chiSquare χ^2 distribution with degree $n = X_1^2 + X_2^2 + \dots + X_n^2$ when X_i is standar. $N(0,1)$

Standard Normal $Z = \frac{X-\mu}{\sigma}$ where X is Normal with mean μ and variance σ^2.

$F(\chi^2_{(n)} < 0) = 0$ | $E(\chi^2_{(n)}) = n$. $Var(\chi^2_{(n)}) = 2n$ | $\chi^2_{(n)} = \frac{(\frac{1}{2})^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$

$\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt$ | $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(1) = 1$ | 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 $n = 0, 1, 2, \dots$ $\Gamma(n) = n!$ | $\Gamma(n+1) = n\Gamma(n)$

Normal distribution $f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ or $= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

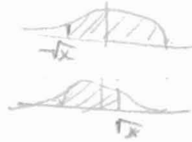
Standard Normal $N(0,1) = \boxed{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}$.

derive $\chi^2_{(1)}$. let $X = Z^2$ where $Z \sim N(0,1)$.

$F_X(x) = P(X < x) = P(Z^2 < x) = P(-\sqrt{x} < Z < \sqrt{x})$

$F_X(x) = P(-\sqrt{x} < Z) + P(Z < \sqrt{x})$

$\Rightarrow \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$



$f_X(x) = f_Z(\sqrt{x}) \cdot \frac{1}{2}\frac{1}{\sqrt{x}} + f_Z(-\sqrt{x}) \frac{1}{2}\frac{1}{\sqrt{x}}$

but $f_Z = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(u)^2}$. let $u = \sqrt{x}$, $u = \sqrt{x}$ in above.

plug you'll get $\chi^2_{(1)}$ formula

CLT sum and mean, standardized, converges to $N(0,1)$

$Z_n = \frac{\sum X_i - n\mu_X}{\sqrt{Var X}\ \sqrt{n}}$

Standardize sum.

each X has same dis.

Standarizing: $Z_n = \frac{X_n - E(X_n)}{\sqrt{Var(X_n)}}$ ⇒ Z_n has 0 mean, and Var 1.

use this

apply this to $\sum X_i$ and $\bar{X}_n$

$Z_n = \frac{\bar{X}_n - \mu_X}{\frac{\sqrt{Var(X)}}{\sqrt{n}}}$

E(X) avg X.

Standarize mean

Chi Square ($\chi^2_{(n)}$).

definition: $\chi^2_{(n)} = \underbrace{Z^2(0,1) + Z^2(0,1) + \dots + Z^2(0,1)}_{n \text{ times}}$

$\chi^2_{(n)} = \frac{(\frac{1}{2})^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$. $\chi^2_{(1)} = \frac{x^{-\frac{1}{2}}}{\sqrt{2\pi}} e^{-\frac{x}{2}}$. $\chi^2_{(2)} = \frac{\frac{1}{2}}{\Gamma(1)_{=1}} x^{\frac{1}{2}} e^{-\frac{x}{2}}$

$E(\chi^2_{(n)}) = n$, $Var(\chi^2_{(n)}) = 2n$

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 $\dfrac{S_n - E(S_n)}{\frac{\sqrt{Var(X)}}{\sqrt{n}}} \longrightarrow N(0,1)$. but $E(S_n) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} |x| e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} x e^{-\frac{x^2}{2}} dx$

$= \sqrt{\frac{2}{\pi}}$

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: $\dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$

$\dfrac{\chi^2_{(m)}/m}{\chi^2_{(n)}/n} \sim F_{(m,n)}$

Charaterization

$\exp(1) \sim \frac{1}{2}\chi^2_{(2)}$

$\dfrac{N(0,1)}{\sqrt{\frac{\chi^2_{(n)}}{n}}} \sim t_{(n)}$

$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$

σ^2 population Variance

$X \sim$ poisson(λ)

$\frac{X-\lambda}{\sqrt{\lambda}} \xrightarrow{\lambda\to\infty} N(0,1)$

$\frac{N(0,2)}{\sqrt{2}} \sim N(0,1)$

for large λ $\underset{\text{discrete}}{X} \approx \underset{\text{Cont.}}{N(\lambda,\lambda)}$

$P(X=1000) = P(999.5 < X < 1000.5)$

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 generating 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 sums 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

$Y = X_1 + X_2 + \dots + X_n$, 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 non linear 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:

$$g(X) \sim g(\mu_X) + (X-\mu_X)\, g'(x) + (X-\mu_X)^2 \frac{g''(x)}{2!} + \dots$$

then write $E(g(X)) = E(\qquad\qquad)$. find $g'(x)$, $g''(x)$ separately

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 Bernulli

$0 \times (1-p) + 1 \times p = p$ 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 $\sum E(\text{Bernulli}) = np$

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 $E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\pi} \frac{1}{(1+x^2)} dx \rightarrow$ 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 Combination of R.V. Then

$Y = a + \sum b_i X_i \Rightarrow E(Y) = a + \sum b_i E(X_i)$

compon Collection Example.

IF asked to identify distribution of Some function of 2 random Variables Then →

Consider Characterization. Example $\frac{X}{\sqrt{Y}}$ — X: $N(0,2)$, Y: $\exp(1)$

$$\boxed{\begin{array}{l} \frac{X-0}{\sqrt{2}} \sim N(0,1) \\ \exp(1) \sim 2\chi^2_{(2)} \end{array}}$$

But $\frac{N(0,1)}{\sqrt{\frac{\chi^2_{(n)}}{n}}} \sim t_{(n)}$. so

otherwize, use Moment generation ← Can require Jacobian →

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. whose we know its expected value.

Binomial = Σ Bernoulli

Chebyshev: $P(|X-\mu_x| \geq t) \leq \dfrac{Var(X)}{t^2}$

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 Expection:

$$E(Y) = E(E(Y|X))$$

or $E(Y) = \sum_{all\,x} E(Y|X)\,P_x = E(Y|X=1)\,P(X=1) + E(Y|X=2)\,P(X=2) + \dots$

Random Sum $Y = \sum_{i=1}^{N} X_i$ (N ← Random)

so $E(Y) = E(E(Y|N=n))$

$E(Y)$ $= E(n\,E(X))$, then

$= E(N\,E(X)) = E(N)E(X)$ $\perp$

Correlation $\rho = \dfrac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$

$$Var(Y) = Var[E(Y|X)] + E[Var(Y|X)]$$

$$Var(Y|X) = E(Y^2|X) - (E(Y|X))^2$$

$E(Y|X)$ 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

Poison distribution: Probability that $X=x$ events in t time

let λ = <u>expected</u> number of events in t time. then we write

$P(X=x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$. For example: calls arrive at a rate of <u>.5 calls</u> per min. Find probability of <u>one</u> call in 5 minutes? λ = expected # of calls in 5 minutes. this is $5 \times .5 =$ <u>2.5 calls</u> so $P(X=1 \text{ call in 5 mins}) = \dfrac{2.5^1 e^{-2.5}}{1!} = \underline{0.205}$

if expected # of events in given time is large, then poisson R.V. (i.e. $P(X=x)$ ~~events~~ in this time) can be approximated by Normal distribution.

<u>Note</u> $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Standard deviation of $\bar{X}$, the Sample mean, is also called standard error. $\Rightarrow \sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}}$

i.e. Variance of Sample mean $= \dfrac{\sigma^2}{n}$.

<u>Central limit</u> (μ, σ^2) ⟶ (...) ⟶ $\sum^n x_i$ ⟶ distribution of sum $\sim N(n\mu, \sigma^2 n)$

sample ⟶ $\bar{X}$ ⟶ distribution $\sim N(\mu, \frac{\sigma^2}{n})$

---

3 girls, 4 boys, committee of 3. What is the Expected number of boys in the committee?

do $1\,P(X=1) + 2\,P(X=2) + 3\,P(X=3)$

$P(X=1) = \dfrac{\text{number of way to have one boy in committee}}{\text{Total number of ways to select committee}}$

$= \dfrac{\binom{3}{2}\binom{4}{1}}{\binom{7}{3}}$, do this for $P(X=2)$, $P(X=3)$ etc...

<u>problem</u> given $f(x,y) = x+y$ $\quad 0<x<1,\ 0<y<1$

Find $f(y|x)$, find Conditional expectation of Y?

<u>solution</u> $f(x) = \int_0^1 f(x,y)\,dy = x + \frac{1}{2}$.

$f(y|x) = \dfrac{f(x,y)}{f(x)} = \dfrac{x+y}{x+\frac{1}{2}}$ $\quad 0<y<1,\ 0<x<1$

$E(Y|X=x) = \int_{\text{all } y} y\, f(y|x)\,dy = \int_0^1 y\,\dfrac{x+y}{x+\frac{1}{2}}\,dy$

Expection problem: $P(\text{hit}) = \frac{3}{10}$. Find $E(X)$, where X is number of hits in N trials.

write $X = Z_1 + Z_2 + \dots + Z_N$ where

$Z_i = \begin{cases} 1 & \text{if } i\text{th hit was good} \\ 0 & \text{if } i\text{th hit was bad} \end{cases}$

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 $E(Z_i^2) = 1^2 \cdot p + (0)^2(1-p) = p$ QED

Expected Value: To find $E(X+Y)$, and X, Y are not $\perp\!\!\!\perp$. Then need joint PDF of X, Y and then do $E(X+Y) = \sum_{\text{all combinations of } X+Y} \quad P(X=?, Y=?) \dots$

$E(Y) = E(E(Y|X=x))$ "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 $Z_i = \begin{cases} 1 & \text{if } i\text{th student pass} \\ 0 & \text{if } i\text{th fail.} \end{cases}$

$E(X) = \sum E(Z_i)$.

but $E(Z_i) = 1 \times .75 + 0 \times .25 = .75$

$\therefore E(X) = \sum^{N} .75 = (.75)(40)$

$Var(X) = Var \sum Z_i = \sum Var(Z_i)$

but $Var(Z_i) = E(Z_i^2) - [E(Z_i)]^2$

but $E(Z_i^2) = 1^2 \times 0.75 + 0^2 \times .25 = .75$

$\therefore Var(Z_i) = .75 - .75^2 = 0.1875$

$\therefore Var(X) = \sum .1875 = 40(.1875)$. 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

$X_1 = 1$, 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 $Var(XY)$

$Var(XY) = E((XY)^2) - [E(XY)]^2$

$E(XY) = E(X)E(Y)$

$E(X^2Y^2) = E(X^2)E(Y^2)$

$= \left(Var(X) + [E(X)]^2\right)\left(Var(Y) + [E(Y)]^2\right)$

$= [\sigma_X^2 + \mu_X^2][\sigma_Y^2 + \mu_Y^2]$

$= \cdots$

so $Var(XY) = \cdots$ $\quad f(x) = \frac{1}{\sigma}e^{-\frac{x}{\sigma}}$

let X be exp R.V. with std σ. Find $P(|X-\mu_X| > k\sigma)$ for $k=2$

$P(|X-\mu_X| > k\sigma) = P(|X-\sigma| > k\sigma)$ $\quad \because \mu_X = \sigma_X$

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

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

$= 1 - \int_0^{\sigma(k+1)} \frac{1}{\sigma} e^{-\frac{x}{\sigma}}dx = e^{-(k+1)}$

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.

$$Y = g(X) \approx g(\mu_X) + (X-\mu)g'(\mu_X) + \frac{1}{2}(X-\mu_X)^2 g''(\mu_X)$$

$$\Rightarrow \boxed{E(Y) \approx g(\mu_X) + \frac{1}{2}\sigma_X^2 g''(\mu_X)} \quad \mu_Y$$

$$\sigma_Y^2 \quad \boxed{Var(Y) \approx \sigma_X^2[g'(\mu_X)]^2}$$