$

IF asked to find expectation of R.V. given its expected conditional use Law of total expectation. $E(Y) = E(E(Y|X))$

IF asked to Find Expected Value or Var of some nonlinear function, then use taylor.

Example. Find $E(\sqrt{x})$ where $X \sim$ poisson (λ)

first let $g(X) = \sqrt{x}$, expand $g(X)$ around μ_X using taylor:

$$g(x) \sim g(\mu_x) + (x - \mu_x) g'(x) + (x - \mu_x)^2 \frac{g''(x)}{2!} + \cdots$$

then write $E(g(x)) = E(\quad\quad\quad)$. find $g'(x)$, $g''(x)$ seperately

for μ_x, μ_x^2 use table.

Expected Value:

IF Just one R.V. Then do

$E(X) = \sum_{all\ x} x_i P(x_i)$ or $\int x f(x)\, dx$

if Summation of R.V. as $\sum_{i=1}^{n} X_i$ Then

$E(S_n) = \sum E(X_i)$ -- need to know $E(X)$.

if R.V.

IF R.V. Binomial then use indicator Variable Bernelli

$0 \times (1-p) + 1 \times p = p$

and sum these we get $E(Y) = np$.

Expected Value.

IF just one R.V. Then

$E(X) = \sum x_i P(x_i)$ or $\int x f(x)\, dx$

IF Binomial Then do use indicator variable Bernulli. $0 \times (1-p) + 1 \times p = p$ and do $\sum E(\text{Bernolli}) = np$

IF geometric, use trick $E(X) = \sum_{k=1}^{\infty} k\, p (1-p)^{k-1} = p \sum k (1-p)^{k-1}$

$k q^{k-1} = \frac{d}{dq} q^k \Rightarrow E(X) = p \frac{d}{dq} \sum q^k \rightarrow p \frac{d}{dq} \frac{q}{1-q} = \frac{p}{(1-q)^2} = \boxed{\frac{1}{p}}$

if poisson Then $\sum_{0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda e^{-\lambda} e^{\lambda} = \boxed{\lambda}$

if Gamma $\int_0^{\infty} x \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \rightarrow E(X) = \boxed{\frac{\alpha}{\lambda}}$

if Normal $E(X) = \mu$.

if Cauchy $E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\pi} \frac{1}{(1+x^2)} dx \longrightarrow$ do Not exits as $\int_{-\infty}^{\infty} \frac{|x|}{1+x^2} = \infty$.

IF Function of R.V. as in $E(g(x))$ Then use theorm A $E(Y) = \sum_{all\ x} g(x) P(x)$ and $E(Y) = \int g(x) f(x)\, dx$

stick example here.

IF Liner Combintion of R.V. Then

$Y = a + \sum b_i X_i \Rightarrow E(Y) = a + \sum b_i E(X_i)$

compr Collection Example.

IF asked to identify distribution of Some function of 2 random Variables Then $\longrightarrow$

Consider Characterization. Eample $\frac{X}{\sqrt{Y}}$ with $X \sim N(0,2)$, $Y \sim \exp(1)$

$$\boxed{\begin{array}{l} \frac{X-0}{\sqrt{2}} \sim N(0,1) \\ \exp(1) \sim 2\chi^2_{(2)} \end{array}}$$

But $\frac{N(0,1)}{\sqrt{\frac{\chi^2_{(n)}}{n}}} \sim t_{(n)}$. so

otherwize, use Moment generation $\leftarrow$ Can require Jacobian $\longrightarrow$

A stick of unit length is broken in 2 places what is expected length of middle piece?

$E|U_1-U_2| = \int_0^1\int_0^1 |u_1-u_2|\, du_1\, du_2$

$= \int_0^1 \int_0^{u_1} (u_1-u_2)\, du_2\, du_1 + \int_0^1 \int_{u_1}^1 (u_2-u_1)\, du_2 du_1$

$= \frac{1}{3}$

To find expected value of a R.V. we sometimes write the R.V. as a sum of another R.V. whom we know its expected value.

Binomial = $\sum$ Bernoulli

Chebyshev: $P(|X-\mu_x| \geq t) \leq \frac{Var(x)}{t^2}$

or $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

identities for $Var(X)$

$Var(X) = E((X-\mu_x)^2)$

$= E(X^2) - (E(X))^2$

$Var(bX) = b^2\, Var(X)$

$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$

$Var(X-Y) = Var(X) + Var(Y) - 2\,Cov(X,Y)$

$Var(X) = Cov(X,X)$

$Var(X-Y) = Cov(X-Y, X-Y) = Cov(X-Y, X) + Cov(X-Y, -Y)$

$= Cov(X,X) - Cov(X,Y)$

$+ Cov(-Y,X) - Cov(-Y,Y)$

$= \boxed{Var(X) + Var(Y) - 2\,Cov(X,Y)}$

covariance

$Cov(X,Y) = E((X-\mu_X)(Y-\mu_Y))$

$= E(XY) - E(X)E(Y)$.

note if $\perp$ then $Cov(X,Y)=0$

if Cov is negative, then if X less than its mean, then Y is more than Y mean.

$Cov(a+X, Y) = Cov(X,Y)$

$Cov(aX, bY) = ab\, Cov(X,Y)$

$Cov(X, Y+Z) = Cov(X,Y) + Cov(X,Z)$

Law of total Expectation:

$E(Y) = E(E(Y|X))$

"$E(Y) = \sum_{all\, x} E(Y|X)\, P_X = E(Y|X=1)\,P(X=1) + E(Y|X=2)\,P(X=2) + \dots$

Random Sum $Y = \sum_{i=1}^{N} X_i$ Random

so $E(Y) = E(E(Y|N=n))$

$= E(n\, E(X))$, then

$E(Y) = E(N\, E(X)) = E(N)\,E(X)$ $\perp$

Correlation $\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)\, Var(Y)}}$

$Var(Y) = Var[E(Y|X)] + E[Var(Y|X)]$

$Var(Y|X) = E(Y^2|X) - (E(Y|X))^2$

$E(Y|X)$ is a function of X! since it depends on X. but since X is R.V. then $E(Y|X)$ is a R.V and it has expectation $E(E(Y|X))$

so $E(Y) = E(E(Y|X))$

Chapter 5 distributions derived from Normal: if $Z \sim N(0,1)$ then $U = Z^2$ is chisquare(1).

if $X \sim N(\mu, \sigma^2)$ then $\frac{X-\mu}{\sigma} \sim N(0,1)$

if $Z \sim N(\mu, \sigma^2)$, then $\frac{(Z-\mu)^2}{\sigma^2} \sim$ Chisquare

Poisson distribution: Probability that $X = x$ events in t time

let λ = expected number of events in t time. then we write

$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$. For example: calls arrive at a rate of .5 calls per min. Find probability of one call in 5 minutes? λ = expected # of calls in 5 minutes. this is $5 \times .5 = 2.5$ calls so $P(X=1 \text{ call in 5 mins}) = \frac{2.5^1 e^{-2.5}}{1!} = 0.205$

if expected # of events in given time is large, then poisson R.V. (i.e. # of ~~events~~ in this time) can be approximated by Normal distribution.

Note $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Standard deviation of $\bar{X}$, the sample mean, is also called standard error. $\Rightarrow \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

i.e. Variance of sample mean $= \frac{\sigma^2}{n}$.

Central limit (μ, σ^2) ... $\sum^n x_i$ → distribution of sum $\sim N(n\mu, \sigma^2 n)$

sample $\bar{X}$ → distribution $\sim N(\mu, \frac{\sigma^2}{n})$

Expected Value: To find $E(X+Y)$, and X,Y are not $\perp\!\!\!\perp$. Then need joint PDF of X,Y and then do $E(X+Y) = \sum_{\substack{\text{all}\\ \text{combinations of}\\ X+Y}} P(X=?,Y=?)$ ..

3 girls, 4 boys, committee of 3. What is the Expected number of boys in the committee?

do $1\,P(X=1) + 2\,P(X=2) + 3\,P(X=3)$

$P(X=1) = \frac{\text{number of ways to have one boy in committee}}{\text{Total number of ways to select committee}}$

$= \frac{\binom{3}{2}\binom{4}{1}}{\binom{7}{3}}$, do this for $P(X=2)$, $P(X=3)$ etc...

$E(Y) = E(E(Y|X=x))$ "average of averages"

$Var(X) = \sum_{i=1}^{n} (x_i - \mu_x)^2 P(X=x_i)$

this is $Var(X) = E((X-\mu_x)^2)$

Expectation of sum:

Suppose 75% of students taking class pass in class of $N=40$, what is expected number who will pass? Find Variance and σ.

$E(X) = E\sum_{i=1}^{N} Z_i$ where $Z_i = \begin{cases} 1 & \text{if ith student pass} \\ 0 & \text{if ith fail} \end{cases}$

$E(X) = \sum E(Z_i)$.

but $E(Z_i) = 1 \times .75 + 0 \times .25 = .75$

so $E(X) = \sum^{N} .75 = (.75)(40)$.

$Var(X) = Var\sum Z_i = \sum Var(Z_i)$

but $Var(Z_i) = E(Z_i^2) - [E(Z_i)]^2$

but $E(Z_i^2) = 1^2 \times 0.75 + 0^2 \times .25 = .75$

so $Var(Z_i) = .75 - .75^2 = 0.1875$

so $Var(X) = \sum .1875 = 40(.1875)$. QED

problem given $f(x,y) = x+y$, $0<x<1$, $0<y<1$

Find $f(y|x)$, find Conditional expectation of Y?

solution $f(x) = \int_0^1 f(x,y)\,dy = x + \frac{1}{2}$.

$f(y|x) = \frac{f(x,y)}{f(x)} = \frac{x+y}{x+\frac{1}{2}}$ $0<y<1$, $0<x<1$

$E(Y|X=x) = \int_{\text{all } y} y\, f(y|x)\,dy = \int_0^1 y \frac{x+y}{x+\frac{1}{2}}\,dy$

Expection problem: $P(hit) = \frac{3}{10}$. Find $E(X)$, where X is number of hits in N trials.

write $X = Z_1 + Z_2 + \cdots + Z_N$ where

$Z_i = \begin{cases} 1 & \text{if ith hit was good} \\ 0 & \text{if ith hit was bad.} \end{cases}$

so $E(X) = E(\sum Z_i) = \sum E(Z_i)$

But $E(Z_i) = 1 \times P(\text{hit good}) + 0 \times P(\text{hit bad})$.

$E(X) = \sum \frac{3}{10} = \boxed{n(p)}$

$Var(X) = Var\sum Z_i \Rightarrow Var(X) = \sum_i^N Var(Z_i)$ because Z_i's $\perp\!\!\!\perp$ → $Var(X) = \sum^N (p - p^2) = \boxed{N\,p(1-p)}$

but $Var(Z_i) = E(Z_i^2) - \{E(Z_i)\}^2 = E(Z_i^2) - p^2$, but $E(Z_i^2) = 1^2 p + (0)^2(1-p) = p$ QED

$Var(X+Y) = Var(X) + Var(Y)$ if X,Y $\perp\!\!\!\perp$

$Var(aX) = a^2 Var(X)$.

$Var(X+Y)$ if X,Y NOT $\perp\!\!\!\perp$ $= Var(X) + Var(Y) + 2\,Cov(X,Y)$.

$Cov(X,Y) = E(XY) - E(Y)E(X)$.

Coupon problem: Collect coupons, n types. How many trials would you expect to do to win all coupons?

$$\sum_{k=1}^{\infty} q^k = \frac{q}{1-q}$$

X_i = # trials up to including i^{th} to win
so $X_1 = 1$, X_2 = # of trials to win second coupon

$$E(X) = \sum E(X_r)$$

but $P(\text{win at trial } r) = \frac{n-r+1}{n}$.

$$E(X_r) = \frac{n}{n-r+1}$$

$$\Rightarrow E(X) = \sum_{r=1}^{n} X_r = \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{1}$$

$$= n \sum \frac{1}{r} = n\left[\log n + \gamma + \varepsilon_n\right]$$

Approximation for $Z = g(X,Y)$

$$E(Z) \approx g(\mu) + \frac{1}{2}\sigma_X^2 \frac{\partial^2 g(\mu)}{\partial x^2} + \frac{1}{2}\sigma_Y^2 \frac{\partial^2 g(\mu)}{\partial y^2} + \sigma_{XY}\frac{\partial^2 g(\mu)}{\partial x \partial y}$$

$$Var(Z) \approx \sigma_X^2 \left(\frac{\partial g(\mu)}{\partial x}\right)^2 + \sigma_Y^2 \left(\frac{\partial g(\mu)}{\partial y}\right)^2 + 2\sigma_{XY}\left(\frac{\partial g(\mu)}{\partial x}\right)\left(\frac{\partial g(\mu)}{\partial y}\right)$$

for $g(x,y) = y/x$

$$\Rightarrow E(Z) \approx \frac{\mu_Y}{\mu_X} + \frac{1}{\mu_X^2}\left(\sigma_X^2 \frac{\mu_Y}{\mu_X} - \rho\sigma_X\sigma_Y\right)$$

$$Var(Z) = \frac{1}{\mu_X^2}\left(\sigma_X^2 \frac{\mu_Y^2}{\mu_X^2} + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y \frac{\mu_Y}{\mu_X}\right)$$

Q if X, Y ⫫, find $Var(XY)$

$$Var(XY) = E((XY)^2) - [E(XY)]^2$$

$$E(XY) = E(X)E(Y)$$

$$E(X^2Y^2) = E(X^2)E(Y^2)$$

$$= \left(Var(X) + [E(X)]^2\right)\left(Var(Y) + [E(Y)]^2\right)$$

$$= [\sigma_X^2 + \mu_X^2][\sigma_Y^2 + \mu_Y^2]$$

$= \ldots$

so $Var(XY) = \ldots$ $\qquad f(x) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}}$

let X be exp R.V. with std σ. Find $P(|X-\mu_X| > k\sigma)$ for $k=2$

$$P(|X-\mu_X| > k\sigma) = P(|X-\sigma| > k\sigma) \quad \because \mu_X = \sigma$$

$$= 1 - P(|X-\sigma| \le k\sigma)$$

$$= 1 - P(X \le \sigma(k+1))$$

$$= 1 - \int_0^{\sigma(k+1)} \frac{1}{\sigma} e^{-\frac{x}{\sigma}} dx = e^{-(k+1)}$$

so for $k=2$ ...

More formulas

Approximation. (δ method)

if we know μ_X, and σ_X but not the distribution. Suppose next we are interested in the mean and Var of some function of X. say $Y = g(X)$.
we can't find mean, Var of Y from X unless $g(x)$ is linear. however, if g is linear in range in which X has high probability, it can be approximated by linear function, and approximate moments of Y can be found.
we expand g around μ_X.

$$Y = g(X) \approx g(\mu_X) + (X-\mu)g'(\mu_X) + \frac{1}{2}(X-\mu_X)^2 g''(\mu_X)$$

$$\Rightarrow \mu_Y \quad \boxed{E(Y) \approx g(\mu_X) + \frac{1}{2}\sigma_X^2 g''(\mu_X)}$$

$$\sigma_Y^2 \quad \boxed{Var(Y) \approx \sigma_X^2 [g'(\mu_X)]^2}$$

On solving this problem: Let U1, U2, U3 be independent random variables uniform on [0,1]. Find the probability of the roots of the quadratic $U1x_2 + U2x + U3$ are real

Small investigation into problem 11

Nasser Abbasi, 9/26/07

The problem :
Let U_1, U_2, U_3 be independent random variables uniform on [0,1]. Find the probability of the roots of the quadratic $U_1 x^2 + U_2 x + U_3$ are real
Answer:
Roots are real when discriminant is ≥ 0

In[25]:=

```
eq = U1 x^2 + U2 x + U3;
expr = First@Solve[eq == 0, x];
f = First@Cases[expr, Sqrt[any_] → any, Infinity] (*Pull out the expression under the sqrt *)
```

Out[27]= $U2^2 - 4\, U1\, U3$

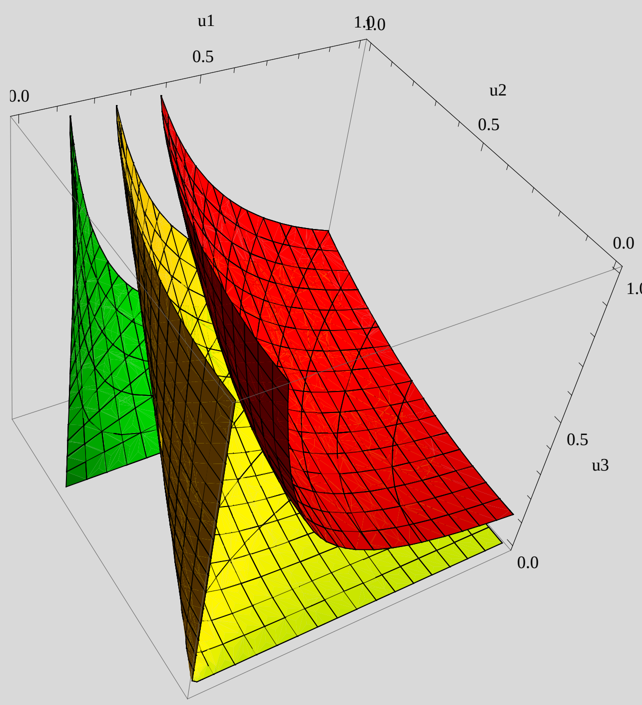
Hence we want to find P($U2^2$ – 4 U1 U3 >0)

This is the VOLUME between the above surface and between a cube of side 1. i.e. a cube of volume 1.

As an initial look, One way to view this is to look at the constant surface contours in 3 D space. We can look at the constant contour surfaces in which the function $U2^2$ – 4 U1 U3 is zero. And then look at the surface in which this function is on the positive side and then on the negative side of it. This will give us an idea where the volume of interest lies in relation to the zero contour surface.

In[28]:=
```
ContourPlot3D[f , {U1, 0, 1}, {U2, 0, 1}, {U3, 0, 1}, AxesLabel → {"u1", "u2", "u3"},
 Contours → {0, -.5, .5}, ContourStyle → {Yellow, Red, Green}]
```

Out[28]=

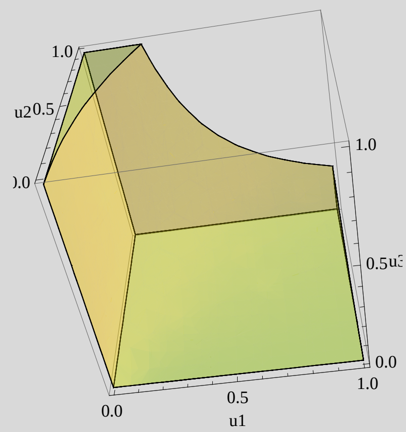


Cell[TextData[{Cell[TextData[{ValueBox["FileName"]}], "Header"], Cell[" ", "Header", CellFrame -> {{0, 0.5}, {0, 0}}, CellFrameMargins -> 4], " ",
Cell[TextData[{CounterBox["Page"]}], "PageNumber"]}], CellMargins -> {{Inherited, 0}, {Inherited, Inherited

We see from the above that surfaces below the yellow surface (the GREEN) are positive, and those above it (RED) are negative. We can get a better view of the volume by getting a plot of the region where such a function is POSITIVE. Next we draw the solid region where this function is POSITIVE

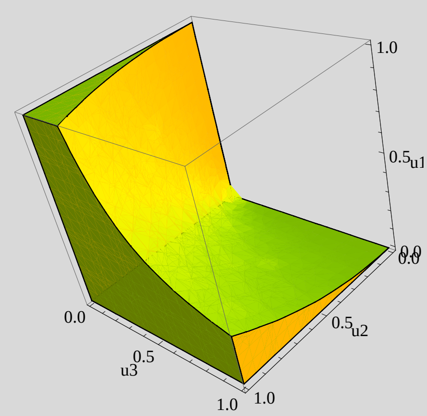
In[29]:=
```
RegionPlot3D[f > 0, {U1, 0, 1}, {U2, 0, 1}, {U3, 0, 1},
 AxesLabel → {"u1", "u2", "u3"}, PlotStyle → Directive[Yellow, Opacity[0.5]], Mesh → None]
```

Out[29]=



In[30]:=
```
RegionPlot3D[f > 0, {U1, 0, 1}, {U2, 0, 1}, {U3, 0, 1},
 AxesLabel → {"u1", "u2", "u3"}, PlotStyle → Directive[Yellow], Mesh → None]
```

Out[30]=



So the solid volume in the above represents the numerical value of the probability we are looking for. It is hard for me now to see the regions of integration analytically, but there is a simulation run which gives an approximate value for the probability we need

In[31]:=
```
n = 3 000 000; SeedRandom[010 101];
u1 = Table[RandomReal[{0, 1}], {i, n}];
u2 = Table[RandomReal[{0, 1}], {i, n}];
u3 = Table[RandomReal[{0, 1}], {i, n}];
r = Select[u2^2 - 4 u1 u3, # ≥ 0 &];
Print["Probability is ", N[Length[r] / n]]
```

```
Probability is 0.253976
```