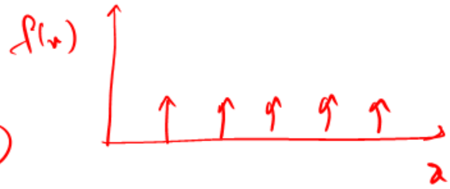
Continuous Random Variables Distributions



1. Continuous Uniform Distribution

$X \sim \text{Unif}(a, b)$

A continuous random variable X with probability density function

$\boxed{f(x) = 1/(b-a)}$ a<=x<=b is a continuous uniform random variable

In this model the total probability of 1 is spread evenly between two limits, so the subintervals of same length have same probability

$$f(x) = \frac{1}{b-a}$$

$$E[X] = \int x f(x)$$

$$= \int_a^b x \cdot \frac{1}{b-a} = \frac{1}{b-a} \frac{x^2}{2}\Big|_a^b = \frac{1}{b-a} \frac{b^2 - a^2}{2}$$

$$E[X] = \frac{b+a}{2}$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 = \frac{(a-b)^2}{12}$$

$$M_X(x) = E[e^{tx}] = \int e^{tx} f(x)\, dx$$

$$E[e^0] = E[1] = \int_a^b f(x)\,dx = \int_a^b e^{tx} \cdot \frac{1}{b-a}\, dx = \frac{e^{tx}}{t(b-a)}\Big|_a^b$$

$$M_X(x) = \begin{cases} \dfrac{e^{bt} - e^{at}}{t(b-a)} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

Let the continuous random variable X denote the current measured in a thin copper wire in milliamperes. Assume that the range of X having uniform distribution[0, 20 mA], and assume that the probability density function of X is $\boxed{f(x) = \frac{1}{20}}$

$X \sim U(\overset{a}{0}, \overset{b}{20})$

$f(x) = \frac{1}{b-a}$

What is the probability that a measurement of current is between 5 and 10 milliamperes?

$$P(5 < X < 10) = \int_5^{10} f(x)\,dx$$

$$= \int_5^{10} \frac{1}{20}\,dx = \frac{1}{4}$$

2. Exponential distribution:

PMF $f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$

$X \sim Poi(\lambda)$

Poisson distribution-models the number of events that ~~occur in a specified in~~terval of time.

Let Poisson distribution defined a random variable to be the number of flaws along a length of copper wire.

number flaws

The distance between flaws is another random variable that is often of interest.

Let the random variable X denote the length from any starting point on the wire until a flaw is detected.

3mm

EX- The distance to the first flaw exceeds 3 millimeters if and only if there are no flaws within a length of 3 millimeters

$P(X>3) = P(N=0)$

In general, let the random variable N denote the number of flaws in x millimeters of wire. If the mean number of flaws ~~is per millimeter~~, N has a Poisson distribution with mean $\boxed{\lambda x}$

We assume that the wire is longer than the value of x. Now,

$N \sim Poi(\lambda x)$

$$P(X>x) = P(N=0) = \frac{(e^{-\lambda x})(\lambda x)^{0}}{0!}$$

PMF : $P_X(x) = \frac{e^{-\lambda x}(\lambda x)^{x}}{x!}$

$$P(X>x) = e^{-\lambda x}$$

CDF

$$F_X(x) = P(X \le x) = \int_0^x f(u)\,du$$
$$= 1 - P(X>x)$$
$$= 1 - e^{-\lambda x}$$

PDF

$$\boxed{f_X(x) = F'_X(x) = \lambda e^{-\lambda x}}$$

$$E[X] = \int_0^\infty x f(x)\,dx$$
$$= \int_0^\infty x \lambda e^{-\lambda x}\,dx$$
$$= 1/\lambda$$

Exponential distribution

The random variable X that equals the distance between successive events of a Poisson process with mean number of events λx per unit interval is an exponential random variable with parameter The probability density function of X is

PDF $\boxed{f_X(x) = \lambda e^{-\lambda x} \quad 0 \le x < \infty}$

$Var(X) = 1/\lambda^2$

Example:
In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean of 25 log-ons per hour. What is the probability that there are nolog-ons in an interval of 6 minutes?

$\boxed{\lambda = 25}$

λ = per hour

$X \sim \exp(\lambda = 25 \text{ logon/hr})$

6 min

$P(X > 6) = P(X > 0.1 \text{ hr})$

$f(x) = \lambda e^{-\lambda x}$

$f_X(x) = 25 e^{-25x}$

$P(X > 0.1) = \int_{0.1}^{\infty} f(x)\,dx$

$$= \int_{0.1}^{\infty} 25 e^{-25x}\,dx = 25 \frac{e^{-25x}}{-25}\Big|_{0.1}^{\infty} = e^{-2.5} = 0.082$$

$\mu = \frac{1}{\lambda} = \frac{1}{25}$, $\sigma^2 = \frac{1}{(25)^2}$

t_1

t_2

Special Property of exponential distribution(Memoryless property)

For an exponential random variable X,

$$\boxed{P\left(\frac{X < t_1 + t_2}{x > t_1}\right) = P(X < t_2)}$$

t_1 t_2

$P\left(\frac{X < 7}{x > 6}\right) =$

EX-

$X \sim Exp(\lambda)$ → detecting particle $E[x] = \frac{1}{\lambda}$

$E[x] = 1.4$ min. $\lambda = \frac{1}{E[x]} = \frac{1}{1.4}$

find the probability that we detected particle within 30sec of experiment.

$P(X < \boxed{30sec}) = \int_0^{0.5min} f(x)\,dx = \int_0^{0.5} \frac{5}{7} e^{-5/7 x}\,dx$

$= 1 - e^{-5/7(0.5)}$

$= 0.30$

pdf

$f_x(x) = \lambda e^{-\lambda x}$

$= \frac{5}{7} e^{-5/7 x}$

→ Suppose we wait for 3 minutes without any detection. What is the probability that a particle is detected in 30sec

$P\left(\frac{x < 3.5}{x > 3}\right) = P\left(\frac{x < 3.5 \cap x > 3}{P(x>3)}\right) = \frac{P(3 < x < 3.5)}{P(x>3)}$

$= 0.30$

Practical Interpretation: After waiting for 3 minutes without a detection, the probability of a detection in the next 30 seconds is the same as the probability of a detection in the 30 seconds immediately after starting the counter. The fact that you have waited 3 minutes without a detection does not change the probability of a detection in the next 30 se

X ~ distance or length from starting point to first flaw

$\lambda \Rightarrow$ mean num

N ~ Poisson (λx)

$\rightarrow$ number of flaws per unit length.

$X \sim \exp(\lambda)$

$$f_X(x) = \lambda e^{-\lambda x} \qquad 0 \le x < \infty$$

$$E[X] = 1/\lambda \quad , \quad Var(X) = \frac{1}{\lambda^2}$$

$f_X(x))$

$$M_X(t) = E[e^{tX}] = \int e^{tx} f_X(x)\, dx$$

$$= \int_0^\infty e^{tx} \lambda e^{-\lambda x}\, dx$$

Gamma Distribution $X \sim gamma(r, \lambda)$

An exponential random variable describes the length until the first count is obtained in a Poisson process.

A generalization of the exponential distribution is the length until r counts occur in a Poisson process. Consider the following example

Gamma function : $\Gamma(r)$

$\int y^{25} e^{-y}$ 3min $\Gamma(24)$

$= 23$

$$\boxed{\Gamma(r) = \int_0^\infty y^{r-1} e^{-y} \, dy}$$

Improper integral of 3rd kind

$$1 = \frac{1}{\Gamma(r)} \int_0^\infty y^{r-1} e^{-y} \, dy$$

$\boxed{y = \lambda x}$

$dy = \lambda \, dx$

$$1 = \frac{1}{\Gamma(r)} \int_0^\infty (\lambda x)^{r-1} e^{-\lambda x} \lambda \, dx$$

λ mean

$$1 = \int_0^\infty \boxed{\frac{\lambda^r x^{r-1} e^{-\lambda x} \, dx}{\Gamma(r)}}$$

$$f_X(x) = \begin{cases} \dfrac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$r = 1$ $\boxed{f_X(x) = \lambda e^{-\lambda x}}$

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x}\, dx$$

B. S. Grewal textbook

Properties

① $\Gamma(1) = 1$

$\Gamma(0, -ve \text{ integers})$

② $\Gamma(n+1) = n\Gamma(n)$

$$\int_0^\infty x^{r-1} e^{-x^2}$$

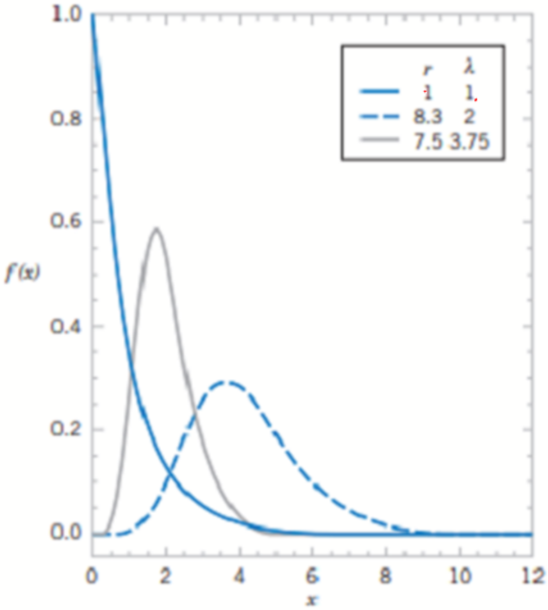
③ $\Gamma(n+1) = n!$ ~~whole~~ positive integers

④ $\boxed{\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}}$

$\Gamma(3/2) = \frac{1}{2}\Gamma(1/2) =$

$\Gamma(-1/2)$ $\quad \Gamma(5/2)$

Gamma density function



r → Changes shape of

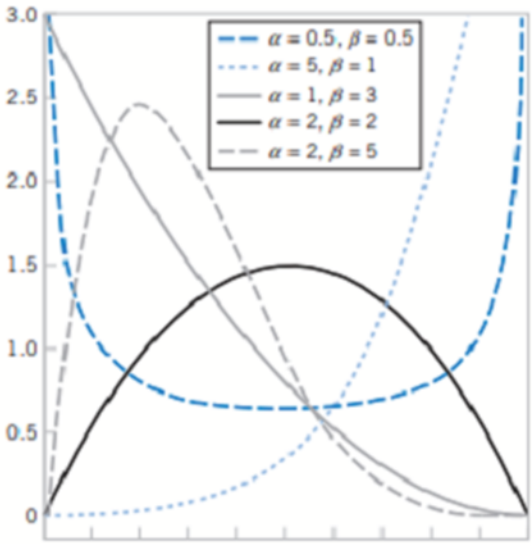
λ → scale will change

Beta Distribution

A continuous distribution that is flexble, but bounded over a finite range, is useful for probability models. The proportion of solar radiation absorbed by a material or the proportion (of the maximum time) required to complete a task in a project are examples of continuous random variables over the interval [0, 1]

$$f_x(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in [0,1]$$

$\alpha > 0, \beta > 0$



$$\int_0^1 x^3(1-x)^4$$

Beta function

$$\beta(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$$

$$\beta(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$1 = \int_0^1 \frac{x^{\alpha-1}(1-x)^{\beta-1}\,dx}{\beta(\alpha,\beta)}$$

$$1 = \int_0^1 \frac{\Gamma(\alpha+\beta)\, x^{\alpha-1}(1-x)^{\beta-1}\,dx}{\Gamma(\alpha)\,\Gamma(\alpha)}$$

EXAMPLE

The time to prepare a micro-array slide for high-throughput genomics is a Poisson process with a mean of two hours per slide. What is the probability that 10 slides require more than 25 hours to prepare? Find the mean and standard deviation of timeto prepare 10 slides?

$E[x] = 2 = 1/\lambda \Rightarrow \lambda = 0.5$

$\lambda = 2hr/slide$

X ~ time required to complete 10 slide

$X \sim gamma(r, \lambda)$, $r = 10$, $\lambda = 2hr/slide$

$$P(x > 25) = \int_{25}^{\infty} f_x(x)\,dx$$

pdf

$$f_x(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)} = \frac{(1/2)^{10} x^{9} e^{-\frac{1}{2}x}}{\Gamma(10)}$$

$$E[x] = \frac{r}{\lambda} = \frac{10}{0.5} = 20$$

$$Var(x) = \frac{r}{\lambda^2}$$

EXAMPLE

Consider the completion time of a large commercial development. The proportion of the maximum allowed time to complete a task is modeled as a beta random variable with alpha= 2.5 and beta=1. What is the probability that the proportion of the maximum time exceeds 0.7?

$\alpha = 2.5 \quad \beta = 1$

$$P(x > 0.7) = \int_{0.7}^{1} f_x(x) = \int_{0.7}^{1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx$$

$$= \int_{0.7}^{1} \frac{\Gamma(3.5)}{\Gamma(2.5)\Gamma(1)} x^{1.5}(1-x)^{0}\,dx$$

$\Gamma(n+1) = n\Gamma(n)$

$\Gamma(3.5) = 2.5\,\Gamma(2.5)$
$= 2.5 \times 1.5\,\Gamma(1.5)$

$$= \frac{2.5\,(1.5)\,\Gamma(0.5)}{(1.5)\,(0.5)\,\Gamma(0.5)} \int_{0.7}^{1} x^{1.5}\,dx$$

$= 0.59$

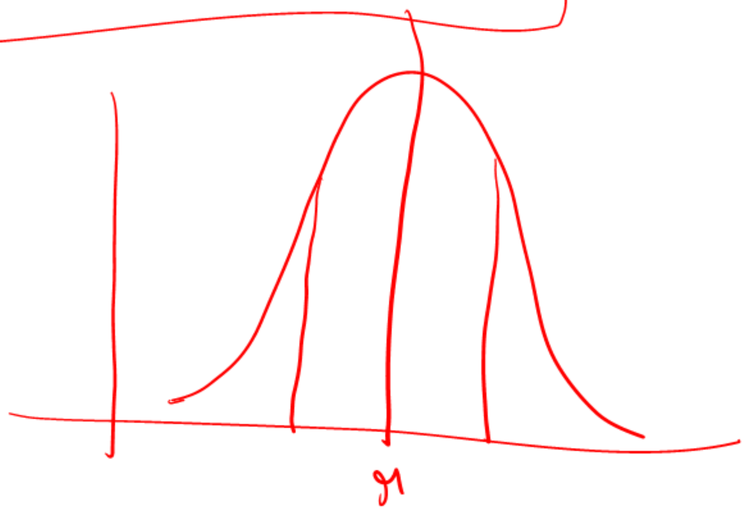
Normal distribution or Gaussian distribution

A continuous r.v x is said to have normal distribution with mean μ, variance σ^2

$X \sim N(\mu, \sigma^2)$ if has pdf

$$\boxed{f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

$$\left\{ \begin{array}{r} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma > 0 \end{array} \right\}$$



$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$Z = \frac{x-\mu}{\sigma}$$

$$dz = \frac{1}{\sigma} dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \cdot \sigma\, dz$$

$$\frac{1}{2} z^2 = t$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\, dz$$

$$\frac{1}{2} \cdot 2z\, dz = dt$$

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad x \in [0,1)$$

$\alpha > 0$

$\beta > 0$

$\Rightarrow$ E[x] for gamma function

$$E[X] = \int_0^\infty x \frac{\lambda^\gamma x^{\gamma-1} e^{-\lambda x}}{\Gamma(\gamma)} dx$$

$$\int_0^\infty x^{\gamma-1} e^{-x}$$

$$= \int_0^\infty \frac{\lambda^\gamma x^\gamma e^{-\lambda x}}{\Gamma(\gamma)} dx$$

$\lambda x = y$

$\lambda dx = dy$

$$= \int_0^\infty \frac{\lambda^\gamma \left(\frac{y}{\lambda}\right)^\gamma e^{-y}}{\Gamma(\gamma)} \frac{dy}{\lambda} = \int_0^\infty \frac{y^\gamma e^{-y}}{\lambda\Gamma(\gamma)} dy$$

$$\frac{\Gamma(\gamma+1)}{\lambda\Gamma(\gamma)} = \frac{\gamma\cancel{\Gamma(\gamma)}}{\lambda\cancel{\Gamma(\gamma)}} = \frac{\gamma}{\lambda}$$

$E[X] = \gamma/\lambda$, $\quad Var(X) = \gamma/\lambda^2$

$$2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-t} \frac{1}{\sqrt{2t}} \, dt$$

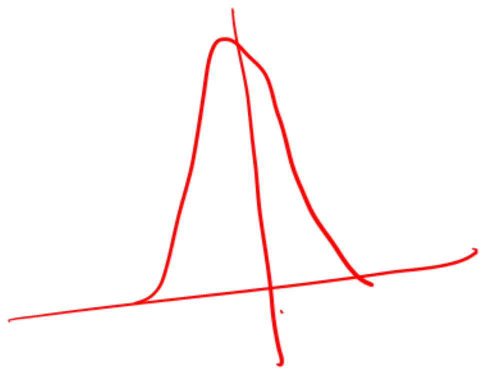
$$t^{1/2-1}$$

$$= \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-t} t^{-1/2} \, dt = \frac{1}{\sqrt{\pi}} \Gamma(1/2)$$

$$= \frac{1}{\sqrt{\pi}} \Gamma(1/2) = \frac{\sqrt{\pi}}{\sqrt{\pi}} = 1$$

Standard normal distribution

$\mu = 0$, $\sigma^2 = 1$



$f_X(x)$

MGF

$$E[e^{tx}] = \int e^{tx} f(x)$$

$$\boxed{M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}}$$

Distribution	PDF $f_X(x)$	Mean $E[X]$	Variance $Var(X)$	MGF $M_X(t)=E[e^{tX}]$
Continuous Uniform Distribution $X \sim U(a,b)$	$\frac{1}{b-a}$ $a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{t(b-a)}$, if $t \neq 0$; 0 if $t=0$
Gamma Distribution $X \sim gamma(r,\lambda)$	$\frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}$ for $x>0$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$\left(1-\frac{t}{\lambda}\right)^{-r}$, $t<\lambda$
Exponential distribution $\lambda>0$ $X \sim exp(\lambda)$	$\lambda e^{-\lambda x}$, $x>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\left(1-\frac{t}{\lambda}\right)^{-1}$; $t<\lambda$
Beta distribution $\alpha>0, \beta>0$ $X \sim Beta(\alpha,\beta)$	$\frac{\Gamma(\alpha+\beta) x^{\alpha-1}(1-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)}$ $0<x<1$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	
Normal Distribution $X \sim N(\mu,\sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ $-\infty<x<\infty$	μ	σ^2	$e^{\mu t+\frac{1}{2}\sigma^2 t^2}$
Standard Normal distribution $X \sim N(0,1)$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ $-\infty<x<\infty$	0	1	$e^{\frac{1}{2}t^2}$

Characteristic equation

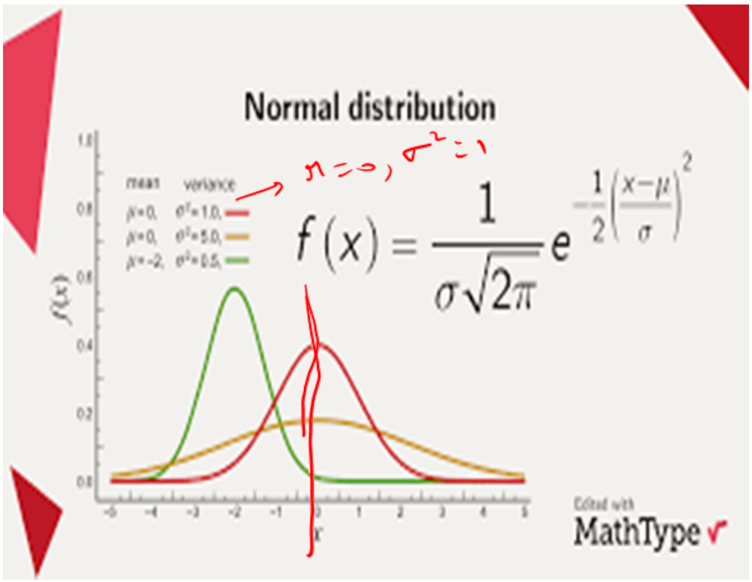
(1) $\phi(t) = E\left[e^{itx}\right]$

$e^{it} =$

Normal Distribution

$$X \sim N(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

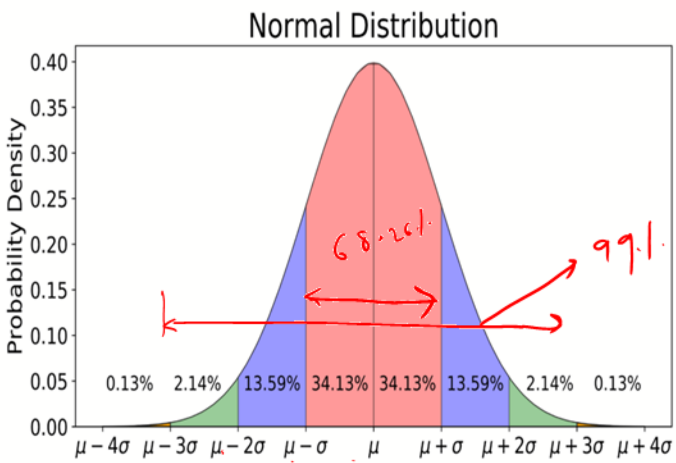


Normal distribution is symmetrical about mean

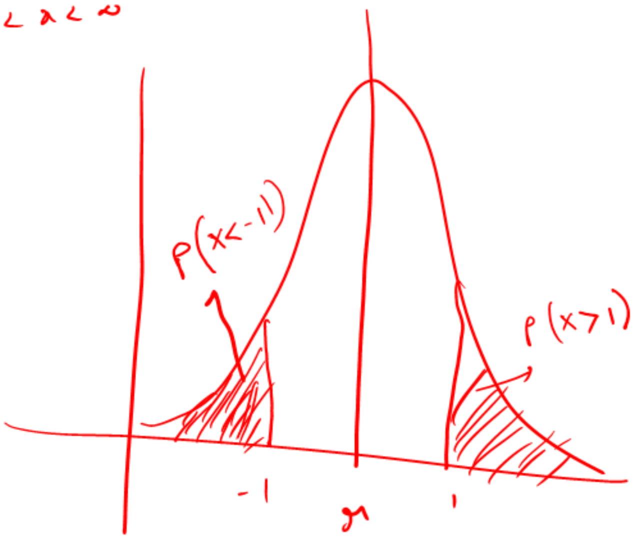
$\beta_1 = 0$ coeff skewness

$\beta_2 = 0$ coeff excess kurtosis

$-\infty < x < \infty$



92%



$$P(X < x) = \int_{-\infty}^{x} f_X(x)\,dx$$

$$X \sim N(\mu, \sigma^2)$$

Standard Normal distribution

Theorem

If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$ then

$$\boxed{Y \sim N(a\mu + b,\ a^2\sigma^2)}$$

$E[X] = \mu$

$Var[X] = \sigma^2$

$E[Y] = E[aX + b]$

$= a\mu + b$

$Var[Y] = a^2 Var(X)$

$= a^2\sigma^2$

$\Rightarrow \mu = 0,\ \sigma^2 = 1 \qquad Z \sim N(0, 1)$

$$\boxed{Z = \frac{X - \mu}{\sigma}} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$$

$X \sim N(\mu, \sigma^2)$

$a = \frac{1}{\sigma}, \quad b = -\frac{\mu}{\sigma}$

$E[Z] = \frac{1}{\sigma}E[X] - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$

$Var(Z) = a^2 Var(X) = a^2\sigma^2 = \frac{1}{\sigma^2} \times \sigma^2 = 1$

$P(a \le X \le b)$

$= \int$

Conv al

'Z'

ex:-

Most graduate schools of business require applicants for admission to take the Graduate Management Admission Council's GMAT examination. Scores on the GMAT are roughly normally distributed with a mean of 527 and a standard deviation of 112. What is the probability of an individual scoring above 500 on the GMAT?

$$X \sim N(527, (112)^2)$$

$$P(X > 500) =$$

$$P\left(\frac{X-\mu}{\sigma} > \frac{500-\mu}{\sigma}\right) = P\left(Z > \frac{500-527}{112}\right)$$

$$P(Z > -0.24) = 1 - P(Z < -0.24)$$

$$= 1 - 0.40517$$

$$= 0.5948$$

ex:-

The average number of acres burned by forest and range fires in a large New Mexico county is 4,300 acres per year, with a standard deviation of 750 acres. The distribution of the number of acres burned is normal. What is the probability that between 2,500 and 4,200 acres will be burned in any given year?

$$X \sim N(4300, (750)^2)$$

$$P(2500 < X < 4200)$$

$$P\left(\frac{2500-4300}{750} < Z < \frac{4200-4300}{750}\right)$$

$$P(-2.4 < Z < -0.13)$$

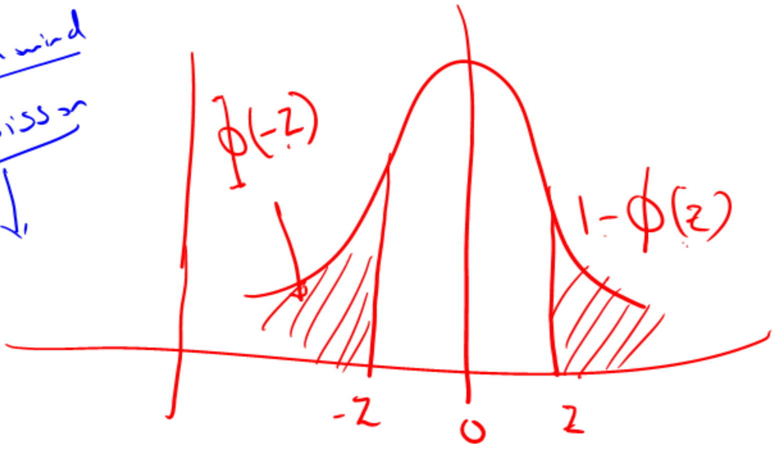
$$P(Z < -0.13) - P(Z < -2.4)$$

$$= 0.44828 - 0.00820$$

$$= 0.44$$

Binomial

Poisson

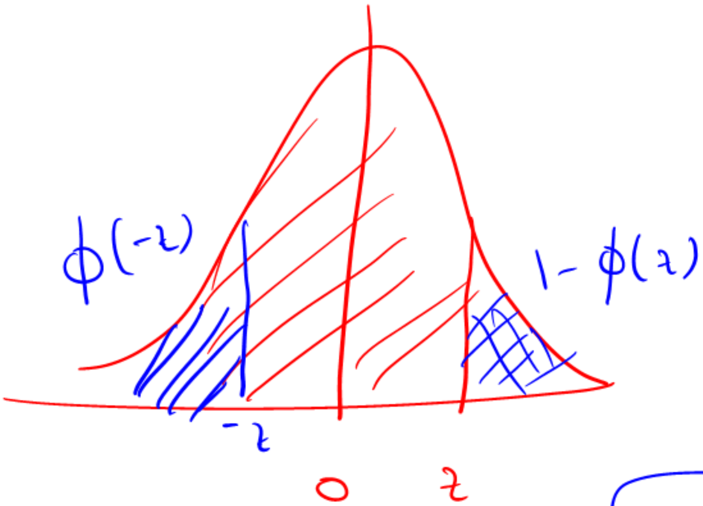


pdf of Standard normal distribution

$$\boxed{\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$$

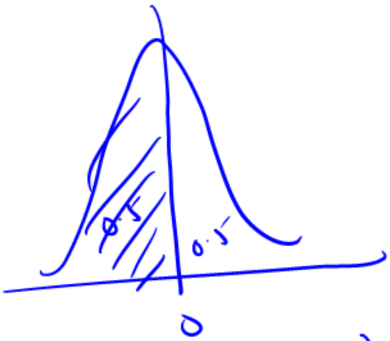
$$\Phi(-z) = P(Z < -z) = \int_{-\infty}^{-z} \phi(z)\,dz$$

$$\Phi(z) = \overbrace{P(Z < z)}^{F_Z(z)} \leq \int_{-\infty}^{z} \phi(z)\,dz$$



$\phi(z)$

$$\boxed{\phi(-z) = 1 - \phi(z)}$$



$$\phi(-z) + \phi(z) = 1$$

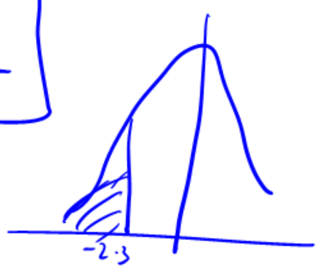
$z = 0$

$P(z<0)$

$$\boxed{\phi(0) = 1/2}$$

$P(-2.3)$

$P(Z < -2.3)$



$X \sim N(25, 36)$

$\mu = 25$

$\sigma^2 = 36$

$\sigma = 6$

$P(X < 28)$

$Z = \frac{X-\mu}{\sigma}$

$P\left(\frac{X-\mu}{\sigma} < \frac{28-\mu}{\sigma}\right) = P\left(Z < \frac{28-25}{6}\right)$

$P(Z < 0.5)$

$= 0.69146$

0.83 0.84

$P(X > 30) = 1 - P(X < 30)$

$1 - P\left(Z < \frac{30-25}{6}\right) = 1 - P(Z < 0.833)$

$P(0.836$

$1 - P(Z < 0.83)$

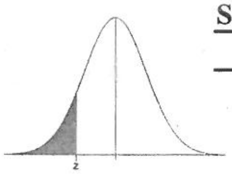
$1 - 0.79673 =$

$P(25 < X < 30)$

$F_X(30) - F_X(25)$

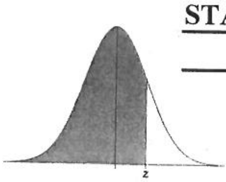
$P(X < 30) - P(X < 25) =$

asc Standard Normal Distribution Tables



STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
-0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414



STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997

Unit 3

Jointly Distributed Random Variable-Conditional Distributions:
Discrete and Continuous case, Joint and Marginal Distribution, Covariance and Correlation, Bivariate and Multivariate Normal Distribution, Functions of random variables and random vectors, distributions of sums of random variables;

Jointly Distributed Random Variables

In many engineering applications situations arise where it is necessary to make use of more than one random variable

Picking card from a deck of 52 card

$X \sim$ corresponding suit card

$X(\text{club}) = 0$

$X(\text{diamond}) = 1$

$X(\text{spade}) = 2$

$X(\text{heart}) = 3$

$Y \sim$ corresponding to denomination of card

$Y(\text{ace}) = 1$

$Y(2) = 2$

$Y(10)$

$Y(J)$

K

Q $= 13$

$(X, Y) \rightarrow$ bivariate

$X, Y, Z \rightarrow$ trivariate

$X, Y, Z \ldots \ldots \rightarrow$ multivariate distribution

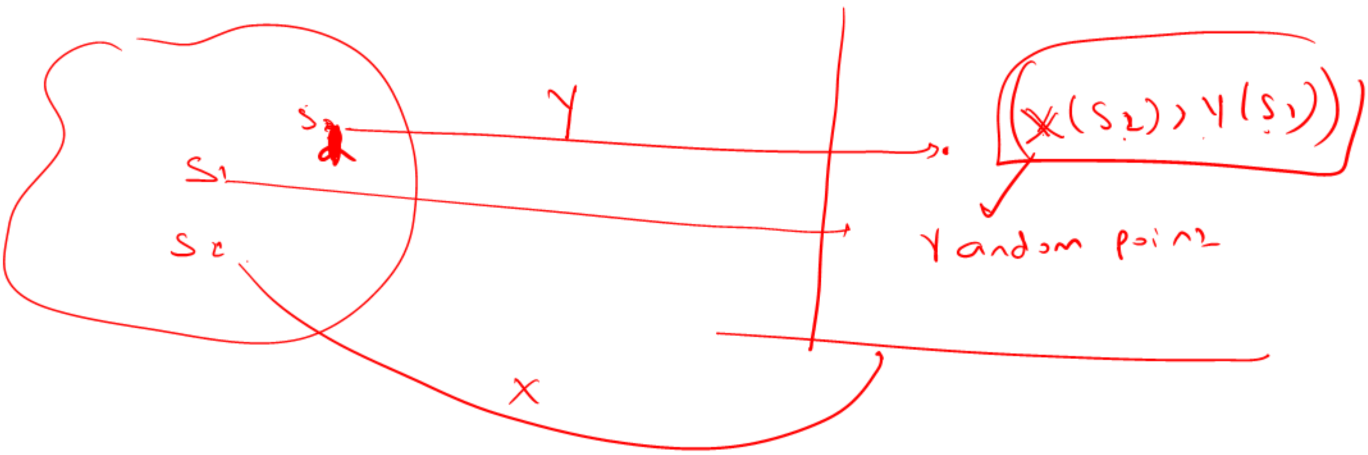
* When outcome of r.v can be characterized in more than 1 way

$X_1 \sim 0, 1, 2, 3 \rightarrow x_1$ (

$X_2 \sim 1, 2, \dots 13 \quad x_2$

Ordered pair $= (x_1, x_2) \rightarrow$ random point

$(0,1) \quad (0,2) \quad (1,1)$



$X($

random vector

$X_1, X_2, \dots X_N$

If N random variables defined on sample space Ω. we

$X = \{X_1, X_2 \dots X_N\}$

$X, Y \rightarrow R^2$

$X_1 =$ time

$X_2 =$ pulse rate

$X_3 =$ BP

$X = \{x_1, x_2, x_3\}$

Recall Univariate r.v

Discrete RV

1. Probability mass function (PMF)

$p_X(x_i) = P(X = x_i)$

1. $p_X(x) \geq 0$

2. $\sum_{\forall x_i} p_X(x_i) = 1$

Continuous R.V

Probability density fn

PDF

$P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

1. $f_X(x) \geq 0$

2. $\int_{-\infty}^{\infty} f_X(x) = 1$

Jointly distributed PMF

X, Y are r.v on sample space

Probability function

1. $p_{X,Y}(x_i, y_i) = P(X = x_i, Y = y_i)$

2. $p_{X,Y}(x_i, y_i) \geq 0$

3. $\sum \sum p_{X,Y}(x_i, y_i) = 1$

Jointly distributed density function

1. $P(a < x < b, c < y < d)$

$$= \int_a^b \int_c^d f_{XY}(x,y)\,dx\,dy$$

2. $f_{XY}(x,y) \geq 0$

3. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy$

Jointly distributed probability function pmf

Discrete $p_{xy}(x_i, y_i) = P(X = x_i, Y = y_i)$

Pdf

Continuous $f_{xy}(x, y)$;

→ Marginal probability distribution $p_x(x_i)$

$p_y(y_i)$

let X, Y be 2 jointly distributed random variables with joint pmf $p_{xy}(x_i, y)$

→ The individual distribution of X, Y are called marginal distribution of X, Y

→ Marginal probability of function of X is $p_x(x)$ (or) $P(x)$

$$\rightarrow \quad P_x(x) = (X = x) = \sum_{y} P(X = x, Y = y)$$

$$= \sum_{y} P(x, y)$$

$$p_x(x_i) = \sum_{y_i} P(X = x_i, y_i)$$

⟹ The marginal probability of y is denoted by $p_y(y)$ (or) $p(y)$

$$p_y(y) = (Y=y) = \sum_{x} p(X=x, Y=y)$$

$p(X=x_1, Y=y_1)$

Bivariate Probability distribution table

X \ Y	y_1	y_2	y_3	...	y_m	$p_x(x)$
x_1	p_{11}	p_{12}	p_{13}	...	p_{1m}	$p_x(x_1)$ → row total
x_2	p_{21}	p_{22}	p_{23}	...	p_{2m}	$p_x(x_2)$
x_3	⋮					
⋮	⋮					
x_n	p_{n1}	p_{n2}	...		p_{nm}	$p_x(x_3)$
$p_y(y)$	$p_y(y_1)$	$p_y(y_2)$			$p_y(y_m)$	1

column total

CDF Recall Univariate

Discrete

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

1. $\lim_{x \to -\infty} F_X(x) = 0$

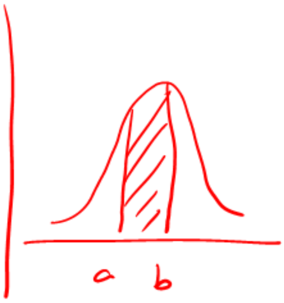
   $\lim_{x \to \infty} F_X(x) = 1$

2. non decreasing

Continuous

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$P(a \le x \le b) = F_X(b) - F_X(a)$$



$$= \int_a^b f(t)\,dt$$

$$P(a < x < b,\ c < y < d)$$

$$= \iint f_{XY}(x,y)\,dx\,dy$$

Jointly distributed CDF

$$F_{XY}(-\infty,-\infty) = 0$$

$$F_{XY}(\infty,\infty) = 1$$

$$F_{XY}(-\infty, x) = 0$$

Non decreasing

$$F_{XY}(x, -\infty) = 0$$

$$F_X(x_1 \le X \le x_2,\ y_1 \le Y \le y_2) = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f_{XY}(x,y)\,dx\,dy$$

Joint CDF

Marginal density function

X, Y are jointly distributed r.v.

Marginal PDF of X

$$f_X(x) = \int_{-\infty}^{\infty} f_X(x,y)\, dy$$

Marginal PDF of Y

$$f_Y(y) = \int_{-\infty}^{\infty} f_X(x,y)\, dx$$

Ex-suppose a car showroom has 10 cars of a brand out of which 5 are good, 2 are of defective transmission(DT) and 3 have defective steering and good ones are denoted by G.Now if 2 cars are selected at random.

Let X denote the number of cars with DT $\rightarrow$ 0, 1, 2

Y denote the numbe of cars with DS $\rightarrow$ 0, 1, 2

Find joint PMf

CDF

MPF

10 $\rightarrow$ 5 Good, 2 DT, 3 DS

$P_{XY}(x,y) = P(X=x, Y=y)$

$P_{XY}(0,0) = P(X=0, Y=0)$

$= \dfrac{{}^5C_2}{{}^{10}C_2} = \dfrac{2}{9}$

X \ Y	0	1	2	
0	$\frac{2}{9}$	$1/3$	$1/15$	$P_X(0)$
1	$2/9$	$6/45$	0	$P_X(1)$
2	$1/45$	0	0	$P_X(2)$
	$P_Y(0)$	$P_Y(1)$	$P_Y(2)$	1

$P_{XY}(0,1)$

$= \dfrac{{}^5C_1 \times {}^3C_1}{{}^{10}C_2}$

$= 1/3$

$P_{XY}(0,2)$ = 2 DS cars $= \dfrac{{}^3C_2}{{}^{10}C_2} = 1/15$

$P_{XY}(1,0) =$

$P(X \le 1, Y \le 1) = P(X=0, Y=0) + P(X=0, Y=1)$

$+ P(X=1, Y=0) + P(X=1, Y=1)$

$= 41/45$

$P(X < 2)$

Conditional pmf/pdf

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)}$$

If X and Y are jointly distributed random variables

The conditional probability that $X = x_i$ given $Y = y_i$

is given by

Discrete case   X, Y are D r.v's

$$p_{\frac{X}{Y}}(x_i, y_i) = \frac{P(X = x_i, Y = y_i)}{P(Y = y_i)} = \boxed{\frac{p_{XY}(x_i, y_i)}{p_Y(y_i)}} ; \quad p_Y(y_i) > 0$$

$$p_{\frac{Y}{X}}(x_i, y_i) = \frac{P(X = x_i, Y = y_i)}{P(X = x_i)} = \frac{p_{XY}(x_i, y_i)}{p_X(x_i)}$$

Continuous case   pdf   X, Y are c r.v

pdf :- $f_{XY}(x, y)$

$$f_{\frac{X}{Y}}(x, y) = \frac{f_{XY}(x, y)}{f_Y(y)}$$

$$f_{\frac{Y}{X}}(x, y) = \frac{f_{XY}(x, y)}{f_X(x)}$$

As we know

$$P(A \cap B) = P(A/B)\, P(B) = P(B/A)\, P(A)$$

$$p_{XY}(x_i, y_i) = p_{\frac{X}{Y}}(x_i, y_i)\, p_Y(y_i) = p_{\frac{Y}{X}}(x_i, y_i)\, p_X(x_i)$$

$$f_{XY}(x, y) = f_{\frac{Y}{X}}\, f_X(x) = f_{\frac{X}{Y}}\, f_Y(y)$$

Independent:

If X,Y are jointly distributed random variables.Then X and Y are independent if

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

$$f_{XY}(x,y) = f_X(x) f_Y(y) \quad \text{continuous case}$$

$$p_{XY}(x,y) = p_X(x) p_Y(y)$$

Example-1

An insurance agency has customers with both home and auto policy.
For each type of policy, a deductible amount must be specified.
For auto policy, choices are $100 and $250,
for home policy, choices are $0, $100, and $200.
Suppose a customer is selected at random. Let:
X = his deductible on the auto policy
Y = his deductible on the home policy

$p_x(100) =$

Suppose the joint pmf is given by the insurance company in the accompanying joint probability table:

$p(x, y)$		y: 0	100	200	$p_x(x)$	
x	100	.20	.10	.20	0.5	$\to p_x(100)$
	250	.05	.15	.30	0.5	$\to p_x(250)$
$p_y(y)$		$p_y(y)=0.25$	0.25	0.5		

1) find $P_{Y/X}\left(\frac{Y=0}{X=100}\right)$ $P_y(200)$ 2) $P_{X/Y}\left(\frac{X=100}{Y=200}\right)$

$$P_{Y/X}\left(\frac{Y=0}{X=100}\right) = \frac{P(X=100, Y=0)}{p_x(100)} = \frac{0.20}{0.50} = \frac{2}{5}$$

$$P_{X/Y}\left(\frac{X=100}{Y=200}\right) = \frac{P(X=100, Y=200)}{P_y(200)} = \frac{0.20}{0.50} = \frac{2}{5}$$

X, Y are they independent ??

$$P(X=100, Y=100) = p_x(100)\, p_y(100)$$

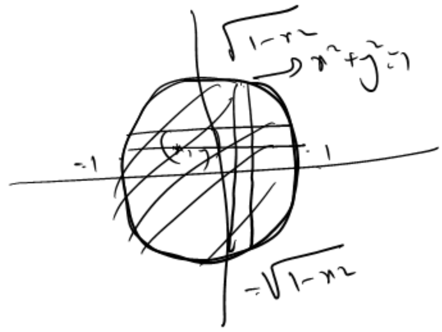
$$= 0.10 \neq (0.5)(0.25$$

X, Y are not independent

2) If x and y are jointly distributed given by pdf

$$f_{xy}(x,y) = \begin{cases} \frac{1}{\pi} \; ; & x^2+y^2 \le 1 \\ 0 & \text{outside} \end{cases}$$

find $f_{\frac{y}{x}}(x,y) = \dfrac{f_{xy}(x,y)}{f_x(x)}$



$x^2+y^2=1$

$y^2 = 1-x^2$

$y = \pm\sqrt{1-x^2}$

$$f_x(x) = \int_y f_{xy}(x,y)\, dy$$

$$= \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\, dy$$

$$= \frac{1}{\pi}\, y \Big|_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} = \frac{2\sqrt{1-x^2}}{\pi}, \quad -1 \le x \le 1$$

$$f_y(y) = \int_x f_{xy}(x,y)\, dx = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}$$

$$f_{Y/X}\left(\frac{y}{x}\right) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{\frac{1}{\pi}}{\frac{2\sqrt{1-x^2}}{\pi}} = \underline{\underline{\frac{1}{2\sqrt{1-x^2}}}}$$

X, Y are independent ??

$$f_{XY}(x,y) = f_X(x)\, f_Y(y)$$

$$\frac{1}{\pi} \neq \frac{2\sqrt{1-x^2}}{\pi} \cdot \frac{2\sqrt{1-y^2}}{\pi}$$

X, Y are not independent.

5)Suppose X and Y both take values in [0,1] with density f(x, y) = 4xy. Show f(x, y) is a valid joint pdf, If event A = 'X < 0.5 and Y > 0.5' find its probability

b)Find its joint cdf

$$f_{XY}(x,y) = 4xy \qquad 0 \le x \le 1, \quad 0 \le y \le 1$$

$\Rightarrow$ check valid pdf $\quad f_{XY}(x,y) \ge 0$

$$\int_0^1 \int_0^1 4xy \, dx \, dy = \int_0^1 4y \frac{x^2}{2} \Big|_0^1 dy = \int_0^1 2y \, dy$$

$$= \frac{2y^2}{2} \Big|_0^1 = 1$$

Joint cdf

$$F_{XY}(x,y) = P(X \le x, Y \le y) = \int_0^x \int_0^y 4xy \, dy \, dx$$

$$= \int_0^x 4x \frac{y^2}{2} dx = 2\frac{x^2}{2} y^2 \Big|_0^x = x^2 y^2$$

$$f_{\underset{X=x}{Y}}(x,y) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{4xy}{2x} = 2y$$

$$f_X(x) = \int_y f_{XY}(x,y) \, dy = \int_0^1 4xy \, dy = 4x\frac{y^2}{2}\Big|_0^1$$

$$= 2x, \quad 0 < x < 1$$

$$f_Y(y) = \int_x f_{XY}(x,y) \, dx = \int_0^1 4xy \, dx = 2y, \quad 0 < y < 1$$

$$f_{\frac{x}{y}}(x,y) = \frac{f_{xy}(x,y)}{f_y(y)} = \frac{4xy}{2y} = 2x$$

$\Rightarrow$ check whether x, y are independent.

$$f_{xy}(x,y) = f_x(x)\, f_y(y)$$

$$4xy = 2x \times 2y$$

$$= 4xy$$

X, Y are independent

Product moments

Moments — Univariate

Central moment: $E\left[(X-\mu)^k\right]$

Non central moment: $E\left[X^k\right]$ about 0

$E\left[(X-c)^k\right]$ about c

⇒ Expected value of two variables

X, Y are jointly distributed r.v

$$E\left[h(x,y)\right] = \begin{cases} \sum_{x}\sum_{y} h(x,y)\, p_{xy}(x,y) & \text{Discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\, f_{xy}(x,y)\, dx\, dy & \text{for continuous} \end{cases}$$

Product moments

$\mu'_{r,s} = E\left[X^r Y^s\right] = (r,s)^{th}$ non central moment

$\mu_{r,s} = E\left[(X-\mu_X)^r (Y-\mu_Y)^s\right] \rightarrow (r,s)^{th}$ central moment

$$E[x^r y^s] = (r,s)^{th}$$

$$E[x] = \mu'_{1,0}$$

$$E[x y^0] = E[x]$$

$$\mu'_{1,1} = E[xy]$$

$$E[y] = \mu'_{0,1}$$

$\Rightarrow$ Central moments

$$\mu_{1,1} = E[(x-\mu_x)(y-\mu_y)] = \text{Covariance}(x,y)$$

Covariance: when the two random variables are not independent.it is frequently of the interest to assess how strongly they are related to one another

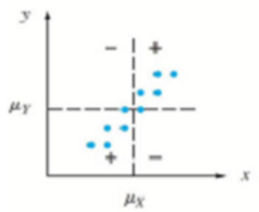
$$\mu_{1,1} = E[(X-\mu_X)(Y-\mu_Y)] = \mathrm{Cov}(X,Y) = \begin{cases} \sum\sum (x-\mu_x)(y-\mu_y)\,p_{xy}(x,y) & \text{Discrete r.v's} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y)\,f_{xy}(x,y)\,dx\,dy & \text{Continuous R.v} \end{cases}$$

$$= E[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y]$$

$$= E[XY] - \mu_X E[Y] - \mu_Y \underbrace{E[X]}_{\mu_X} + \cancel{\mu_X}\mu_Y$$

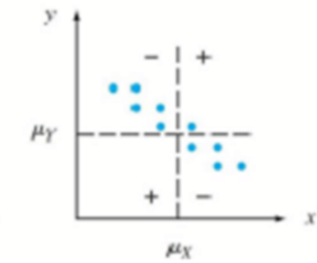
$$\mathrm{Cov}(X,Y) = E[XY] - \mu_X\mu_Y$$

If both variables tend to deviate in the same direction (both go above their means or below their means at the same time), then the covariance will be positive.



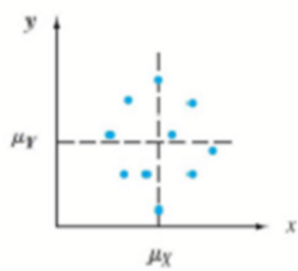
Positive covariance

If they tend to deviate in the opposite direction from their means at the same time – ie, the signs of (x –) and (y –) tend to be opposite the product (and covariance) will be negative



Negative covariance

If X and Y are not strongly related, positive and negative products will tend to cancel one another, yielding a covariance near 0.



covariance near 0

Properties of Covariance

1. $\text{Cov}(X,X) = \text{Var}(X) = E\left[\left(X-\mu_X\right)\left(X-\mu_X\right)\right] = E\left[\left(X-\mu_X\right)^2\right]$

2. $\text{Cov}(X,Y) = \text{Cov}(Y,X)$.

3. $\text{Cov}(X,c) = 0$, c is constant

4. $\text{Cov}(cX,Y) = c\,\text{Cov}(X,Y)$

5. $\text{Cov}(X, Y+Z) = \text{Cov}(X,Y) + \text{Cov}(X,Z)$

6. $\text{Cov}(X+Y, Z+W) = \text{Cov}(X,Z) + \text{Cov}(X,W) + \text{Cov}(Y,Z)$

$$+ \text{Cov}(Y,W)$$

7. $\text{Cov}\left(\sum_{i=1}^{m} a_i x_j, \sum_{j=1}^{n} b_j y_j\right) = \sum_{i,j} a_i b_j \,\text{Cov}(x_i, y_j)$

8. $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2) + 2\,\text{Cov}(X_1, X_2)$

9. $\text{Var}(X_1 + X_2 + \cdots X_n)$

$$= \text{Var}(X_1) + \text{Var}(X_2) \cdots + \text{Var}(X_n)$$

$$+ 2\sum_{i<j} \text{Cov}(x_i, y_j)$$

Correlation:

Covariance (x, y)

$$\text{Corr}(x, y) = \rho(x, y) = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y}$$

Interpretation of correlation

$\Rightarrow$ $$-1 \le \rho(x, y) \le 1$$

1. If $\rho_{XY} = 1$, then X and Y are perfectly, positively, linearly correlated.
2. If $\rho_{XY} = -1$, then X and Y are perfectly, negatively, linearly correlated.
3. If $\rho_{XY} = 0$, then X and Y are completely, un-linearly correlated. That is, X and Y may be perfectly correlated in some other manner, in a parabolic manner, perhaps, but not in a linear manner.
4. If $\rho_{XY} > 0$, then X and Y are positively, linearly correlated, but not perfectly so.
5. If $\rho_{XY} < 0$, then X and Y are negatively, linearly correlated, but not perfectly so.

* X, Y are independent $\text{cov}(x, y) = 0$

As we know $E[xy] = E[x] E[y]$ when x & y are independent

$\text{Cov}(x, y) = E[xy] - \mu_x \mu_y = E[x] E[y] - \mu_x \mu_y = \mu_x \mu_y - \mu_x \mu_y$

$= 0$

$X, Y,$

$Y = g(X)$

X_1, X_2, X_3

$Y = g(X_1, X_2)$, $Z = g(X_1, X_2, \dots X_n)$

$\Longrightarrow$ X, Y are random Variables with pmf

$x \backslash y$	1	2	3	$P_X(x)$
1	0.25	0.25	0	$0.5 \rightarrow P_X(1)$
2	0	0.25	0.25	$0.5 \rightarrow P_X(2)$
$P_Y(y)$	0.25 $P_Y(1)$	0.5 $P_Y(2)$	0.25 $P_Y(3)$	

Find $\operatorname{cov}(X, Y)$, $\operatorname{corr}(X, Y) = \rho_{XY}$

$$\operatorname{Cov}(X, Y) = E[XY] - \mu_X \mu_Y$$

$$\mu_X = E[X] = \sum_x x \, P_X(x) = 1 \times 0.5 + 2 \times 0.5$$
$$= 1.5$$

$$\mu_Y = E[Y] = \sum_y y \, P_Y(y) = 1 \times 0.25 + 2 \times 0.5 + 3 \times 0.25$$
$$= 2$$

$$E[XY] = \sum_x \sum_y xy \, P_{XY}(x, y) = 1 \times 1 \times 0.25 + 2 \times 0.25 + 0 + 0$$
$$+ 4 \times 0.25 + 6 \times 0.25$$
$$= 3.25$$

$$\mathrm{cov}(x,y) = E[xy] - \mu_x \mu_y$$

$$= 3.25 - 1.5 \times 2 = \underline{\underline{0.25}}$$

$$\mathrm{Corr}(x,y) = \rho_{xy} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}$$

$$\sigma_x = \sqrt{0.25} \qquad \sigma_y = \sqrt{0.5}$$

$$\mathrm{Var}(x) = E[x^2] - \underbrace{(E[x])^2}_{\mu_x^2} = 2.5 - (1.5)^2$$

$$= \underline{\underline{0.25}}$$

$$E[x^2] = \sum x^2 p_x(x) = 1^2 \times 0.5 + 2^2 \times 0.5$$

$$= 2.5$$

$$E[y^2] = \sum y^2 p_y(y) = 4.5$$

$$\mathrm{Var}(y) = E[y^2] - \mu_y^2 = 4.5 - 2^2 = 0.5$$

$$\mathrm{corr}(x,y) = \rho_{xy} = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y} = \frac{0.25}{\sqrt{0.25}\sqrt{0.5}} = \underline{\underline{\sqrt{0.5}}}$$

X, Y are Positively not linearly Perfect relation

$\Rightarrow$ Let $X \sim U(-1,1)$ $\rightsquigarrow$ CRV Let $Y = X^2$ Then. $X \sim U(a,b)$

$f_X(x) = \frac{1}{b-a}$

$\mathrm{cov}(X,Y)$

$$f_X(x) = \frac{1}{1-(-1)} = \frac{1}{2} \quad -1 < x < 1$$

$-1 < y < 1$

$$\mathrm{cov}(X,Y) = E[XY] - \mu_X \mu_Y$$

$$\mu_X = E[X] = \int_a x f_X(x)\,dx = \int_{-1}^{1} x \cdot \frac{1}{2}\,dx = 0$$

$$\mu_Y = E[Y] = E[X^2] = \int x^2 f_X(x)\,dx = \int_{-1}^{1} x^2 \cdot \frac{1}{2}\,dx$$

$$E[Y] = \int y\, f_Y(y)$$

$$= \frac{1}{3}$$

$$E[XY] = E[X \cdot X^2] = E[X^3] = \int_{-1}^{1} x^3 f_X(x)\,dx = 0$$

$$\mathrm{cov}(X,Y) = E[XY] - \mu_X \mu_Y = 0 - 0 = 0$$

$$\mathrm{corr}(X,Y) = \rho_{XY} = 0$$

Functions of random Variables

Discrete

1. unif
2. geo

Continuous

$Y = g(X)$ X distribution is known

$Y = g(X_1, X_2)$

$Y = g(X_1, X_2 \cdots X_n)$

1. There are 3 types of techniques

1. CDF technique → simple
2. Transformation technique → General technique
3. Moment generating function technique

Y sum of r.v's or

X_1, X_2 linear combination of r.v which are iid

$Y = X_1 + X_2$

ex:- let X be random Variable with

pdf $f_X(x) = 2e^{-2x}$, $x > 0$

Determine pdf of $Y = \sqrt{X}$

CDF technique

$$F_X(x) \longrightarrow F_Y(y)$$

$$F_X(x) = P(X \leq x) = \int_0^x 2e^{-2x}\,dx \quad , \quad x > 0$$

$$F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2)$$

$$= \int_0^{y^2} f_X(x)\,dx = \int_0^{y^2} 2e^{-2x}\,dx$$

$$-e^{-2x}\Big|_0^{y^2} = 1 - e^{-2y^2}$$

$$F_Y(y) = 1 - e^{-2y^2}$$

Find pdf $\quad f_Y(y) = F'_Y(y) = \frac{d}{dy}\left(1 - e^{-2y^2}\right)$

$$f_Y(y) = 4y\,e^{-2y^2} \quad , \quad y > 0$$

ex:- let $X \sim geo(p)$ for $x = 1, 2, \ldots$

find pmf of $Y = X^2$ $\longrightarrow$ $y = 1, 4, 9, 16 \ldots$

$\rightarrow$ $X \sim geo(p)$

$\rightarrow$ pmf $p_X(x) = p\,q^{x-1} = p\,(1-p)^{x-1}$

$\rightarrow$ CDF $F_X(x) = P(X \le x) = 1 - (1-p)^{x}$

$P(X=1) + P(X=2) \ldots P(X=$

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(X \le \sqrt{y})$$

f

$$= F_X(\sqrt{y}) = 1 - (1-p)^{\sqrt{y}}$$

pmf:-

$p_Y(y) = ?$

→ let X be r.v having uniform distribution $\{1, 2, \dots n\}$

$P_X(X=x) = \frac{1}{n}$ let $Y = X + a$

find $P_Y(y)$

Transformation technique

→ Discrete random Variable

$$\boxed{p_X(x) \longrightarrow p_Y(y)}$$

Let X be DRV with PMF $p_X(x)$ on A.

Let $Y = g(X)$. be 1-1 function from A to B

If $X = g^{-1}(y)$ is inverse transformation for every $y \in B$. then PMF of Y is

$$p_X(x) = P(X = x)$$

$$p_Y(y) = P(Y = y) = P(g(X) = y) = P(X = g^{-1}(y))$$

$$= p_X(g^{-1}(y)) \quad \text{for } y \in B$$

ex:- Let $X \sim geo(p)$, $x = 1, 2, \ldots\ldots$ find PMF of $Y = X^2$

$X = 1$

$X = -1$

$X \sim geo(p)$

$$p_X(x) = p(1-p)^{x-1} \quad \text{for } x = 1, 2, \ldots\ldots$$

$A = \{x \mid x = 1, 2, \ldots\}$

$Y = X^2$ is 1-1 transformation from A to B

$Y = g(X) = X^2$ $\quad B = \{y \mid y = 1, 4, 9, \ldots\}$

$$X = g^{-1}(y) = \sqrt{y}$$

$$p_Y(y) = p_X(g^{-1}(y)) = p_X(\sqrt{y}) = p(1-p)^{\sqrt{y}-1}$$

for $y = 1, 4, 9, \ldots$

bivariate Discrete R.V $\quad p_{X_1,X_2}(x_1,x_2) \longrightarrow p_{Y_1,Y_2}(y_1,y_2)$

let X_1 and X_2 are DRV. with Joint pmf $p_{XY}(x,y)$. defined A:

$p_{Y_1}(y_1)$

$Y_1 = g_1(X_1, X_2) \leftarrow$

$Y_2 = g_2(X_1, X_2) \leftarrow$ define by yourself

y_1 and y_2 are 1-1 transformation defined on $A \to B$

$p_{Y_1}(y_1) = ?$

$x_1 = g_1^{-1}(y_1, y_2)$

$x_2 = g_2^{-1}(y_1, y_2)$

$$p_{Y_1,Y_2}(y_1,y_2) = p_{X_1X_2}\left(g_1^{-1}(y_1,y_2),\ g_2^{-1}(y_1,y_2)\right)$$

$\downarrow$

marginal pmf for y_1

ex:- let $X_1 \sim Poi(\lambda_1)$, $X_2 \sim Poi(\lambda_2)$ be independent random variables. find pmf of $Y_1 = X_1 + X_2$

$X_1 \sim Poi(\lambda_1) \qquad p_{X_1}(x_1) = \dfrac{\lambda_1^{x_1} e^{-\lambda_1}}{x_1!}$

$X_2 \sim Poi(\lambda_2) \qquad p_{X_2}(x_2) = \dfrac{\lambda_2^{x_2} e^{-\lambda_2}}{x_2!}$

$\boxed{Y_1 = X_1 + X_2}$

$Y_2 = X_2$

1-1 transformation.

$g_2^{-1}(y_1, y_2) = X_2 = Y_2$ ✓

$g_1^{-1}(y_1, y_2) = X_1 = Y_1 - X_2 = Y_1 - Y_2$ ✓

$$p_{X_1X_2}(x_1, x_2) = p_{X_1}(x_1)\, p_{X_2}(x_2)$$

$$p_{X_1X_2}(x_1, x_2) = \frac{\lambda_1^{x_1} e^{-\lambda_1}}{x_1!} \frac{\lambda_2^{x_2} e^{-\lambda_2}}{x_2!}$$

$$p_{Y_1Y_2}(x_1, x_2) = p_{X_1X_2}\left(g_1^{-1}(y_1, y_2),\ g_2^{-1}(y_1, y_2)\right)$$

$$= p_{X_1X_2}(y_1 - y_2,\ y_2)$$

$$p_{Y_1Y_2}(x_1, x_2) = \frac{\lambda_1^{y_1-y_2} e^{-\lambda_1}}{(y_1-y_2)!} \frac{\lambda_1^{y_2} e^{-\lambda_2}}{y_2!}$$

X \ Y	1	2	3	$p_X(\cdot)$
1				$p_X(1)$
2				

Marginal pmf of Y_1.

$$p_{Y_1}(y_1) = \sum_{y_2} p_{Y_1Y_2}(y_1, y_2) =$$

$\Longrightarrow$ Transformation technique for Continuous r.v

Let X be a continuous r.v with pdf $f_X(x)$ defined on A. Let $g(x)$ be 1–1 transformation that maps from A to B, with inverse $x = g^{-1}(y)$ for every $y \in B$. Let $\frac{dx}{dy}$ be continuous and non zero for every $y \in B$. Let $Y = g(X)$.

$X \sim$

$Y = g(X)$

$$F_Y(y) = P(Y \le y) = P(g(x) \le y) = P(x \le g^{-1}(y))$$

$$= F_X(g^{-1}(y))$$

diff w.r.t y on both sides

$$\frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(g^{-1}(y)) \qquad \frac{d}{dy} = \frac{dF_X}{dx}\frac{dx}{dy}$$

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left|\frac{dx}{dy}\right| \qquad |J|$$

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

$\Rightarrow$ let $U \sim unif(0,1)$ and $V = 1-u$ find its pdf. of V.

$U \sim unif(0,1) \Rightarrow f_u(u) = 1 \quad 0<u<1$

$V = g(u) = 1-u \Rightarrow u = 1-v$

$u = g^{-1}(v) = 1-v \qquad \frac{du}{dv} = -1$

$$\Rightarrow f_v(v) = f_u(g^{-1}(v)) \left| \frac{du}{dv} \right| = f_u(1-v)\,|-1|$$

$$= 1$$ ← pdf of V.

ex:- let y be CRV with pdf

$$f_y(y) = \left\{ \frac{2+3y}{10} \quad y \in [1,3] \right\}$$

find pdf $w = y^2$

$f_w(w) = ??$

$w = g(y) = y^2 \Rightarrow y = \sqrt{w} \rightarrow$ 1-1 transformation

$y = g^{-1}(w) = \sqrt{w} \qquad \frac{dy}{dw} = \frac{1}{2\sqrt{w}}$

$$f_w(w) = f_y(g^{-1}(w)) \left| \frac{dy}{dw} \right| = f_y(\sqrt{w}) \left| \frac{1}{2\sqrt{w}} \right|$$

$$= \frac{2+3\sqrt{w}}{10} \cdot \frac{1}{2\sqrt{w}} \qquad w = [1,9]$$

bivariate continuous R.V

X_1, X_2 are CRV having $f_{X_1 X_2}(x_1, x_2)$ defined on A

$Y_1 = g_1(X_1, X_2)$

$Y_2 = g_2(X_1, X_2) \rightarrow$ you need defined

$\Big\}$ are 1-1 transformation

$X_1 = g_1^{-1}(Y_1, Y_2)$ $\quad$ $X_2 = g_2^{-1}(Y_1, Y_2)$

$$f_{Y_1 Y_2}(y_1, y_2) = f_{X_1 X_2}\left(g_1^{-1}(y_1, y_2),\ g_2^{-1}(y_1, y_2)\right)|J|$$

Jacobian

$Y_1 = g_1(x_1, x_2)$

$Y_2 = g_2(x_1, x_2)$

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}_{2\times 2}$$

$x = r\cos\theta$

$y = r\sin\theta$

$\vec{X} = (X_1, X_2 \cdots X_n)$

$\vec{Y} = g(X_1, X_2 \cdots X_n)$

$$f_{\vec{Y}}(\vec{y}) = f_{\vec{X}}(g^{-1}(y))|J|$$

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & & \\ \frac{\partial x_n}{\partial y_1} & \cdots & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

ex:- find the pdf of sum of 2 independent $U(0,1)$ r.v's.

$X_1 \sim U(0,1) \longrightarrow f_{x_1}(x_1) = 1 \quad 0 < x_1 < 1$

$X_2 \sim U(0,1) \longrightarrow f_{x_2}(x_2) = 1 \quad 0 < x_2 < 1$

$① \leftarrow Y_1 = X_1 + X_2 \rightarrow [0,2] \qquad f_{x_1x_2}(x_1 x_2) = f_{x_1}(x_1) f_{x_2}(x_2) = 1$

$② \leftarrow Y_2 = X_1 - X_2 \rightarrow [-1,1] \longrightarrow$ Assuming 1-1 transformation

$\cancel{A+} =$ ① + ②

$Y_1 + Y_2 = 2X_1 \Rightarrow X_1 = \frac{Y_1 + Y_2}{2} = g_1^{-1}(y_1, y_2)$

① − ②

$Y_1 - Y_2 = 2X_2 \Rightarrow X_2 = \frac{Y_1 - Y_2}{2} = g_2^{-1}(y_1, y_2)$

$$f_{y_1y_2}(y_1, y_2) = f_{x_1x_2}\left(g_1^{-1}(y_1, y_2), g_2^{-1}(y_1, y_2)\right)|J|$$

$x_1 = \frac{y_1 + y_2}{2}$

$x_2 = \frac{y_1 - y_2}{2}$

$$= f_{x_1x_2}\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right)|J|$$

$$= 1 \times \left|\frac{-1}{2}\right| = \frac{1}{2}$$

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{vmatrix}$$

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

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{1}{2}$$

$$\longrightarrow f_{Y_1}(y_1) = \int_{y_2} f_{Y_1,Y_2}(y_1, y_2)\, dy_2$$

ex:- 1. let $X \sim N(0,1)$ find pdf $Y = X^2$

2. $X \sim N(0,1)$ find pdf $Y = e^x$

3. let $U \sim \text{Unif}(0,1)$ for $[0,1]$ find pdf $Y = -\ln U$

$\Rightarrow$ $X \sim N(0,1)$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad \forall \quad -\infty < x < \infty$$

$Y = X^2 \longrightarrow$ Not 1-1 transformation

use CDF technique

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X < \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\,dx = \int_{x=-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx$$

$x^2 = t$

$2x\,dx = dt$

$$= 2\int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx$$

$$= \cancel{2}\int_0^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t} \frac{dt}{\cancel{2}\sqrt{t}}$$

$$= \int_0^{y} \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}t}\,dt$$

$$F_Y(y) = P(Y \le y) = \int_0^y \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}t}\, dt$$

$$F_Y(y) = \int f_Y(y)\, dy$$

$$\boxed{f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2} \quad \text{for } y > 0}$$

Ex:- If $X \sim N(0,1)$ $\quad y = e^x$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad -\infty < x < \infty$$

Transformation technique

$Y = g(x) = e^x$

$x = \ln y = g^{-1}(y)$

$\frac{dx}{dy} = \frac{1}{y}$

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

$$= f_X(\ln y) \left| \frac{1}{y} \right|$$

$$f_Y(y) = \frac{1}{y\sqrt{2\pi}} e^{-\frac{1}{2}(\ln y)^2} \longrightarrow \text{log normal distribution}$$

③ $U \sim$ Unif distribution $[0,1]$ $\quad Y = -\ln U$

find $f_Y(y) = ?$ $\quad e^{-y}$

$$f_U(u) = 1 \qquad 0 \le u \le 1,$$

$Y = -\ln u$

$u = +e^{-y} = g^{-1}(y)$

$\frac{du}{dy} = -e^{-y}$

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

$$= f_U(+e^{-y}) \left| -e^{-y} \right|$$

$$1 \cdot e^{-y} = e^{-y}$$

exponential dist with mean $=1$

Sum of random variables

⟹ Discrete case

Let X and Y be 2 independent random variables having distribution $p_x(x)$ and $p_y(y)$. find $Z = X+Y$

Transformation

$Z_1 = X+Y$

$Z_2 =$ Assume

$p_{Z_1,Z_2}(z_1,z_2)$

$p_Z(z)$

$$p_Z(z) = P(Z=z) = P(X+Y=z)$$

fix $X = k$
$Y = Z-k$

Pair wise disjoint events

$$p_Z(z) = \sum_{k=0}^{\infty} p_x(k)\, p_y(z-k) = p_x * p_y$$

$X+Y$ → Convolution

ex:- A die is rolled twice. let x_1 and x_2 be outcome on 1st & 2nd roll. let $S = X_1+X_2$

X_1 and X_2 are independent.

PMF of x_1

x	1	2	3	4	5	6
$p_x(x)$	1/6	1/6	1/6	1/6	1/6	1/6

$$p_S(s) = \sum_{k=1}^{6} p_x(k)\, p_y(s-k)$$

$p_S(2) = \boxed{p_x(1)\, p_y(1)} + p_x(2)\, p_y(0)$

$= 1/6 \times 1/6 = 1/36$

$X_1, X_2 \ldots\ldots X_n$

$Y_n = X_1 + X_2 \ldots\ldots + X_n$

$\rightarrow Y_n = Y_{n-1} + X_n$

$$\boxed{p_{Y_n}(y) = p_{Y_{n-1}} * p_{X_n}}$$

general formula.

$X_1 \; \& \; X_2$

$Y = X_1 + X_2$

$p_Y = p_{X_1} * p_{X_2}$

$Y_2 = X_1 + X_2$

$Y_3 = \boxed{X_1 + X_2} + X_3$

$Y_3 = Y_2 + X_3$

ex:- Let $X_1 \sim \text{Poi}(\lambda_1)$, $X_2 \sim \text{Poi}(\lambda_2)$ are independent. find pmf of $Y = X_1 + X_2$

$$p_{X_1}(x) = \frac{e^{-\lambda_1} \lambda_1^x}{x!} \qquad p_{X_2}(x) = \frac{e^{-\lambda_2} \lambda_2^x}{x!}$$

$x = 0, 1, 2 \ldots\ldots$

$Y = X_1 + X_2$

$$p_Y(y) = p_{X_1} * p_{X_2} = \sum_{k=0}^{y} p_{X_1}(k)\, p_{X_2}(y-k)$$

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

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

$\rightarrow \binom{y}{k} \lambda_1^k \lambda_2^{y-k}$

Binomial

$= \dfrac{y!}{(y-k)! \times k!}$

$\sum_{k=0}^{y} \binom{y}{k} \lambda_1^k \lambda_2^{y-k} = (\lambda_1+\lambda_2)^y$

$$f_Y(y) = \frac{e^{-(\lambda_1+\lambda_2)}\,(\lambda_1+\lambda_2)^y}{y!}$$

$$Y = \overset{\lambda_1}{X_1} + \overset{\lambda_2}{X_2}.$$

$$\sim \mathrm{Poi}(\underline{\underline{\lambda_1+\lambda_2}})$$

$\Rightarrow$ Sum of 2 independent continuous r.v's

let X and Y be 2 independent CRV having pdf $f_X(x)$ and $f_Y(y)$. Then $Z = X+Y$.

$$f_Z(z) = \underline{f_X * f_Y}$$

$$\left.\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx \\ &= \int_{-\infty}^{\infty} f_Y(y)\, f_X(z-y)\, dy \end{aligned}\right\}$$

$\Rightarrow$ ex:- Let X and Y be CRV having uniform distribution $[0,1]$. find $Z = X+Y$

$X \in [0,1]$

$Y \in [0,1]$

$Z \in [0,2]$

$$\boxed{f_X(x) = f_Y(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}}$$

$$f_Z(z) = f_X * f_Y = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

$$= \int_0^1 1 \cdot f_Y(z-x)\, dx$$

$$= \int_0^1 f_Y(z-x)\, dx.$$

$$f_Z(z) = \int_0^1 f_Y(z-x)\,dx \qquad z \in [0,2]$$

$f_Y(z-x) = 1$ when $0 \le z-x \le 1$

$z-x \le 1$

$z-1 \le x$

$0 \le z-x$

$z-1 \le x \le z$

I) $0 \le z \le 1$

$1 \le z \le 2$

$$f_Z(z) = \int_0^z 1\,dx = z$$

II) $1 \le z \le 2$

$$f_Z(z) = \int_{z-1}^1 1\,dx = 2-z$$

$$f_Z(z) = \begin{cases} z & 0 \le z \le 1 \\ 2-z & 1 \le z \le 2 \\ 0 & \text{otherwise} \end{cases}$$

1. Let X and Y be independent binomial dist

   $X \sim bin(n, p)$ Then find distribution of

   $Y \sim bin(m, p)$ $T = X + Y$

2. find the dist of sum of 2 independent exponential r.v.

3. find the dist of sum of 2 independent standard normal r.v's

ex:- If $X \sim Bin(n, p)$, $Y \sim Bin(m, p)$. find the distribution of $T = X+Y \sim Bin(n+m, p)$

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2 \ldots$$

$$p_Y(y) = \binom{m}{y} p^y (1-p)^{m-y} \qquad y = 0, 1, 2, \ldots$$

$$p_T = p_X * p_Y$$

$$= \sum_{x=0}^{t} p_x(x)\, p_y(t-x) = \sum_{x=0}^{t} \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{t-x} p^{t-x} (1-p)^{m-(t-x)}$$

$$= \boxed{\sum_{x=0}^{t} \binom{n}{x}\binom{m}{t-x}}\, p^t (1-p)^{m+n-t}$$

$$\boxed{\sum_{x=0}^{t} \binom{n}{x}\binom{m}{t-x} = \binom{n+m}{t}}$$

$$p_T(t) = \binom{n+m}{t} p^t (1-p)^{n+m-t}$$

$$\sim Bin(n+m, p)$$

ex:- If $X \sim exp(\lambda)$ $\quad Y \sim exp(\lambda)$ $\quad T = X+Y$

$$f_X(x) = \lambda e^{-\lambda x} \quad x > 0 \qquad f_Y(y) = \lambda e^{-\lambda y} \quad y > 0$$

$$T = X+Y \qquad f_T(t) = f_X * f_Y$$

$$f_T(t) = \int_{-\infty}^{\infty} f_X(x) f_Y(t-x)\, dx \longrightarrow \int_{-\infty}^{\infty} f_Y(y) f_X(t-y)\, dy$$

$$f_T(t) = \int_0^t \lambda e^{-\lambda x} \lambda e^{-\lambda(t-x)} dx$$

$$= \int_0^t \lambda^2 e^{-\lambda t} \underline{\underline{dx}} = \lambda^2 x\Big|_0^t e^{-\lambda t} = \underline{\underline{\lambda^2 t e^{-\lambda t}}}$$

$\Rightarrow$ ③ $X \sim N(0,1) \qquad Y \sim N(0,1) \qquad Z = X+Y.$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

$-\infty < x < \infty \qquad -\infty < y < \infty$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx = f_X * f_Y$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-x)^2} dx$$

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

$$= \frac{1}{2\pi} e^{-\frac{1}{4}z^2} \sqrt{\pi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-\left(x-\frac{z}{2}\right)^2} dx \qquad \sigma = \sqrt{2}$$

$\searrow 1 \qquad \longrightarrow \sim N(0,2)$

$\mu = 0, \ \sigma = \sqrt{2}$

$$= \boxed{\frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}z^2}} \qquad f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\underline{\underline{\sim N(0,2)}}$$

Moment generating technique

$X \sim N(\mu_1, \sigma_1^2)$ $\qquad$ $Y \sim N(\mu_2, \sigma_2^2)$

$M_x(t) = e^{\mu_1 t + \frac{1}{2}\sigma_1^2 t^2}$ $\qquad$ $M_y(t) = e^{\mu_2 t + \frac{1}{2}\sigma_2^2 t^2}$

→ ** If X and Y are independent

$M_{X+Y}(t) = M_x(t) \cdot M_y(t)$

$M_Z(t) = e^{\mu_1 t + \frac{1}{2}\sigma_1^2 t^2} \, e^{\mu_2 t + \frac{1}{2}\sigma_2^2 t^2}$

$= e^{(\mu_1+\mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2}$

$\sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

$X \sim N(\underset{\mu_1}{0}, \underset{\sigma_1^2}{1})$ $\qquad$ $Y \sim N(\underset{\mu_2}{0}, \underset{\sigma_2^2}{1})$

$Z = X+Y \sim N(0, 2)$

$$\Sigma_{p \times p} = \begin{pmatrix} \sigma_{11} & \overset{\downarrow\, Cov(x_1, x_2)}{\sigma_{12}} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & & \ddots & \vdots \\ \sigma_{p1} & \cdots & \cdots & \sigma_{pp} \end{pmatrix}$$

$$X \sim N\left(\begin{pmatrix} 5 \\ 10 \end{pmatrix}, \begin{pmatrix} 16 & 12 \\ 12 & 36 \end{pmatrix}\right)$$

$\rightarrow \mu' \qquad \sigma^2$

$$\rightarrow \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

$a' X$

$$Y = 3X_1 + 2X_2 = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\mu_y = a' \mu = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 10 \end{bmatrix} = 15 + 20 = 35$$

$$Y \sim N(35,$$

$$\Sigma = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{pmatrix} \overset{16}{\sigma_{11}} & \overset{12}{\sigma_{12}} \\ \overset{12}{\sigma_{21}} & \overset{36}{\sigma_{22}} \end{pmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \cancel{4}32$$

2: Sub-sets of Variables

If $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then all sub-sets of $\mathbf{X}$ are (multivariate) normally distributed.

For example, let's partition $\mathbf{X}$ into two sub-sets

$$\mathbf{X}_{p\times 1} = \left(\begin{array}{c} X_1 \\ \vdots \\ X_q \\ \hline X_{q+1} \\ \vdots \\ X_p \end{array}\right) = \left(\begin{array}{c} \mathbf{X}_{1(q\times 1)} \\ \hline \mathbf{X}_{2((p-q)\times 1)} \end{array}\right) \text{ and } \boldsymbol{\mu} = \left(\begin{array}{c} \mu_1 \\ \vdots \\ \mu_q \\ \hline \mu_{q+1} \\ \vdots \\ \mu_p \end{array}\right) = \left(\begin{array}{c} \boldsymbol{\mu}_{1(q\times 1)} \\ \hline \boldsymbol{\mu}_{2((p-q)\times 1)} \end{array}\right)$$

$$\boldsymbol{\Sigma}_{p\times p} = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{11(q\times q)} & \boldsymbol{\Sigma}_{12(q\times(p-q))} \\ \hline \boldsymbol{\Sigma}_{21((p-q)\times p)} & \boldsymbol{\Sigma}_{22((p-q)\times(p-q))} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right)$$

Then for

$$\mathbf{X} = \left(\frac{\mathbf{X}_{1(q\times 1)}}{\mathbf{X}_{2((p-q)\times 1)}} \right)$$

The distributions of the sub-sets are

$$\mathbf{X}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}) \qquad \text{and} \qquad \mathbf{X}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$$

The result means that

- Each of the $\mathbf{X}_i$'s are univariate normals (next page)
- All possible sub-sets are multivariate normal.

CJ Anderson, ILLinois, https://education.illinois.edu/docs/default-source/carolyn-anderson/edpsy584/lectures/MultivariateNormal-beamer-online.pdf

Little Example on Sub-sets

Suppose that

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N}_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

Due to the result on sub-sets of multivariate normals,

$$\begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_3 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{13} \\ \sigma_{31} & \sigma_{33} \end{pmatrix}\right)$$

$$X_1 \sim \mathcal{N}(\mu_1, \sigma_{11})$$
$$X_2 \sim \mathcal{N}(\mu_2, \sigma_{22})$$
$$X_3 \sim \mathcal{N}(\mu_3, \sigma_{33})$$

Also

$$\begin{pmatrix} X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_2 \\ \mu_3 \end{pmatrix}, \begin{array}{c|cc} & X_2 & X_3 \\ X_2 & \sigma_{22} & \sigma_{23} \\ X_3 & \sigma_{32} & \sigma_{33} \end{array}\right)$$

3: Zero Covariance & Statistical Independence

There are three parts to this one:

$\sigma_{12} = \Sigma_{11} = 0$

- If $\mathbf{X}_1$ is $(q_1 \times 1)$ and $\mathbf{X}_2$ is $(q_2 \times 1)$ are statistically independent, then $\text{cov}(\mathbf{X}_1, \mathbf{X}_2) = \mathbf{\Sigma}_{12} = \mathbf{0}$.

- If

$$\begin{pmatrix} \mathbf{X}_1 \\ \hline \mathbf{X}_2 \end{pmatrix} \sim \mathcal{N}_{q_1+q_2}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{pmatrix}, \left(\begin{array}{c|c} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \hline \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{array} \right) \right),$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \hline x_4 \end{pmatrix}$$

Then $\mathbf{X}_1$ and $\mathbf{X}_2$ are statistically independent if and only if $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}'_{21} = \mathbf{0}$.

- If $\mathbf{X}_1$ and $\mathbf{X}_2$ are statistically independent and distributed as $\mathcal{N}_{q_1}(\boldsymbol{\mu}_1, \mathbf{\Sigma}_{11})$ and $\mathcal{N}_{q_2}(\boldsymbol{\mu}_2, \mathbf{\Sigma}_2)$, respectively, then

$$\begin{pmatrix} \mathbf{X}_1 \\ \hline \mathbf{X}_2 \end{pmatrix} \sim \mathcal{N}_{q_1+q_2}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{pmatrix}, \left(\begin{array}{c|c} \mathbf{\Sigma}_{11} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{\Sigma}_{22} \end{array} \right) \right).$$

Example

$$\mathbf{Y}_{4\times 1} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} \text{ and } \mathbf{\Sigma}_{\mathbf{Y}} = \begin{pmatrix} 2 & 1 & 0 & .5 \\ 1 & 3 & 0 & .5 \\ 0 & 0 & 4 & 0 \\ .5 & .5 & 0 & 1 \end{pmatrix}$$

and $\mathbf{Y} \sim \mathcal{N}_4(\boldsymbol{\mu}, \mathbf{\Sigma})$.
Let's take $\mathbf{X}_1' = (Y_1, Y_2, Y_4)$ and $\mathbf{X}_2' = (Y_3)$.

Then

$$\left(\frac{\mathbf{X}_1}{\mathbf{X}_2}\right) \sim \mathcal{N}_4\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_4 \\ \hline \mu_3 \end{pmatrix}, \begin{pmatrix} 2 & 1 & .5 & 0 \\ 1 & 3 & .5 & 0 \\ .5 & .5 & 1 & 0 \\ \hline 0 & 0 & 0 & 4 \end{pmatrix}\right)$$

So set $\mathbf{X}_1$ is statistically independent of $\mathbf{X}_2$.

4: Conditional Distributions

Let $\mathbf{X}' = (\mathbf{X}'_{1(q1\times 1)}, \mathbf{X}'_{2(q_2\times 1)})$ be distributed at $\mathcal{N}_{q1+q2}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\mu} = \left(\begin{array}{c} \boldsymbol{\mu}_1 \\ \hline \boldsymbol{\mu}_2 \end{array}\right) \quad \text{and} \quad \boldsymbol{\Sigma} = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right)$$

and $|\boldsymbol{\Sigma}| > 0$ (i.e., positive definite). Then the <u>conditional distribution</u> of $\mathbf{X}_1$ given $\mathbf{X}_2 = \mathbf{x}_2$ is (multivariate) normal with mean and covariance matrix

$$\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \quad \text{and} \quad \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$$

Let's look more closely at this for a simple case of $q_1 = q_2 = 1$.

Special Case

Conditional Distribution for $q_1 = q_2 = 1$

Bivariate normal distribution

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right)$$

$$f(x_1|x_2) \quad \text{is} \quad \mathcal{N}_1\left(\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2), \sigma_{11} - \sigma_{12}\left(\frac{\sigma_{12}}{\sigma_{22}}\right)\right)$$

CJ Anderson, ILLinois, https://education.illinois.edu/docs/default-source/carolyn-anderson/edpsy584/lectures/MultivariateNormal-beamer-online.pdf

Review

$X \sim N(\mu, \sigma^2)$

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

$-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \to -\frac{1}{2}(x-\mu)(\sigma^2)^{-1}(x-\mu)$

$\longrightarrow$ univariate

$(x-\mu)^T \Sigma^{-1} (x-\mu)$

$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right) \Sigma^{-1} (x-\mu)$

Multivariate

$X = (X_1, X_2, \ldots, X_p)$

$X_1 \sim N(\mu_1, \sigma_1^2)$, $X_2 \sim N(\mu_2, \sigma_2^2)$, $X_p \sim N(\mu_p, \sigma_p^2)$

$\sim N$

Bivariate

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \underset{2\times 2}{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

$$X \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma\right)$$

$$f_X(x) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}}} e^{\frac{-1}{2(1-\rho_{12})^2}\left[\left(\frac{x-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 \boxed{- 2\rho_{12}\left(\frac{x-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x-\mu_2}{\sqrt{\sigma_{22}}}\right)}\right]}$$

$x \to \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

If x_1 and x_2 are independent $\Longrightarrow \rho_{12} = 0$ ✓

If $\rho_{12} = 0$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_{11}}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sqrt{\sigma_{11}}}\right)^2} \cdot \frac{1}{\sqrt{2\pi\sigma_{22}}} e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sqrt{\sigma_{22}}}\right)^2}$$

$$f_X(x) = f_X(x_1) \times f_X(x_2)$$

$$\Rightarrow \quad Y = a_1 x_1 + a_2 x_2 \qquad X \sim N_2\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}\right)$$

$$a_1 x_1 + a_2 x_2 \rightarrow [a_1 \;\; a_2]\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

$$Y \sim N(a\mu, \; a^T \Sigma a)$$

$$\rightarrow \quad Y_1 = a_{11}x_1 + a_{12}x_2 \qquad A = \begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$$

$$Y_2 = a_{21}x_1 + a_{22}x_2$$

$$Y \sim N\left(A\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}, \; A\Sigma A^T\right)$$

$\rightarrow$ $X \sim N_3(\mu, \Sigma)$ find the distribution of

$$Y = \begin{bmatrix}x_1 - x_2\\ x_2 - x_3\end{bmatrix} \qquad Y_1 = x_1 - x_2, \quad Y_2 = x_2 - x_3$$

$$X \sim N_3\left(\begin{pmatrix}\mu_1\\ \mu_2\\ \mu_3\end{pmatrix}, \begin{array}{c|ccc} & x_1 & x_2 & x_3\\ \hline x_1 & \sigma_{11} & \sigma_{12} & \sigma_{13}\\ x_2 & \sigma_{21} & \sigma_{22} & \sigma_{23}\\ x_3 & \sigma_{31} & \sigma_{32} & \sigma_{3}\end{array}\right)$$

$$Y = \begin{bmatrix}Y_1\\ Y_2\end{bmatrix} = \underbrace{\begin{bmatrix}x_1 - x_2\\ x_2 - x_3\end{bmatrix}}_{} = \underbrace{\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & -1\end{bmatrix}}_{A}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$$

$$Y \sim N(A\mu, \; A\Sigma A^T)$$

$$\mu_y = A\mu = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \end{bmatrix}$$

Covariance matrix $= \Sigma_y = A \Sigma A^T$

$$= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} \sigma_{11} - \sigma_{12} & \sigma_{12} - \sigma_{13} \\ \sigma_{21} - \sigma_{22} & \sigma_{22} - \sigma_{23} \\ \sigma_{31} - \sigma_{32} & \sigma_{32} - \sigma_{33} \end{bmatrix} = \begin{bmatrix} \sigma_{11} - \sigma_{12} - \sigma_{21} + \sigma_{22} & \sigma_{12} - \sigma_{13} - \sigma_{22} + \sigma_{23} \\ \sigma_{21} - \sigma_{22} - \sigma_{31} + \sigma_{32} & \sigma_{22} - \sigma_{23} - \sigma_{32} + \sigma_{33} \end{bmatrix}$$

$$\Sigma_y = \begin{bmatrix} \sigma_{11} - 2\sigma_{12} + \sigma_{22} & \sigma_{12} - \sigma_{13} - \sigma_{22} + \sigma_{23} \\ \sigma_{21} - \sigma_{22} - \sigma_{31} + \sigma_{32} & \sigma_{22} - 2\sigma_{23} + \sigma_{3} \end{bmatrix}$$

$$Y \sim N_2\left(\begin{pmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \end{pmatrix}, \begin{pmatrix} \sigma_{11} - 2\sigma_{12} + \sigma_{22} & \sigma_{12} - \sigma_{13} - \sigma_{22} + \sigma_{23} \\ \sigma_{21} - \sigma_{22} - \sigma_{31} + \sigma_{32} & \sigma_{22} - 2\sigma_{23} + \sigma_{3} \end{pmatrix}\right)$$

$$X_{p\times 1} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_q \\ \hline X_{q+1} \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} X_{1(q\times 1)} \\ \hline X_{2(p-q)\times 1} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \hline \mu_2 \end{pmatrix}, \left(\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right) \right)$$

$$Y_{4\times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} \qquad \Sigma_Y = \begin{array}{c|cccc} & \begin{matrix}Y_1\\ \sigma_{11}\end{matrix} & \begin{matrix}Y_2\\ \sigma_{12}\end{matrix} & \begin{matrix}Y_3\\ \sigma_{13}\end{matrix} & \begin{matrix}Y_4\\ \sigma_{14}\end{matrix} \\ \hline Y_1 & 2 & 1 & 0 & 0.5 \\ Y_2 & 1 & 3 & 0 & 0.5 \\ Y_3 & 0 & 0 & 4 & 0 \\ Y_4 & 0.5 & 0.5 & 0 & 1 \end{array}$$

$$\mu_Y = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$$Y \sim N_4(\mu_Y, \Sigma_Y)$$

find $\begin{pmatrix} Y_1 \\ Y_2 \\ Y_4 \\ \hline Y_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 1 \\ 2 \\ 4 \\ \hline 3 \end{pmatrix}, \quad \right)$

$$Y_{New} = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_4 \\ \hline Y_3 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_4 \\ \hline \mu_3 \end{pmatrix}_{x_1}, \begin{pmatrix} \overset{\Sigma_{11}}{\begin{matrix} 2 & 1 & 0.5 \\ 1 & 3 & 0.5 \\ 0.5 & 0.5 & 1 \end{matrix}} & \overset{\Sigma_{12}}{\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}} \\ \hline \underset{\Sigma_{21}}{\begin{matrix} 0 & 0 & 0 \end{matrix}} & \underset{\Sigma_{22}}{4} \end{pmatrix}\right)$$

$$\Sigma_{Y_{New}} = \begin{array}{c|cccc} & Y_1 & Y_2 & Y_4 & Y_3 \\ \hline Y_1 & \sigma_{11} & \sigma_{12} & \sigma_{14} & \sigma_{13} \\ Y_2 & \sigma_{21} & \sigma_{22} & \sigma_{24} & \sigma_{23} \\ Y_4 & \sigma_{41} & \sigma_{42} & \sigma_{44} & \sigma_{43} \\ Y_3 & \sigma_{31} & \sigma_{32} & \sigma_{34} & \sigma_{33} \end{array}$$

$\Sigma_{12} = 0$

(Y_1, Y_2, Y_4) and Y_3 are statistically independent.

$\Longrightarrow \quad X \sim N_3(\mu, \Sigma) \qquad \Sigma = \left(\begin{array}{cc|c} 4 & 1 & 0 \\ 1 & 3 & 0 \\ \hline 0 & 0 & 2 \end{array}\right)$

$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$

$\Rightarrow$ Verify whether X_1, X_2 are independent?

(X_1, X_2) and X_3 are independent?

$$\begin{pmatrix} X_1 \\ \hline X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{array}{c} \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{cc} x_1 & x_2 \end{array} \\ \left(\begin{array}{c|c} 4 & \overbrace{1}^{\Sigma_{12}} \\ \hline \underbrace{1}_{\Sigma_{21}} & 3 \end{array}\right) \end{array} \right) \qquad \underline{\underline{\Sigma_{12} \neq 0}}$$

X_1 and X_2 are dependent

(b)

$$\begin{pmatrix} X_1 \\ X_2 \\ \hline X_3 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \hline \mu_3 \end{pmatrix}, \left(\begin{array}{cc|c} 4 & 1 & 0 \\ 1 & 3 & 0 \\ \hline 0 & 0 & 2 \end{array}\right) \right)$$

Σ_{12}, Σ_{21}

$\Sigma_{12} = 0$

So (X_1, X_2) and X_3 are statistically Independent.

Unit – 4 [5 Hours]: Limit Theorems and Inequalities-The Central Limit Theorem, Law of Large Numbers, Boole's Inequality, Bonferroni's Inequality, Chebyshev's and Markov's Inequality, Cauchy-Schwartz inequality, Jensen's Inequality;

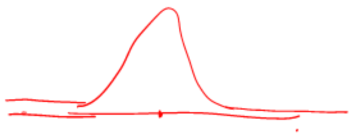
Unit - 4

Limit theorems and Inequalities

1) Markov's Inequality

Suppose X is a non negative random variable then for any $a > 0$ we have

$$P(X \geq a) \leq \frac{E[X]}{a} \longrightarrow$$



Proof:- We consider continuous case

Let X be a positive continuous random variable

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx \xrightarrow{\text{as } X \geq 0} \int_{0}^{\infty} x f_X(x)\, dx \quad (X \geq 0)$$

$$= \int_{0}^{a} x f_X(x)\, dx + \int_{a}^{\infty} x f_X(x)\, dx$$

$$\geq \int_{a}^{\infty} x f_X(x)\, dx$$

x takes values from a to ∞

$x f_X(x) = a f_X(x)$

$x \geq a$

$x f_X(x) \geq a f_X(x)$

$$\geq \int_{a}^{\infty} a f_X(x)\, dx$$

$$= a \int_{a}^{\infty} f_X(x)\, dx = a P(X \geq a)$$

$$E[X] \geq P(X \geq a) \times a$$

$$\Longrightarrow \boxed{P(X \geq a) \leq \frac{E[X]}{a}}$$

Another Proof

$P(X \geq a) \leq \frac{E[X]}{a}$

Let $I = \begin{cases} 1 & X \geq a \\ 0 & \text{otherwise} \end{cases}$

$$X \geq a \Longrightarrow \frac{X}{a} \geq 1$$

$$\Longrightarrow \frac{X}{a} \geq I$$

$$E\left[\frac{X}{a}\right] \geq E[I] = 1 \cdot P(X=a) + 1 \cdot P(X=a+1)$$

$$+ 0 \times P(\text{other})$$

$$= 1 \cdot P(X \geq a)$$

$$E\left[\frac{X}{a}\right] \geq P(X \geq a)$$

$$\frac{E[X]}{a} \geq P(X \geq a) \Longrightarrow \boxed{P(X \geq a) \leq \frac{E[X]}{a}}$$

2) <u>Chebyshev's Inequality</u>

If X is a random variable with finite mean 'μ' and variance as σ^2. then for any value $k > 0$

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

$|x-\mu| \geq k$

<u>Proof:-</u> $\Rightarrow |x-\mu|^2 \geq k^2$ → Positive

$E[x]$

$E[(x-\mu)^2] = E[x^2] - \mu^2$

by Markov's inequality

$$P(|x-\mu|^2 \geq k^2) \leq \frac{E[|x-\mu|^2]}{k^2} = \frac{Var(x)}{k^2} = \frac{\sigma^2}{k^2}$$

$$P(|x-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

ex:

Suppose that if it is known that number of items produced in a factory during a week is a r.v. with mean = 50

a) What can be said about the probability that this week's production will exceed 75.

(b) If the variance of a week's production is 25. What is the probability that this week's production will be between 40 and 60.

(a) $X \sim$ No. of items produced during a week

using Markov's

$$P(X > 75) \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3}$$

(b) $P(40 < X < 60)$

$P(X > 60)$

$P(X > 40)$

$$= P(40 - 50 < X - 50 < 60 - 50)$$

$$= P(-10 < X - 50 < 10) = P(|X - 50| < 10)$$

$$= 1 - P(|X - 50| > 10)$$

$$P(40 < X < 60) \geq 1 - \frac{1}{4} = \frac{3}{4}$$

$$P(|X - 50| > 10) \leq \frac{Var(X)}{10^2} = \frac{25}{100} = \frac{1}{4}$$

Ex:

$$X \sim \exp(\lambda = 1) \longrightarrow f_X(x) = \lambda e^{-\lambda x}$$

$$= e^{-x} \quad \text{for } x > 0$$

using Markov

$$P(X \geq a) \leq \frac{E[X]}{a}$$

$$P(X \geq a) \leq \frac{1}{a}$$

exact

$$P(X \geq a) = \int_a^\infty f_X(x)\,dx$$

$$= \int_a^\infty e^{-x}\,dx = e^{-a}$$

ex:- $X \sim \text{unif}$ over $[0, 10]$

find $P(|X - 5| > 4)$

$X \sim$ uniform $[\overset{a}{0}, \overset{b}{10}]$

$$f_X(x) = \frac{1}{b-a} = \frac{1}{10}, \quad E[X] = \frac{a+b}{2} = \frac{10}{2} = 5$$

$$\text{Var}(X) = \frac{(b-a)^2}{12} = \frac{10^2}{12} = 25/3$$

using Chebyshev's Inequality

$$P(|X-5| > 4) \leq \frac{\text{Var}(X)}{4^2} = \frac{25}{3 \times 16} = \frac{25}{48} = 0.52$$

using def

$[0, 10]$

$$P(|X-5| > 4) = P(X > 9 \text{ and } X < 1)$$

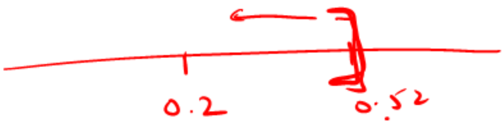
$X - 5 > 4$, $\quad -(X-5) > 4$

$X > 9$ $\quad X - 5 < -4$

$\quad X < 1$

$$= \int_0^1 \frac{1}{10}\,dx + \int_9^{10} \frac{1}{10}\,dx$$

$$= \frac{1}{10} + \frac{1}{10} = \frac{2}{10} = 0.2$$



3) <u>Cauchy Schwartz Inequality</u>

Suppose x, y are 2 R.V then

$$\left[E[xy]\right]^2 \le E[x^2]\, E[y^2]$$

<u>Proof:</u> let $(ax-y)^2 \ge 0$

$$E\left[(ax-y)^2\right] \ge 0$$

$$E\left[a^2x^2 + y^2 - 2axy\right] \ge 0$$

$$a^2 E[x^2] + E[y^2] - 2aE[xy] \ge 0$$

$$a^2 E[x^2] + E[y^2] - \frac{E[xy]^2}{E[x^2]} - 2aE[xy] + \frac{E[xy]^2}{E[x^2]}$$

$$\underbrace{E[x^2]}_{\ge 0}\left(\underbrace{\left(a - \frac{E[xy]}{E[x^2]}\right)^2}_{\ge 0} + \frac{E(y^2)}{E(x^2)} - \left(\frac{E[xy]}{E[x^2]}\right)^2\right] \ge 0$$

$$\frac{E[y^2]}{E[x^2]} - \left(\frac{E[xy]}{E[x^2]}\right)^2 \ge 0$$

$$E[y^2]\, E[x^2] \ge \left(E[xy]\right)^2$$

Inequalities

① Markov's Inequality :- If $X \ge 0$

$$P(X \ge a) \le \frac{E[X]}{a}$$

② Chebyshev's Inequality

$$P(|X-\mu| \ge k) \le \frac{\sigma^2}{k^2}$$

③ Cauchy Schwartz's $\qquad X = ay$

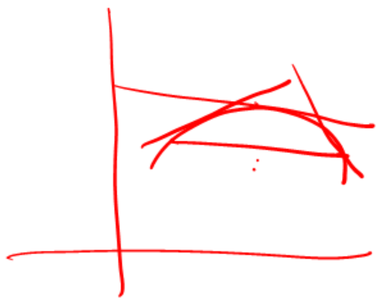
$$(E[XY])^2 \le E[X^2]\,E[Y^2]$$

④ Jensen's Inequality

$\Longrightarrow$ If g is convex function

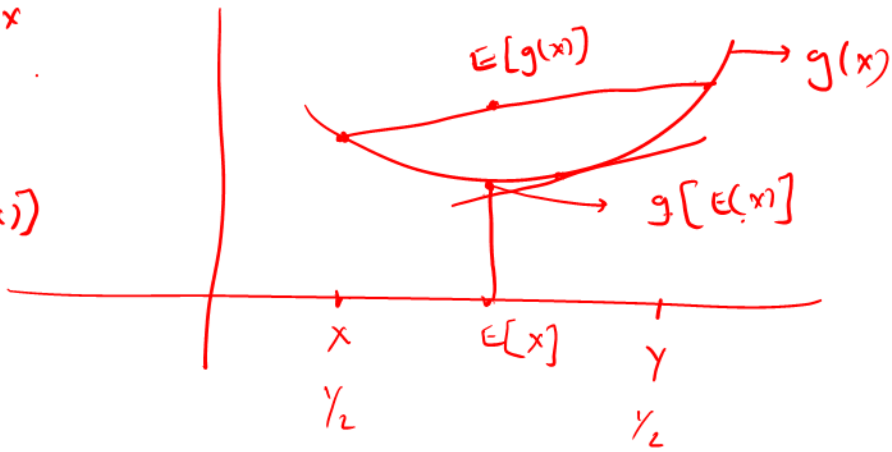
① $g''(x) \ge 0$

②



If g is convex function then

$$E[g(x)] \ge g(E(x))$$



Proof:- g is a convex function

(a) $g''(x) \geq 0$

$\downarrow$

(b) $g(x) \geq g(c) + g'(c)(x-c)$

$\downarrow$ $E[x]$

$$g(x) \geq g(E[x]) + g'(E[x])(x-E[x])$$

$$E[g(x)] \geq E[\underline{g(E(x))} + g'(E(x))(x-E[x])]$$

$E[\pi]$

$$\geq g(E(x)) + g'(E(x))\,\underline{E[x-E[x]]}$$

$$\geq g(E[x]) + g'(E(x))\underbrace{[E(x)-E[x]]}_{\to 0}$$

$$\boxed{E[g(x)] \geq g(E[x])}$$

ex:- $g(x) = x^2$

$g''(x) = 2$ convex function

using Jensen's Inequality

$$E[g(x)] \geq g(E(x))$$

$$\boxed{E[x^2] \geq (E[x])^2}$$

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

$$E[x^2] = \underbrace{\underline{Var(x)}}_{\geq 0} + (E[x])^2 \geq (E[x])^2$$

$$\boxed{E[x^2] \geq (E[x])^2}$$

Ex:- $g(x) = x^4$

$g'(x) = 4x^3$

$g''(x) = 12x^2 \geq 0 \ \forall x$ → Convex function

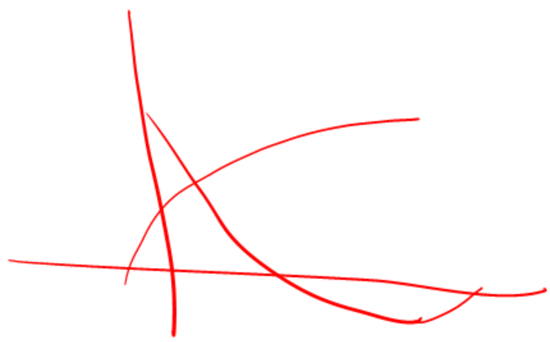
$$E[\underline{g(x)}] \geq g(E[x])$$

$$E[x^4] \geq (E[x])^4$$

$\Longrightarrow$ $g(x) = -\log x$

$g''(x) = \frac{1}{x^2} \geq 0 \ \forall x$

→ convex



$$E[g(x)] \geq g(E(x))$$

$$E[-\log x] \geq -\log(E(x))$$

$$-E[\log(x)] \geq -\log(E(x))$$

$$E[\log(x)] \leq \log(E[x])$$

$f(x) = \log(x)$

$$E[f(x)] \leq f(E[x])$$

$\Rightarrow$ If f is a concave function $(f''(x) \le 0)$

then by Jensen's Inequality

$$E[f(x)] \le f(E(x))$$

$\Rightarrow$ A coin is weighted so that Probability of getting head is 20%. Suppose the coin is flipped 20 times. find a bound for the Probability that it lands on head atleast 16 times

X ~ Number of heads)

$$X \sim Bin(20, 0.2)$$

$$p_X(x) = {}^{20}C_x (0.2)^x (0.8)^{20-x}$$

$$E[X] = np = 20 \times 0.2 = 4$$

$$Var(X) = npq = 20 \times 0.2 \times 0.8 = \underline{3.2}$$

exact :

$$P(X \ge 16) = \sum_{x=16}^{20} {}^{20}C_x (0.2)^x (0.8)^{20-x}$$

$\Rightarrow$ using Inequality (Markov's)

$$P(X \ge 16) \le \frac{E[X]}{16} = \frac{4}{16} = \underline{1/4}$$

Law of large numbers

1. The weak law of large numbers
2. The strong law of large numbers

⟹ The weak law of large numbers

1. Let $X_1, X_2 \cdots X_n$ iid r.v (independent and identically distributed r.v) with finite mean μ
$E[X_i] = \mu$, $Var(X_i) = \sigma^2$. for $\varepsilon > 0$

$$P\left(\left|\frac{X_1 + X_2 \cdots X_n}{n} - \mu\right| \geq \varepsilon\right) \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

Tossing of ^{100} Coin

$X_i \to$ getting head

$E[X_i] =$

Proof:-

$$\bar{X}_n = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \cdots x_n}{n}$$

$$E[\bar{X}_n] = E\left[\frac{X_1 + X_2 \cdots X_n}{n}\right] = \frac{n\mu}{n} = \mu$$

$$Var(\bar{X}_n) = Var\left(\frac{X_1 + X_2 + \cdots X_n}{n}\right) = \frac{1}{n^2} Var(X_1 + X_2 + \cdots X_n)$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Chebyshev inequality

$$P\left(\left|\bar{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

$$\frac{x_1 + \cdots x_{n^*}}{n^*}$$

$n > n^*$

The Strong law of large numbers

Let $X_1, X_2 \cdots X_n$ be a sequence of iid r.v

each having mean $\mu = E[X_i]$, $Var(X_i) = \sigma^2$

Then with probability 1

$$\frac{X_1 + X_2 \cdots X_n}{n} \longrightarrow \mu \quad \text{as} \quad n \to \infty$$

$$P\left\{ \left| \frac{X_1 + X_2 \cdots X_n}{n} - \mu \right| = 0 \right\} = 1$$

$$\Rightarrow \quad \bar{X}_n = \frac{X_1 + X_2 \cdots X_n}{n} \sim \qquad X_i :$$

$\Sigma X_i + \cdots X_n$

$X_1 + X_2$

$X_1 + X_2 + \cdots X_n \sim 1$

The central limit theorem

Let $X_1, X_2 \dots X_n$ be a sequence iid r.v.

each having mean μ, variance σ^2 then distribution

of $\overline{X_n}$

$$\boxed{\frac{X_1 + X_2 + \cdots X_n - n\mu}{\sigma\sqrt{n}}} \longrightarrow \text{standard normal distribution as } n \to \infty$$

$$P\left\{ \frac{X_1 + X_2 \cdots X_n}{\sigma\sqrt{n}} \le a \right\} \longrightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{1}{2}x^2} dx$$

X_i with $E[X_i] = \mu$, $Var(X_i) = \sigma^2$

$$\overline{X_n} = \frac{X_1 + X_2 \cdots X_n}{n} \overset{\cdot}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

Standardizing

$$Z = \frac{X - \mu}{\sigma}$$

$$\frac{\frac{X_1 + X_2 \cdots X_n}{n} - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{X_1 + X_2 \cdots X_n - n\mu}{\sqrt{n}\,\sigma} \overset{\cdot}{\sim} N(0,1)$$

$\Longrightarrow$ $X_i \sim Ber(p)$

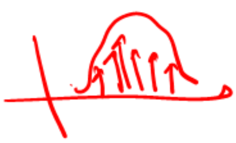
unif

$\sum X_i \sim$

$X_1 + X_2$

$X_1 + X_2$

$X_1 + X_2 \cdots X_n$



$X_1 + \cdots X_n$

$\bar{X}_n \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$ ✓ $\longrightarrow$

$\bar{X}_n = \frac{X_1 + X_2 + \dots X_n}{n}$

$E[\bar{X}_n] = \mu$

$Var(\bar{X}_n) = \frac{\sigma^2}{n}$

$\sum X_i \dot{\sim} N(n\mu, n\sigma^2)$ ✓

$Var(X_1) = \sigma^2$

$Var(X_2) = \sigma^2$

$E[X_1 + X_2 + \dots X_n] = n\mu$

$Var(X_1 + X_2 \dots X_n) = n\sigma^2$

$E[X_i] =$

$E[$

$Var(X_1 + X_2) = E[(X_1 + X_2 - \mu)]$

Ex:- number of claims in a insurance company per day with $\mu = 40$, $\sigma = 12$ Survey is conducted for 50 working days

X_i – no of claims per day

$P(\bar{X}_n < 35)$

$$\bar{X} = \frac{x_1 + x_2 \cdots x_n}{n}$$

$$P\left(\frac{x_1 + x_2 \cdots x_n}{50} < 35\right) \quad \dot\sim N\left(40, \frac{12^2}{50}\right)$$

$$\downarrow$$

$$P\left(\frac{\frac{x_1 + x_2 \cdots x_n}{50} - \mu}{\sqrt{\sigma^2/n}} < \frac{35 - 40}{\sqrt{12^2/50}}\right)$$

$$P\left(z < \frac{35 - 40}{12/\sqrt{50}}\right) = P(z < -2.95)$$

$$= \underline{\underline{0.00159}}$$

Normal distribution to Binomial distribution

$X_i \rightarrow$ be iid rv having bernoulli distribution

$\sum X_i$ $\quad X_i \sim Ber(p)$ $\quad X_i =$ number of success in one bernoulli trial

$X_1, X_2, \ldots\ldots X_n$ iid bernoulli r.v

$\sum X_i \sim Bin(n, p)$ $\quad n = 100$

$\sum X_{100} = Bin(100, 1/2)$

$P(X \leq 50)$

$E[X_i]$ $\quad Var[X_i]$

$\sum X_i \;\dot\sim\; N(n\mu, n\sigma^2) \;\dot\sim\; N(np, npq)$

$\bar{X}_n \;\dot\sim\; N(\mu, \sigma^2/n) \;\dot\sim\; N(p, pq/n)$

$X_i \sim Ber(p)$	$X = \sum X_i \sim Bin(n, p)$
$E[X_i] = p = \mu$	$E[\sum X_i] = np$
$Var(X_i) = pq = \sigma^2$	$Var[X] = npq$

How large n is

$\Longrightarrow$ $\left. n \geq 30 \right\}$ (or) $\left\{ \begin{array}{l} np \\ npq \end{array} \right. > 5$

Normal distribution to Poisson distribution

let $X_1, X_2 \dots X_n$ be iid Poisson r.v with mean (λ)

$X_i \sim Poi(\lambda_i)$

$X_1 + X_2 \sim Poi(\lambda_1 + \lambda_2)$

$E[X_i] = \lambda$

$Var[X_i] = \lambda$

$$\sum X_i \sim Poi(n\lambda) \;\dot{\sim}\; N(n\lambda, n\lambda)$$

$$\bar{X}_n \;\dot{\sim}\; N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\dot{\sim}\; N\left(\lambda, \frac{\lambda}{n}\right)$$

$\Rightarrow$ $X_1, X_2 \dots X_n$ be iid exponential r.v

$$\sum X_i \sim gamma(n, \lambda) \;\dot{\sim}\; N(\;\; , \;\;)$$

?

$$\bar{X}_n \;\dot{\sim}\; N(\qquad)$$

?

Central limit theorem

$x_1, x_2 \ldots x_n$ iid r.v.

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots X_n}{n} \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right) \longrightarrow \sim N(0,1)$$

$$\sum x_i = x_1 + x_2 \ldots x_n \dot{\sim} N(n\mu, n\sigma^2) \longrightarrow \sim N(0,1)$$

$\Rightarrow$ Normal approximation to binomial

$n \geq 30$

(or) $np, nq > 5$

$x_1, x_2 \ldots x_n$ iid Bernoulli r.v

$E[x_i] = \mu$

$Var[x_i] = \sigma^2$

$$\sum x_i \sim Bin(np, npq) \dot{\sim} N(np, npq)$$

$$n\mu, \quad n\sigma^2$$

$$\overline{X_n} \dot{\sim} N\left(p, \frac{pq}{n}\right)$$

$$\mu, \quad \sigma^2/n$$

$\Rightarrow$ Normal approximation to poisson

$x_1, x_2 \ldots x_n$ are iid Poisson r.v (λ)

$E[x_i] = \lambda$

$Var[x_i] = \lambda$

$$\sum x_i \sim Poi(n\lambda) \dot{\sim} N(n\lambda, n\lambda)$$

$$n\mu, \quad n\sigma^2$$

λ as large

$$\overline{X_n} \dot{\sim} N\left(\lambda, \frac{\lambda}{n}\right)$$

★★ $Poi(\lambda) \dot{\sim} N(\lambda, \lambda)$

Normal approximation to exponential

$X_1, X_2 \ldots X_n$ are iid exponential r.v. (λ)

$E[X_i] = \frac{1}{\lambda}$

$Var[X_i] = \frac{1}{\lambda^2}$

$\sum X_i \sim gamma(n, \lambda) \overset{\cdot}{\sim} N\left(\frac{n}{\lambda}, \frac{n}{\lambda^2}\right)$

** The continuity correction

$P(X = 50) = 0$

$P(X = 50) \neq 0$ ord

$X = 50 \Longrightarrow 49.5 < X < 50.5$

$X \geq 50 \Longrightarrow X > 49.5$

$X > 50 \Longrightarrow X > 50.5$

$X < 15 \Longrightarrow X < 14.5$

$X \leq 15 \Longrightarrow X \leq 15.5$

$P(X < 40) = P(X < 39.5)$

$P(X < 40.5$

$P(X=5) = 0$

The continuity correction

$P(X=4) = 0 \longrightarrow P(3.5 < X < 4.5)$

$X = 4 \qquad 3.5 < X < 4.5$

$X \leq 15 \qquad X < 15.5$

$X \geq 15 \qquad X > 14.5$

Continuity correction table

- If P(X=n) use P(n – 0.5 < X < n + 0.5)
- If P(X > n) use P(X > n + 0.5)
- If P(X ≤ n) use P(X < n + 0.5)
- If P (X < n) use P(X < n – 0.5)
- If P(X ≥ n) use P(X > n – 0.5)

Discrete	Continuous
x = 6	5.5 < x < 6.5
x > 6	x > 6.5
x ≤ 6	x < 6.5
x < 6	x < 5.5
x ≥ 6	x > 5.5

A bank teller serves customers standing in the queue one by one. Suppose that the service time Xi for customer i has mean EXi=2 (minutes) and Var(Xi)=1. We assume that service times for different bank customers are independent. Let Y be the total time the bank teller spends serving 50 customers. Find P(90<Y<110).

X_i = Time spent with ith customer.

$X_1, X_2 \ldots X_{50}$ iid r.v.

$E[X_i] = 2 = \mu$, $Var[X_i] = 1$

$n = 50$ using CLT

$Y = X_1 + X_2 + \cdots X_{50} = \sum X_i \dot\sim N(n\mu, n\sigma^2)$

$\bar{X}_n = \frac{X_1 + X_2 \cdots X_{50}}{n} \dot\sim N(\mu, \frac{\sigma^2}{n})$ $\dot\sim N(100, 50)$

$P(90 < X < 110) \xrightarrow{\text{Standardizing}} N(0,1)$ $P\left(\frac{90-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{110-\mu}{\sigma}\right)$

$= P\left(\frac{90-100}{\sqrt{50}} < Z < \frac{110-100}{\sqrt{50}}\right)$

$= P(-\sqrt{2} < Z < \sqrt{2})$

$= P(-1.41 < Z < 1.41)$

$= P(Z < 1.41) - P(Z < -1.41)$

$= 0.92073 - 0.07927$

$= 0.84146$ //

The mean number of kidney transplants performed per day in the United States in a recent year was about 45. Find the probability that on a given day,

a. exactly 50 kidney transplants will be performed,
b. at least 65 kidney transplants will be performed, and
c. no more than 40 kidney transplants will be performed.

$X \sim$ Number of kidney transplants per day.

$X \sim Poi(45)$

$$P_X(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(45)^x e^{-45}}{x!}$$

$$P(X = 50) = \frac{(45)^{50} \cdot e^{-45}}{50!} = \_\_\_\_\_\_\_\_$$

<u>Using CLT</u>

$$X \sim Poi(45) \overset{\cdot}{\sim} N(45, 45)$$

$$P(X = 50) \xrightarrow[\text{Correction}]{\text{Continuity}} P(49.5 < X < 50.5)$$

$\downarrow$ Standardize

$$= P\left(\frac{49.5 - 45}{\sqrt{45}} < Z < \frac{50.5 - 45}{\sqrt{45}}\right)$$

$$= P(0.67 < Z < 0.82)$$

$$= P(Z < 0.82) - P(Z < 0.67)$$

$$= 0.79389 - 0.74875$$

$$= \underline{\underline{0.0453}}$$

(b) $P(X \geq 65) \xrightarrow[\text{Correction}]{\text{Continuity}} P(X \geq 64.5) = 1 - P(X < 64.5)$

$$= 1 - P\left(Z < \frac{64.5 - 45}{\sqrt{45}}\right)$$

$$\simeq 0.00181$$

(c) $P(X \leq 40) = P(X < 40.5)$

?

ex:- The avg calls received per hour by an insurance company is 5. Calculate the probability in working day of 8 hrs. The number of calls received will be

(a) exactly 36 per day

(b) b/w 42 and 45 inclusive

$X \sim$ number of calls received per day

$\lambda = 8 \times 5 = 40$ calls per day

$X \sim Poi(40) \; \dot{\approx} \; N(40, 40)$

(a) $P(X = 36) \xrightarrow{\text{continuity}} P(35.5 < X < 36.5)$

$\downarrow$ Standardize

(b) $P(42 \leq X \leq 45) = P(41.5 < X < 45.5)$

Unit – 5 [6 Hours]: Statistical Inference-I-Distributions of the sample mean and the sample variance for a normal population, Chi-Square, t and F distributions. Estimation, Unbiasedness, consistency, the method of moments and the method of maximum likelihood estimation;

Unit-5

Statistical Inference - I

⇒ **Population** :- Collection of all possible observations of an experiment or survey is termed as population

Population

- Finite population
  - ✓ number of observations are finite
- Infinite population

Ex:- Population in world
→ tigers population in the world

Sample :- subset of observations of populations

⇒ **data point** :- A particular data from a given population.

⇒ **Sampling** :- The process of obtaining the sample from the population is termed as Sampling

→ Since the sampling is the random process associated with population. It is also termed as random sampling

① Classification of Sampling

(1) Simple random sampling :- (SRS)

This a method of selecting n observations out of N observations in the population such that every observation $^{N}C_{n}$ samples has equal chance of being chosen

(2) Probability Sampling :- Each member of the population gets definite probability of being selected

(3) Simple random sampling with replacement (SRSWR)

If every member is drawn from the population one by one. after drawing the data is noted and also replaced. So one data point may occur several times

(4) SRSWOR :- one data point can occur only one time.

⇒ Parameter and Statistic

⇒ Calculating mean, variance, s.d.

→ These are calculated on the basis of population is called as Parameter. Population mean, population variance, sd

→ These are calculated on the basis of sample is called as Statistic

⇒ Sample mean, sample variance, sample sd

Parameter is constant

Statistic is different for different samples from the same population.

⇒ Statistical distribution of $\bar{x}$. → sample mean, variance

⇒ Statistical Inference

estimation

Testing of hypothesis

Ex:- Population : {2, 3, 4, 5}

Take random sample of size 2 (with replacement)

→ { (2,2), (2,3) (2,4), (2,5)
(3,2) (3,3) (3,4) (3,5)
(4,2) (4,3) (4,4) (4,5)
(5,2) (5,3) (5,4) (5,5) }

$x_1 = 2$
$x_2 = 3$
$x_3 = 4$
$x_4 = 5$

⟹ Population mean :- $\mu = \frac{2+3+4+5}{4}$

$= 3.5$

⟹ Sample mean

Sample	Sample mean
(2,2)	2
(2,3) (3,2)	2.5
(2,4) (4,2)	3
(2,5) (5,2)	3.5
(3,3)	3
(3,4) (4,3)	3.5
(3,5) (5,3)	4
(4,4)	4
(4,5) (5,4)	4.5
(5,5)	5

$\mu_{\bar{x}} =$

Sampling distribution

$\bar{x}$	$p(\bar{x})$
2	1/16
2.5	2/16
3	3/16
3.5	4/16
4	3/16
4.5	2/16
5	1/16

$$\mu_{\bar{x}} = \sum \bar{x}\, p_x(\bar{x})$$

$$= 2 \times \frac{1}{16} + 2.5 \times \frac{2}{16} + \cdots$$

$$E[\overline{X}_n] = \underline{\underline{3.5}}$$

Variance :- $\sigma_{\bar{x}}^2$: $= \dfrac{\sigma^2_{population}}{n}$

Sample Variance :- $\dfrac{\sum(\bar{x}_i - \bar{x})^2}{n-1}$ ✓ $\bar{x} = \dfrac{\sum X_i}{n}$

Population Variance $\dfrac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$ ✓

n is large — CLT

Population $\longrightarrow$ mean $= \mu$

Variance $= \sigma^2$

$$\bar{x} = \frac{x_1 + x_2 + \cdots x_n}{n}$$

Sample mean $\mu_{\bar{x}} = E[\bar{x}] = E\left[\frac{x_1+x_2 \cdots x_n}{n}\right] = \mu$

Sample Variance $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$

ex:- A high school class of 100 students taken an IQ test and average is 105.

$P(\bar{x} \geq 105) =$

Normal population

$\Rightarrow \mu = 100$, $\sigma = 25$

$\mu_{\bar{x}}$ $\qquad \sigma_{\bar{x}}^2 = \frac{25^2}{100}$

$= 0.022$

Example 1: Consider a population consisting of only four numbers 1, 2, 3 and 4 with 30% 1's, 40% 2's, 20% 3's and the remaining 10% 4's. That is the probability distribution in the population may be expressed in terms of the p.m.f. $p_Y(y)$ of the population random variable Y as follows:

y	1	2	3	4
$p_Y(y)$	0.3	0.4	0.2	0.1

Sample : n = 2

Now consider drawing a random sample of size 2 from this population and the three statistics $\overline{Y} = (Y_1 + Y_2)/2$, $s_2^2 = 1/2\sum_{i=1}^{2}(Y_i - \overline{Y})^2$ and $s_1^2 = \sum_{i=1}^{2}(Y_i - \overline{Y})^2$, where Y_1 and Y_2 are the two observations. The sampling distributions of these three statistics can be figured out by considering all possible samples of size 2 that can be drawn from this population, the corresponding probabilities of drawing each such sample, and the values of each of the statistics for every such sample

①	$Y_1\ Y_2$ {1, 1}	{1, 2}	{1, 3} {1, 4} {2, 2} {2, 3} {2, 4}, {3, 3} {4, 4} {3, 4}
$P_Y(y)$	0.09	0.3×0.4	
$\overline{Y} = \frac{y_1+y_2}{2}$	1	1.5	

Possible Samples	$Y_1 Y_2$ {1,1}	{1,2}	{1,3}	{1,4}	{2,2}	{2,3}	{2,4}	{3,3}	{3,4}	{4,4}
Probability	0.09	0.24	0.12	0.06	0.16	0.16	0.08	0.04	0.04	0.01
$\frac{y_1+y_2}{2} = \overline{Y}$	1	1.5	2	2.5	2	2.5	3	3	3.5	4
$\frac{1}{2}\sum(y_i-\bar{y})^2$ s_2^2	0	0.25	1	2.25	0	0.25	1	0	0.25	0
$\sum(y_i-\bar{y})^2$ s_1^2	0	0.5	2	4.5	0	0.5	2	0	0.5	0

Y_1 Y_2 Y_3 Y_4
1 2 3 4

Sampling Distribution of the Sample Mean $\overline{Y}$

$\bar{y}$	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$p_{\overline{Y}}(\bar{y})$	0.09	0.24	0.28	0.22	0.12	0.04	0.01

Sampling Distribution of s_1^2

s_2^2	0	0.5	2	4.5
prob.	0.30	0.44	0.20	0.06

Sampling Distribution of s_2^2

s_2^2	0	0.25	1	2.25
prob.	0.30	0.44	0.20	0.06

Ex2- The average weekly earning for the employees in general automotive shops was \$406 nd sd for weekly earnings was \$55. A random sample of size 100 is selected

$X_i \sim N(\mu, \sigma^2)$ $\mu = 406$

a) find the mean of $\overline{X}n$

b) find the sd of $\overline{X}n$ $\sigma = 55$

c) Find the sampling distribution of $\overline{X}n$ $\overline{X}_n = \frac{X_1 + X_2 \cdots X_n}{n}$

d) Compute $P(\overline{xn}<400)$

(a) $E[\overline{X}_n] = E\left[\frac{X_1 + \cdots X_n}{n}\right] = \mu = 406$

$Var(\overline{X}_n) = \frac{\sigma^2}{n} = \frac{55^2}{100}$

using CLT,

(c) $\overline{X}_n \dot\sim N\left(\mu, \frac{\sigma^2}{n}\right) \dot\sim N\left(406, \frac{55^2}{100}\right)$

$P(\overline{X}_n < 400) \xrightarrow{\text{standardize}} = P\left(\frac{X_n - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{400-406}{5.5}\right)$

$= P(Z < -1.09)$

Estimation

Suppose population having pdf $f(y/\theta)$

$\sim N(\mu, \sigma^2)$ pmf $P(y/\theta)$

$\sim Bin(n, p)$

$\sim Exp(\lambda)$

Random Sample – n iid r.v $Y_1, Y_2, \cdots Y_n$

from the population r.v $Y \rightarrow \theta$

Estimation

Point estimation

we need single valued estimation for θ.

Interval estimation

we need interval of values which supposed to contain true unknown value θ

Point estimate of parameters

- Let θ be an unknown parameter of the distribution of a random variable Y.
- Want to estimate θ on the basis of a random sample $Y_1, Y_2, \ldots, Y_n$.
- Using a particular statistic T where T is itself a random variable.
- T is the estimator of θ. Value of T obtained from a given sample is its estimate.
- We want **T** to be a good estimator. The difference $|T-\theta|$ should be as small as possible.

$|T - \theta|$

$\Rightarrow$ X $[x_1, x_2 \ldots x_n]$

ex:- Suppose the r.v $X \sim$ Normally distributed with unknown mean. Suppose a sample is selected $x_1 = 25$, $x_2 = 30$, $x_3 = 29$, $x_4 = 31$

$X \sim N(\mu, \sigma^2)$

Point estimate for population mean $\rightarrow \mu$

Sample mean $T = \dfrac{x_1+x_2+x_3+x_4}{4} = \dfrac{25+30+29+31}{4}$

$= 28.75$

$\theta = 28.75 = \hat{\mu}$

$\Rightarrow$ Unknown Population Variance

Sample Variance $= \dfrac{\sum (x_i - \bar{x})^2}{n-1} = 6.9$

$= \hat{\sigma}^2$

$\dfrac{\sum (x_i - \mu)^2}{n}$

How good is your estimate

- In point estimation how to decide the nature of a “good” estimator?

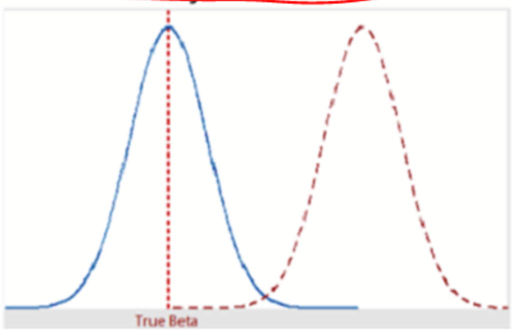
- The kind of properties, characteristics or behavior a reasonable estimator should possess. What are the desirable criteria for “good” estimators.

Unbiasedness

$X = [x_1, x_2 \dots x_n]$

$\bar{X}_n = \frac{x_1 + x_2 \dots x_n}{n}$

- Goodness can be defined in many ways.
- T is an estimator of θ and T has a sampling distribution
- What if the sampling distribution of T has a central tendency towards θ ?
- A statistic T is called unbiased if $E(T) = \theta$.
- $E(T) - \theta = b(\theta)$ where $b_\theta(T)$ is the bias of T.



$X \sim N(\mu, \sigma^2)$

$x_1, x_2 \dots x_n$

$E_i[X] = \mu$

$E[x_i] = \mu$

$x_1 \sim N(\mu, \sigma^2)$

$E\left[\frac{x_1 + x_2}{2}\right] = \mu$

$\rightarrow \sim N\left(\mu, \frac{\sigma^2}{2}\right)$

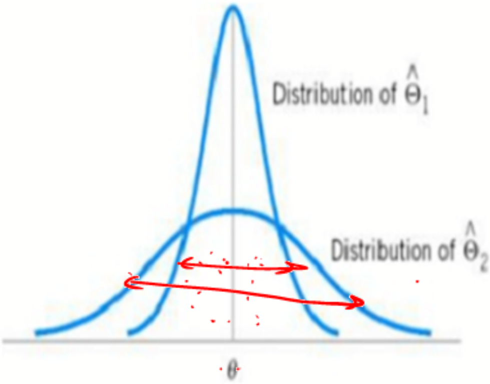
- Minimum variance property: The sampling distribution of estimator T should ~~also have a small dis~~persion

$E[T] = \theta$

- Among all unbiased estimators, T should have the smallest variance.

$E[T'] = \theta$

- $Var(T) \leq Var(T')$ where T' is any other unbiased estimator
- A statistic T following these conditions is called MVUE



Sampling distribution

Population $\Rightarrow$ $X \sim N(\mu, \sigma^2)$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \rightarrow \text{r.v}$$

$T = (x_1, x_2 \cdots x_n)$

$\rightarrow$ statistic

Mean Squared Error and Standard Error

- MSE: MSE of an estimator $\hat{\theta}$, MSE($\hat{\theta}$) = $E(\hat{\theta} - \theta)^2$
- MSE($\hat{\theta}$)= Var($\hat{\theta}$)+ $[B(\hat{\theta})]^2$
- If $\hat{\theta}$ is unbiased for θ, then MSE($\hat{\theta}$)= Var($\hat{\theta}$)
- $\sqrt{Var(\hat{\theta})}$ is known as standard error of the estimator $\hat{\theta}$

$Var(\hat{\theta})$

standard deviation

$\sqrt{Var(\hat{\theta})}$

Estimation :-

Population – Unknown Parameter = θ

Point Estimation

Single value

$\hat{\theta}$

Interval estimation

$[\quad , \quad]$

$\hat{\theta}$

Properties of estimation

Unknown = θ

$T(x_1, \dots x_n) : \hat{\theta}$

$E[\hat{\theta}] = \theta$

① unbiasedness $\Rightarrow$

② Minimum Variance Property

③ Standard error, Mean Square error

④ consistency

$\Rightarrow$ unbiasedness

$$E[\hat{\theta}] = \theta \rightarrow \text{unbiased estimate}$$

$$\rightarrow b_{\hat{\theta}}(\theta) = 0$$

$$E[\hat{\theta}] - \theta = b_{\hat{\theta}}(\theta) \rightarrow \text{bias of estimate}$$

$\Rightarrow$ T_1, T_2 are unbiased estimates

$$Var(T_1) < Var(T_2) \rightarrow \text{Minimum Variance Property}$$

iid $x_1, x_2 \dots x_n$ are s.r.s from population $X \sim N(\mu, \sigma^2)$

$$E[X_i] = \mu, \quad Var[x_i] = \sigma^2 \quad . \quad x_i \overset{iid}{\sim} N(\mu, \sigma^2)$$

① $Y_1 = X_2$ ② $Y_2 = \dfrac{x_1 + x_3 + x_5 + x_7}{4}$

③ $Y_3 = \dfrac{x_1 + x_2 \cdots x_n}{n}$

$$E[\hat{\theta}] = \theta$$

① $E[Y_1] = E[x_2] = \mu$, $Var[Y_1] = \sigma^2$

② $E[Y_2] = E\left[\dfrac{x_1 + x_3 + x_5 + x_7}{4}\right] = \dfrac{4\mu}{4} = \mu$

$Var[Y_2] = \dfrac{\sigma^2}{4}$

③ $E[Y_3] = E\left[\dfrac{x_1 + \cdots x_n}{n}\right] = \mu$

unbiased estimator

$\rightarrow Var[Y_3] = \dfrac{\sigma^2}{n}$

$\Rightarrow$ Y_1, Y_2, Y_3 are unbiased estimator

$\Rightarrow$ Y_3 is having minimum variance

$\Rightarrow$ Sample mean is the unbiased estimator having minimum variance.

$\Rightarrow$ Mean Square error

$$E[\hat{\theta}] - \theta = b_{\hat{\theta}}(\theta)$$

$$\mu - \theta = b_{\hat{\theta}}(\theta)$$

$$E[(\hat{\theta}-\theta)^2] = E[(\hat{\theta}-\mu+\mu-\theta)^2] \qquad E[\hat{\theta}]=\mu$$

$E[(x-\mu)^2]$

$$= E[(\hat{\theta}-\mu)^2+(\mu-\theta)^2+2(\hat{\theta}-\mu)(\mu-\theta)]$$

$$E[(\hat{\theta}-\theta)^2] = E[(\hat{\theta}-\mu)^2] + E[(\mu-\theta)^2]+2E[(\hat{\theta}-\mu)(\mu-\theta)]$$

$$MSE(\theta) = Var[\hat{\theta}] + (\mu-\theta)^2 + 2(\mu-\theta)\underbrace{E[\hat{\theta}-\mu]}_{0}$$

$$MSE(\theta) = Var[\hat{\theta}] + (b\,\hat{\theta}(\theta))^2$$

unbiased estimate

$$MSE[\theta] = Var[\hat{\theta}]$$

$\hat{\theta}_1, \hat{\theta}_2 \ldots \hat{\theta}_n$

having minimum mean square error (MSE

is good estimate

Consistency :- $X_1, X_2 \cdots X_n$ from Population X.

$$T(X_1, \ldots X_n) = T_n \qquad \psi(\theta)$$

$$T_n \longrightarrow \psi(\theta) \text{ as } n \to \infty$$

WLLN iid $X_1, \ldots X_n$ s.r.s from population X,

$$P\left(\left|\frac{X_1 + X_2 \cdots X_n}{n} - \mu\right| \geq \varepsilon\right) \longrightarrow 0 \text{ as } n \to \infty$$

$$P(|T_n - \psi(\theta)| \geq \varepsilon) \longrightarrow 0 \text{ as } n \to \infty$$

$$\Rightarrow P(|T_n - \psi(\theta)| \leq \varepsilon) \longrightarrow 1 \text{ as } n \to \infty$$

Def :- Let $X_1, X_2 \cdots X_n$ be a sequence of iid r.v with common cdf F_θ. A sequence of point estimators T_n will be called as consistent for $\psi(\theta)$ if $T_n \to \psi(\theta)$ as $n \to \infty$

Iff

$$P(|T_n - \psi(\theta)| \geq \varepsilon) \longrightarrow 0 \text{ as } n \to \infty \quad \forall \varepsilon > 0$$

(or)

$$P(|T_n - \psi(\theta)| < \varepsilon) \longrightarrow 1 \text{ as } n \to \infty \quad \forall \varepsilon > 0$$

Ex: Let $X_1, X_2 \ldots X_n$ be iid r.v such that $X_i \sim N(\mu, \sigma^2)$ from $X \sim N(\theta, \sigma^2)$

Verify if $\overline{X}_n$ is consistent estimator?

$$P\left(\left|\overline{X}_n - \theta\right| \geq \varepsilon\right) = \quad \text{as } n \to \infty$$

$$\overline{X}_n = \frac{X_1 + X_2 \cdots X_n}{n}$$

$$E[\overline{X}_n] = \mu$$

$$Var[\overline{X}_n] = \frac{\sigma^2}{n}$$

$$\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\downarrow$$

$$Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$$

$$= P\left(\frac{\left|\overline{X}_n - \theta\right|}{\frac{\sigma}{\sqrt{n}}} \geq \frac{\varepsilon}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= P\left(|Z| \geq \frac{\sqrt{n}\varepsilon}{\sigma}\right)$$

$$= 1 - P\left(|Z| \leq \frac{\sqrt{n}\varepsilon}{\sigma}\right)$$

$$= 1 - P\left(-\frac{\sqrt{n}\varepsilon}{\sigma} < Z < \frac{\sqrt{n}\varepsilon}{\sigma}\right)$$

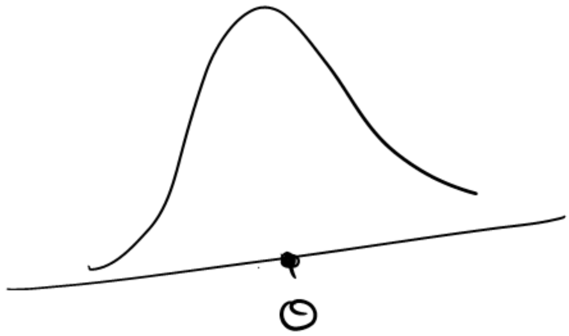
$$= 1 - \int_{\frac{\sqrt{n}\varepsilon}{\sigma}}^{\frac{\sqrt{n}\varepsilon}{\sigma}} f(z)\, dz \quad \text{as } n \to \infty$$

$$= \lim_{n \to \infty}\left[1 - \int_{-\frac{\sqrt{n}\varepsilon}{\sigma}}^{\frac{\sqrt{n}\varepsilon}{\sigma}} f(z)\, dz\right] = 1 - \int_{-\infty}^{\infty} f(z)\, dz = 1 - 1$$

$$= 0$$

$\overline{X_n}$ is consistent estimator as

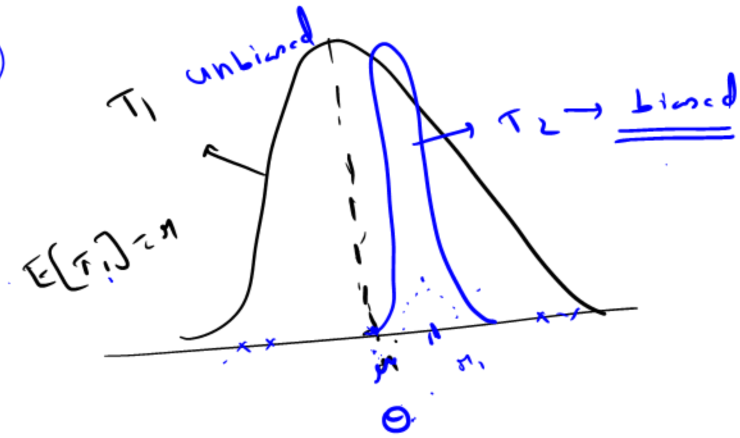
$$P(|X_n - \theta| \geq \varepsilon) \longrightarrow 0 \quad \text{as} \quad n \to \infty$$



$$\frac{\bar{X}_n - \theta \ \mu}{\sigma}$$

$(0, 1)$

ex:



$\Rightarrow$ Methods for finding estimation

$T(x_1, \dots x_n)$

✓ ① Maximum likelyhood estimation $\rightarrow$ ① Very popular.

② Method of moment

② Works well for large samples

ex: Tossing 2 coins independently

$X_i \sim$ number of heads $\quad X \sim Bin(2, P)$ ✓

Pmf

$\theta = 0.2$ (or) $\underline{0.8}$

	$x=0$	$x=1$	$x=2$
✓ $\theta = 0.2$	0.64	0.32	0.04
✓ $\theta = 0.8$	0.04	0.32	0.64

$$P_X(x) = \binom{2}{x} \theta^x (1-\theta)^{2-x}$$

$$= \binom{2}{0} (0.2)^0 (0.8)^2$$

- The concept of the maximum likelihood method was exercised in the last example. The unknown parameter θ is estimated with the value that maximizes the probability of obtaining an observed sample realization. The most probable value of θ is required to be chosen as the estimate according to the available data.

X is population $\rightarrow$ unknown parameter θ

$\downarrow$

$x_1, \ldots, x_n$ $\quad$ X_i having pdf / pmf $f(x_i, \theta)$

Joint pdf

$$f(x_1, x_2 \ldots x_n/\theta) = f(x_1, \theta)\, f(x_2, \theta) \ldots f(x_n, \theta)$$
$$= L(\theta) \rightarrow \text{Likelihood function}$$

- Let $f(x_1, x_2, \ldots, x_n/\theta)$ be the joint pdf or pmf of the sample observations. For fixed θ, this is a function of sample observations. But when $x_1, x_2, \ldots, x_n$ are observed and θ is unknown, it can be considered as a function of θ, $L(\theta)$ which is known as likelihood function. So in order to be a maximum likelihood estimator of θ, $L(\hat{\theta}) = max(L(\theta))$ should be satisfied and $\hat{\theta}$ is the MLE.

$L > 0$ $\quad$ $\ln L$

$$\frac{\partial L}{\partial \theta} = 0 \quad , \quad \frac{\partial^2 L}{\partial^2 \theta} < 0$$

$$\frac{\partial \ln L}{\partial \theta} = 0 \quad , \quad \frac{\partial^2 L}{\partial^2 \theta} < 0$$

- In general it is convenient to deal with $\ln L(\theta)$ rather than $L(\theta)$. $\ln L(\theta)$ gets its highest value for the same value of θ as $L(\theta)$ does.

- Consider $x_1, x_2, \ldots, x_n$ to be a realization of random sample of independent observations from a Poisson distribution with parameter λ. The likelihood function can be derived as:

$$L(\lambda) = \frac{exp[-n\lambda]\lambda^{\sum x_i}}{\prod (x_i!)}$$

$$ln\ L(\lambda) = -n\lambda + \sum x_i\ ln\ \lambda - \sum ln(x_i!)$$

$$\frac{d\ ln\ L(\lambda)}{d\lambda} = -n + (\sum x_i)\frac{1}{\lambda}$$

The maximum likelihood estimate of λ is $\hat{\lambda} = \sum x_i / n$.

$x_i \sim Poi(\lambda)$

$$p_x(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$L(\lambda)$ = Joint pdf of $(x_1, x_2 \ldots x_n)$

$$L(\lambda) = f(x_1, \lambda)\ f(x_2, \lambda) \cdots\cdots f(x_n, \lambda)$$

$$= \frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \frac{e^{-\lambda}\lambda^{x_2}}{x_2!} \cdot \cdots\cdots \frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$$

Likelihood function

$$L(\lambda) = \frac{e^{-n\lambda}\lambda^{x_1+x_2\cdots x_n}}{x_1!\ x_2! \cdots x_n!}$$

$$\frac{\partial L(\lambda)}{\partial\lambda} = 0 \quad \Big| \quad \frac{\partial Ln(\lambda)}{\partial\lambda} = 0$$

Apply $\underline{\underline{Ln}}$

$$\underline{\underline{Ln\, L(\lambda)}} = Ln\left\{\frac{e^{-n\lambda}\lambda^{\sum x_i}}{\prod x_i!}\right]$$

$$= -n\lambda + \sum x_i\, Ln\,\lambda - Ln(x_1!\ x_2! \cdots x_n!)$$

$$Ln\, L(\lambda) = -n\lambda + \sum x_i\, Ln\,\lambda - \sum Ln\, x_i!$$

diff λ

$$\frac{\partial\, Ln\, L(\lambda)}{\partial\lambda} = -n + \sum x_i \cdot \frac{1}{\lambda} + 0 = 0$$

Method of Moments for finding estimation

$X \sim N(\mu, \sigma^2)$ $\qquad \theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_k \end{bmatrix}$

- An unknown population X has parameter vector $\theta = [\theta_1, \ldots, \theta_k]'$. The population moments are functions of the unknown parameters. That is, population moment of order r is given by: $\alpha_r = \alpha_r(\theta_1, \ldots, \theta_k) = E(X^r)$. → Theoritical value

$\alpha_1 = E[X] = \int x f(x)\,dx = \sum x f(x)$

- Given an i.i.d. sample $X_1, X_2, \ldots, X_n$, let us denote the sample non-central moment of order r by a_r:

$$a_r = \frac{1}{n}\sum X_i^r, \text{ r=1,2,3...}$$

$a_1 = \frac{\sum x_i}{n} = \bar{x}$

$a_2 = \frac{\sum x_i^2}{n}$

$\alpha_2 = E[x^2] = Var[x] + \mu^2$

The sample moment(s) are then equated with the theoretical population moment(s) and are then solved for $\theta_1, \ldots, \theta_k$. If there are k unknowns, one typically considers the first k non-central moments, whichever is convenient.

Example

- Population $X \sim N(\mu, \sigma^2)$, i.i.d sample $X_1, X_2, \ldots, X_n$ from the population. Compute the MME of μ and σ^2.
- As there are two parameters, at least two equations are required.

$\hat{\mu}$, $\hat{\sigma}^2$

Population moments

$\alpha_r = E[X^r]$

$\alpha_1 = E[X] = \mu$

$\alpha_2 = E[X^2] = Var[X] + (E[X])^2$

$\alpha_2 = \sigma^2 + \mu^2$

Sample moments

$a_r = \frac{1}{n} \sum x_i^r$

$a_1 = \frac{1}{n} \sum x_i = \bar{x}$

$a_2 = \frac{1}{n} \sum x_i^2$

Solving eq α_1, a_1

$\alpha_1 = a_1 \Rightarrow \mu = \bar{x} \Longrightarrow \hat{\mu} = \bar{x}$

Solving α_2 and a_2

$$\sigma^2 + \mu^2 = \frac{1}{n} \sum x_i^2$$

$$\sigma^2 = \frac{1}{n} \sum x_i^2 - \mu^2$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum x_i^2 - (\bar{x})^2$$

Sample Variance is the unbiased estimator for population variance

$S^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ → unbiased

Let $Y_1, Y_2, \ldots, Y_n$ are n i.i.d random sample from an unknown population with mean μ and variance σ^2. Then $S^2 = \frac{1}{n-1}\sum(Y_i - \bar{Y})^2$ is an unbiased estimator of σ^2.

$s^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$

biased

$$S^2 = \frac{1}{n-1}\sum(Y_i - \bar{Y})^2 = \frac{1}{n-1}\sum(Y_i^2 - 2Y_i\bar{Y} + \bar{Y}^2)$$

$$= \frac{1}{n-1}(\sum Y_i^2 - 2\bar{Y}\sum Y_i + n\bar{Y}^2) = \frac{1}{n-1}(\sum Y_i^2 - 2n\bar{Y}^2 + n\bar{Y}^2)$$

$$= \frac{1}{n-1}(\sum Y_i^2 - n\bar{Y}^2) = \frac{1}{n-1}\sum Y_i^2 - \frac{n}{n-1}(\frac{1}{n}\sum Y_i)^2$$

$$= \frac{1}{n-1}\sum Y_i^2 - \frac{1}{n(n-1)}(\sum Y_i)^2$$

$$E(S^2) = \frac{1}{n-1}\sum E(Y_i^2) - \frac{1}{n(n-1)}E[(\sum Y_i)^2]$$

$$= \frac{1}{n-1}\sum(Var(Y_i) + [E(Y_i)]^2) - \frac{1}{n(n-1)}E[(\sum Y_i)^2]$$

Let, $Z = \sum Y_i$, $E(Z^2) = Var(Z) + [E(Z)]^2$

$$= \frac{1}{n-1}\sum(Var(Y_i) + [E(Y_i)]^2) - \frac{1}{n(n-1)}\{Var[(\sum Y_i)] + (E[\sum Y_i])^2\}$$

$$= \frac{1}{n-1}\sum(\sigma^2 + \mu^2) - \frac{1}{n(n-1)}\{\sum Var(Y_i) + (\sum E[Y_i])^2\}$$

[by independence of samples]

$$= \frac{1}{n-1}(n\sigma^2 + n\mu^2) - \frac{1}{n(n-1)}\{n\sigma^2 + n^2\mu^2\}$$

$$= \frac{1}{n(n-1)}(n^2\sigma^2 + n^2\mu^2 - n\sigma^2 - n^2\mu)^2\} = \sigma^2$$

MLE vs MME

- MLE is statistically well established and has often lower variance than other methods
- MLE may be computationally very expensive and slow
- MMEs are easy to derive
- MMEs are often used to provide starting values while computing MLE
- MME may not be unique for a given sample

Sampling Distributions

$T\ (x_1, x_2 \ldots x_n)$

- Statistic is a random variable. The value of a statistic may vary from sample to sample
- Imagine collecting a number of samples (each of size n) from a population and calculating the statistic for each sample. This series of realizations of statistic may be plotted in a histogram to see its frequency distribution.
- For large number of samples (each of size n), this distribution is called the sampling distribution of the statistic.
- Sampling distribution: mean, s.d, moments of higher orders
- Standard deviation of statistic is known as standard error of the statistic

$(x_1, x_2 \ldots x_n)$

$X_i \sim N(\mu, \sigma^2)$

$\bar{X}$ = sample mean

Standard error

$\bar{x} = \sqrt{Var(\bar{x})} = \frac{\sigma}{\sqrt{n}}$

Expected value

$\bar{x} = \mu$

$X_i \sim N(\mu, \sigma^2)$

$x_1, x_2 \ldots x_n =$

$T(x_1, \ldots x_n) = \bar{x}_n$

standard error $(\bar{x}_n) = \sqrt{Var(\bar{x}_n)} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$

$E\left[\frac{x_1 + x_2 \ldots x_n}{n}\right] = \mu$

Expected value $E[\bar{x}_n] = \mu$

$Var(\bar{x}) = \frac{\sigma^2}{n}$

Four fundamental distributions derived from normal

- Standard normal distribution (Z)
- Chi-square (χ^2) distributiom
- t distribution
- F distribution

Def:- Standard Normal Variable is normal distribution with mean = 0, Variance = 1

---

ex: $X \sim N(\mu, \sigma^2)$

$\downarrow$ Standardize

$$\frac{X-\mu}{\sigma} \sim N(0, 1)$$

$\Rightarrow$ If x is normally distributed with mean μ and Variance σ^2. Then $y = a + bx$ when $b \neq 0$ is also normally distributed with

$$Y \sim N(a + b\mu, b^2\sigma^2)$$

$X \sim N(\mu, \sigma^2)$

$E[X] = \mu$, $Var[X] = \sigma^2$

$E[y] = E[a + bx] = a + b\,E[X] = a + b\mu$

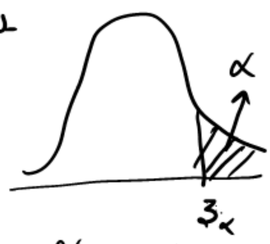
$Var[y] = Var[a + bx] = b^2 Var[x] = \underline{\underline{b^2\sigma^2}}$

$\Rightarrow$ <u>Upper α point</u> and <u>lower α point</u>

$$\frac{X-\mu}{\sigma} \sim N(0, 1)$$

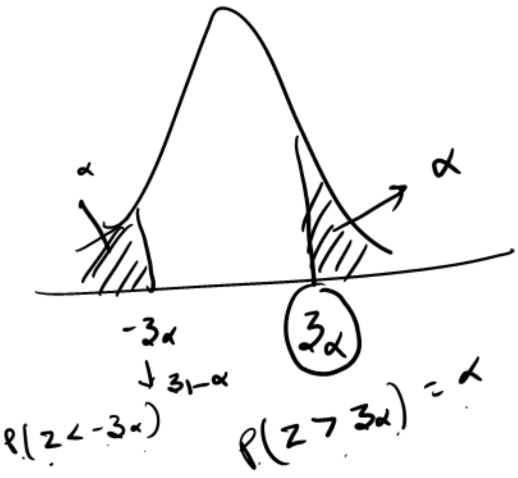
$\Rightarrow$ $z_\alpha \rightarrow$ The value of z such that

$$P(Z > z_\alpha) = \alpha$$



z_α is called upper α point

$P(z > z_\alpha)$



z_α = upper α - point

upper 100α % point of Z

$z_{1-\alpha} = -z_\alpha \Rightarrow$ by symmetry

$z_{1-\alpha}$ is lower α point.

$$P(Z < z_{1-\alpha}) = \alpha$$

$$P(Z > -z_\alpha) = 1-\alpha = P(Z > z_{1-\alpha})$$

$$P(Z > z_\alpha) = 1 - P(Z < z_\alpha)$$

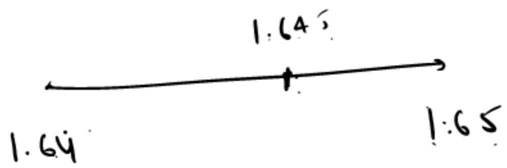
$$P(Z < 1.96) = 0.975$$

$$P(Z > 1.96) = \overset{1-0.975}{0.025}$$

$Z_{0.025}$

$z_{0.05}$ $\quad P(Z > z_\alpha) = 0.05$

1.645



1.645 $\quad$ 1.64

0.94950 $\quad$ 0.95053 $\quad$ 0.95

Interpolation

$$\text{Starting probability} + \left(\frac{1.645 - 1.64}{1.65 - 1.64}\right)(\text{diff b/w probabilities})$$

Chi-Square (χ^2) distribution

Let $x_1, x_2 \cdots x_\nu$ be ν mutually independent

ν nu

$\boxed{\nu \to \text{nu}}$

Standard normal variables

$\sum_{i=1}^{\nu} x_i^2$ is known here χ^2 distribution with

ν degree freedom (df $= \nu$)

pdf: $$f(\chi^2) = \frac{1}{2^{\nu/2}\,\Gamma(\nu/2)}\, e^{-\chi^2/2} (\chi^2)^{\nu/2 - 1}$$

$0 < \chi^2 < \infty$

Mean $= \nu$

Variance $= 2\nu$

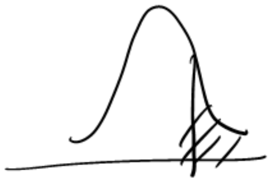
Suppose 1 standard normal variable x_1

$x_1^2 \sim \chi^2(1)$ mean $= 1$

Variance $= 2$

ex: Take iid random samples $x_1, x_2 \ldots x_k$ from some $N(\mu, \sigma^2)$ population

$$X_i \sim N(\mu, \sigma^2)$$



$$\frac{x_i - \mu}{\sigma} \sim N(0,1)$$

$$\sum_{i=1}^{k} \left(\frac{x_i - \mu}{\sigma}\right)^2 \sim \chi^2(k)$$

$\Rightarrow$ Suppose μ is unknown $\quad \bar{x} = \frac{x_1 + \cdots x_k}{k}$

$$\sum_{i=1}^{k} \left(\frac{x_i - \bar{x}}{\sigma}\right)^2 = \left(\frac{x_1 - \bar{x}}{\sigma}\right)^2 + \left(\frac{x_2 - \bar{x}}{\sigma}\right)^2 + \cdots \left(\frac{x_k - \bar{x}}{\sigma}\right)^2$$

$$S^2 = \frac{1}{k-1} \sum (x_i - \bar{x})^2 \qquad = \frac{(k-1)S^2}{\sigma^2} \sim \chi^2(k-1)$$

$\underline{\underline{k=2}}$ $\quad$ 2 independent variables x_1, x_2

$$\bar{x} = \frac{x_1 + x_2}{2}$$

$$x_1 - \bar{x} = x_1 - \frac{x_1 + x_2}{2} = \frac{x_1 - x_2}{2}$$

$$x_2 - \bar{x} = x_2 - \frac{x_1 + x_2}{2} = \frac{x_2 - x_1}{2} \quad \frac{x_1 - x_2}{2} = -(x$$

$$-\left(\frac{x_1 - x_2}{2}\right)$$

$$\left(\frac{x_1 - \bar{x}}{\sigma}\right)^2 + \left(\frac{x_2 - \bar{x}}{\sigma}\right)^2$$

$$= \left(\frac{x_1 - x_2}{2\sigma}\right)^2 + \left(\frac{x_2 - x_1}{2\sigma}\right)^2 = \left(\frac{x_1 - x_2}{\sqrt{2}\,\sigma}\right)^2$$

$$\left(\frac{X_1 - X_2}{\sqrt{2}\sigma}\right)^2 \sim \chi^2(1) \longrightarrow$$

$E[X_1 - X_2] = 0$

$Var[X_1 - X_2] = Var[X_1] + Var[-X_2]$

$= \sigma^2 + \sigma^2$

$= \underline{\underline{2\sigma^2}}$

Result

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

Note

$\Rightarrow$ Suppose Y_1 and Y_2 are 2 Independent rv distributed with χ^2 with df ν_1 and ν_2 respectively

$Y_1 = X_1^2 + X_2^2 \dots X_{\nu_1}^2$

$$Y_1 + Y_2 \sim \chi^2(\underbrace{\nu_1 + \nu_2}_{\text{df}})$$

$Y_2 = \cancel{Z_1^2} + Z_2^2 \dots Z_{\nu_2}^2$

$\Rightarrow$ **Note**

$\chi^2_{\alpha,\nu}$: Value of $\chi^2(\nu)$ for which

$$P(\chi^2 > \chi^2_{\alpha,\nu}) = \alpha$$

$\chi^2_{\alpha,\nu}$: upper α point of χ^2 distribution with df ν

$\chi^2_{1-\alpha,\nu}$: lower α point of χ^2 distribution

Graph

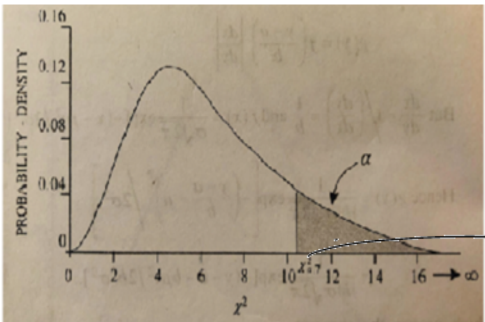
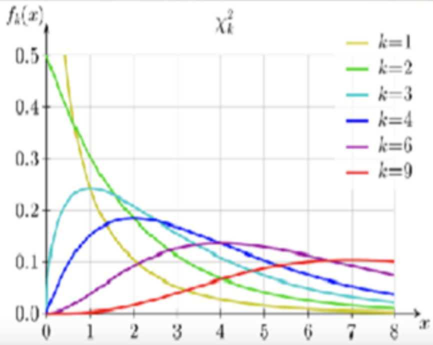
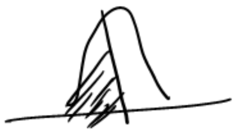


Fig. 13.2 χ^2 distribution with 7 degrees of freedom ($\nu = 7$).

$\chi^2_{\alpha,7}$

$P(X > \chi^2_{\alpha,7}) = \alpha$





$P(Z < 3.)$

ex:- Calculate the probability that for a random sample of 5 values taken from population $N(100, 25^2)$

(i) find probability that $\bar{X}$ is between 80 and 120

$P(80 < \bar{X} < 120)$

(ii) Find probability that S will exceed 41.7

$P(S > 41.7)$

Population $X \sim N(100, 25^2)$

$$\bar{X} = \frac{x_1 + x_2 + \dots x_5}{5} \overset{\text{using CLT}}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$= N(100, 125)$$

$$P(80 < \bar{X} < 120) \xrightarrow{\text{Standardize}} P\left(\frac{80-100}{\sqrt{125}} < \frac{\bar{X} - 100}{\sqrt{125}} < \frac{120-100}{\sqrt{125}}\right)$$

$$= P(-1.789 < z < 1.789)$$

$$=$$

(b) $P(S > 41.7)$ $\qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(4)$

$$P\left(\frac{4S^2}{(25)^2} > \frac{(41.7)^2 \times 4}{(25)^2}\right) \sim \chi^2(4)$$

$P(S > 41.7)$

$$\longrightarrow P(\chi^2 > 11.14) = \underline{\underline{0.025}}$$