Weibull distribution:

$$f(x,\lambda,\beta) = \begin{cases} \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(x/\beta)^{\alpha}} & ; x \geq 0 \\ 0 & , x < 0 \end{cases}$$

Pareto distribution:

$$f(x) = \frac{a b^a}{x^{a+1}} \quad ; x \geq 0$$

$$F(x) = CDF(x) = \int_b^x \frac{a b^a \, x^{-a}}{-a} \, dx = -b^a (x^{-a} - b) = 1 - \left(\frac{b}{x}\right)^a$$

$$\mu_n' = E(x^n) = \int_b^{\infty} x^{n-a-1} \, a \cdot b^a$$

$$= ab^a \cdot \left(\frac{x^{n-a}}{n-a}\right)_b^{\infty} = \frac{ab^a \cdot b^{n-a}}{n-a}$$

$$= \frac{ab^n}{n-a} \qquad 0 \leq n < a$$

$b < \lambda$

05/03/19

In the case of continuous distribution, let X denote the life of a component.

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

$= P($component fails before time $x)$

$$R_X(t) = P(X > t)$$

$= P($system is functioning at time $t)$

$=$ Reliability of the system at time t.

$= 1 - F_X(t)$

$P($system fails immediately sometime after time t given that it was working at time $t) =$

$= P(t < X \le t+h \mid X > t)$

$$H_X(t) = \lim_{h \to 0} \frac{P(t < X \le t+h)}{P(X > t)} \left(\frac{1}{h}\right)$$

= Instantaneous rate of failure at time '$\underline{t}$'.
/Hazard rate at time '$\underline{t}$'.

$$= \lim_{h \to 0} \frac{1}{h} \cdot \frac{F_X(t+h) - F_X(t)}{R_X(t)}$$

$$= \frac{f_X(t)}{R_X(t)}$$

$$\Rightarrow H_X(t) = -\frac{d}{dt} \log(1 - F_X(t))$$

$$\log(1 - F_X(t)) = -\int H_X(t)\,dt$$

$$\rightarrow 1 - F_X(t) = e^{-\int H_X(t)\,dt}$$

<u>Examples:</u>

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

$F_X(x) = 1 - e^{-\lambda x}$; $x > 0$

$H_X(t) = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda$, which is free from $\underline{t}$.

These are distributions which are increasing hazard rate. there can also be distribution with with decreasing Hazard.

$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} \quad , x, \alpha, \beta > 0$$

$$F_X(x) = \begin{cases} 1 - e^{-\alpha x^\beta} & ; x > 0 \\ 0 & ; x \le 0 \end{cases}$$

$$H_X(t) = -\frac{d}{dt}(e^{-\alpha t^\beta}) = (+\alpha\beta) t^{\beta-1}$$

For $\beta = 1$, reduces to exponential distribution.

$$\mu_K' = E(x^K) = \int_0^\infty \alpha\beta x^{\beta+K-1} e^{-\alpha x^\beta} dx$$

$$x^\beta = y \ , x = y^{1/\beta} , dx = \frac{1}{\beta} y^{\frac{1}{\beta}-1} dy$$

$$= \int_0^\infty \alpha\beta . \frac{1}{\beta} . y^{\frac{1}{\beta}-1} y^{\frac{K+\beta-1}{\beta}} e^{-\alpha y} dy$$

$$= \alpha \int_0^\infty y^{K/\beta} e^{-\alpha y} dy = \frac{\alpha \Gamma\left(\frac{K}{\beta}+1\right)}{\alpha^{K/\beta+1}}$$

$$= \alpha^{-K/\beta} \Gamma\left(\frac{K}{\beta}+1\right)$$

$$E(X) = \alpha^{-1/\beta} \Gamma\left(\frac{K}{\beta}+1\right)$$

$$E(X^2) = \alpha^{-2/\beta} \Gamma\left(\frac{K}{\beta}+1\right)$$

$$V(X) = \alpha^{-2/\beta}\left(\Gamma\left(\frac{2}{\beta}+1\right) - \left(\Gamma\left(\frac{1}{\beta}+1\right)\right)^2\right)$$

Example 1:

$$z(t) = 0.029 + 0.00025\,(t-40)^2, \quad t \geq 40.$$

Derive density func of life.

$P(X \geq 50)$ $\quad P\left(\frac{50 < X < 60}{X \geq 50}\right)$.

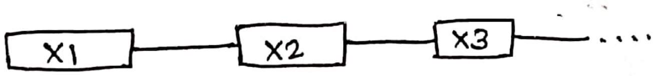
$$R(t) = e^{-\int_{40}^{t} z(s)\,ds}$$

$$R(50) = e^{-0.3533} = 0.702343$$

$$P(X > 60 \mid X > 50) = \frac{R(60)}{R(50)} \simeq 0.426$$

$$f_T(t) = -\frac{d}{dt} R(t) = \left(0.029 + 0.00025\,(t-40)^2\right) e^{-\left(0.0029\,(t-40) + \frac{0.00025}{3}(t-40)^3\right)}$$

series and parallel systems:



A series system functions if and only if all components in the system function.

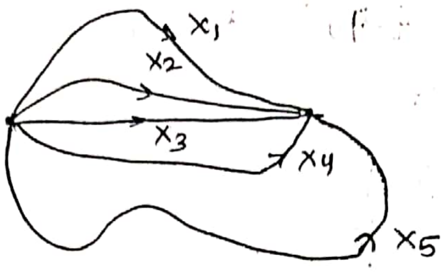
Let x be life of entire system, & $x_1, x_2, \ldots, x_n$ be life of n components.

$$R_X(t) = P(x > t)$$
$$= P(x_1 > t, x_2 > t, \ldots, x_n > t)$$

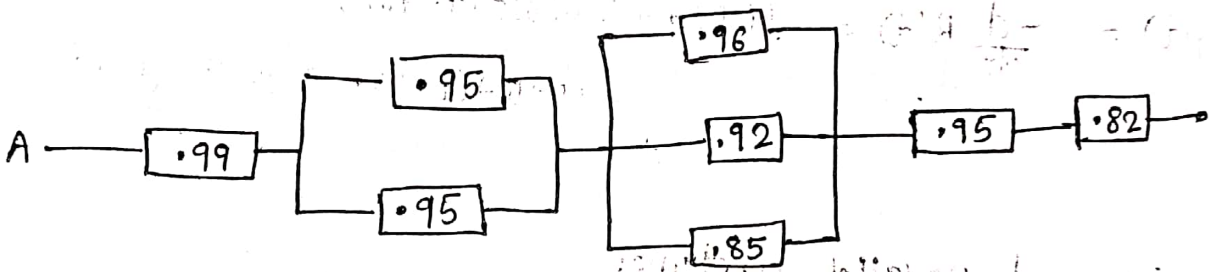
If we assume that the component lives are independently distributed then,

$$R_X(t) = \prod_{i=1}^{n} P(x_i > t) = \prod_{i=1}^{n} R_{X_i}(t)$$

parallel systems:



$$R_X(t) = P(X > t)$$
$$= 1 - P(X \le t)$$
$$= 1 - P(X_1 \le t)P_1(X_2 \le t) \ldots\ldots, P(X_n \le t)$$
$$= 1 - \prod_{i=1}^{n}(1 - R_{X_i}(t))$$



$R_s(t) = 0.7689$

$R_2(t) = 0.9975$

$R_3(t) = 0.99952$

Beta distribution:

A continuos r.v. X is said to have beta distribution if

$$f_X(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1} (1-x)^{\beta-1}$$

For different values of α and β, we get different distributions.

For $\alpha=1$ $\beta=1$, $f_X(x)=1$

For $\alpha=2$, $\beta=1$, $f_X(x) = 2x$

$0<x<1$

For $\alpha=1$, $\beta=2$ $f_X(x) = 2(1-x)$

For $\alpha=2$, $\beta=2$ $f_X(x) = 6x(1-x)$

Beta distributions is related to proportions etc.,

$$\mu_k' = E(x^k) = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{k+\alpha-1}(1-x)^{\beta-1}$$

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

$$\mu_1' = \frac{B(1+\alpha,\beta)}{B(\alpha,\beta)} = \frac{\Gamma 1+\alpha\ \Gamma\beta}{\Gamma 1+\alpha+\beta} \cdot \frac{\Gamma\alpha+\beta}{\Gamma\alpha\ \Gamma\beta}$$

$$= \frac{\alpha}{\alpha+\beta}$$

$$\mu_2' = \frac{\Gamma 1+\alpha (1+\alpha)\Gamma\beta}{1+\alpha+\beta\ \Gamma 1+\alpha+\beta} \cdot \frac{\Gamma\alpha+\beta}{\Gamma\alpha\ \Gamma\beta} = \frac{\alpha}{\alpha+\beta}\left(\frac{1+\alpha}{1+\alpha+\beta}\right)$$

$$V(X) = \mu_2' - (\mu_1')^2$$

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

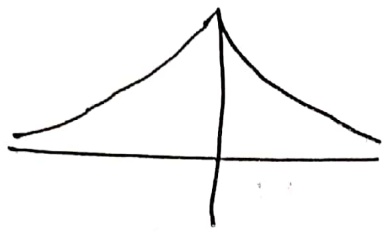
$$= \frac{(\alpha^2+\alpha)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Scanned by CamScanner

Double exponential (Laplace distribution)

$$f(x) = \frac{1}{2\sigma} e^{-\frac{1}{2}\frac{|x-\mu|}{\sigma}}$$

$-\infty < x < \infty$
$-\infty < \mu < \infty$
$\sigma > 0$



$E(X) = \mu$

$V(x) = 2\sigma^2$

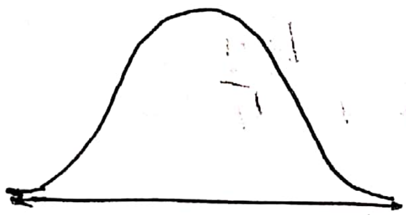
Median $(X) = \mu$

I. Normal (Gaussian) distribution:

A continuos r.v. X is said to have normal distribution with mean μ and variance σ^2 if it has density function of the form,

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

$-\infty < x < \infty$
$-\infty < \mu < \infty$
$\sigma > 0$



$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}(|x-\mu|^2)} dx$$

$$z = \frac{x-\mu}{\sigma} \qquad dx = \sigma dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} . \sigma dz \; e^{-z^2/2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz$$

$$= \underline{1}$$

Scanned by CamScanner

$$\mu_1' = E(X) = \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}\left(\frac{|x-\mu|}{\sigma}\right)^2}$$

$$= \int_{-\infty}^{\infty} \frac{\mu+\sigma z}{\sqrt{2\pi}} e^{-z^2/2} \, dz$$

$$= \frac{\mu}{\sqrt{2\pi}} \sqrt{2\pi} + \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} d\left(z^2/2\right) e^{-z^2/2} dz$$

$$= \cancel{\mu} + \frac{1}{\sqrt{2\pi}}(0) = \underline{\underline{\mu}}$$

$$\mu_K = E(X-\mu)^K$$

$$= \int_{-\infty}^{\infty} (x-\mu)^K \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx.$$

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

$$= \frac{\sigma^K}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^K e^{-z^2/2} \qquad = 0 \text{ if } K \text{ is odd.}$$

$$\frac{z^2}{2} = l \qquad z\,dz = dl$$

$$dz = \frac{dl}{\sqrt{2l}}$$

$$= 2\frac{\sigma^K}{\sqrt{2\pi}} \int_{0}^{\infty} (2l)^{K/2} \cdot e^{-l} \frac{dl}{\sqrt{2l}}$$

$$= \frac{2\sigma^K}{\sqrt{2\pi}} 2^{(K-1)/2} \int_{-\infty}^{\infty} e^{-l} dl \cdot e^{(K-1)/2}$$

$$= \frac{2\sigma^K 2^{(K-1)/2}}{\sqrt{2\pi}} \cdot \Gamma\left(\frac{K+1}{2}\right)$$

$$= \underline{\underline{\frac{\sigma^K 2^{K/2}}{\sqrt{\pi}} \cdot \Gamma\left(\frac{K+1}{2}\right)}}$$

$$\mu_2 = Var(X) = \underline{\underline{\sigma^2}} \qquad \underline{\underline{\mu_3 = 0}} \qquad \beta_1 = 0 \qquad \mu_4 = \frac{2^2 \cdot \sigma^4}{\sqrt{\pi}} \Gamma\left(\frac{5}{2}\right)$$

$$= \underline{\underline{3\sigma^4}}$$

Scanned by CamScanner

$$\beta_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{3\sigma^4}{\sigma^4} - 3 = 0$$

$$M_X(t) = E(e^{tX})$$

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

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

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{t\sigma z} e^{t\mu} e^{-z^2/2}$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dz.\ e^{t\sigma z - z^2/2 - \frac{t^2\sigma^2}{2}} e^{\frac{t^2\sigma^2}{2}} dz$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} e^{t^2\sigma^2/2} \int_{-\infty}^{\infty} dz.\ e^{-\frac{1}{2}(z - t\sigma)^2}$$

$$= \frac{e^{\mu t}}{\sqrt{2\pi}} e^{t^2\sigma^2/2} \sqrt{2\pi} = e^{\mu t + t^2\sigma^2/2}$$

Theorem: Let $X \sim N(\mu, \sigma^2)$ and let $Y = aX + b$.

Then $Y \sim N(a\mu + b, a^2\sigma^2)$

pf: $M_Y(t) = E\{e^{t(aX+b)}\}$

$$= e^{bt} E(e^{atX})$$

$$= e^{bt} M_X(at)$$

$$= e^{(b+\mu)t + a^2t^2\sigma^2/2}$$

$Z \sim N(0,1)$ is called standard normal distribution.

$$P(X \le t) = P\left(\frac{X-\mu}{\sigma} \le \frac{t-\mu}{\sigma}\right) = P\left(Z \le \frac{t-\mu}{\sigma}\right)$$

$\downarrow$

$Z \sim N(0,1)$

· pdf of a standard normal r.v. is denoted by

$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$

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

The cdf of z is denoted by:

$$\Phi(z) = \int_{-\infty}^{z} \phi(t)\,dt = \int_{-z}^{\infty} \phi(t)\,dt$$

$\Rightarrow \Phi(z) + \Phi(-z) = 1.$

$\Phi(0) = 1/2$

Let $X \sim P(\lambda)$, $\lambda > 0$

$M_X(t) = e^{\lambda(e^t - 1)}$

Let us define, $z = \frac{X-\lambda}{\sqrt{\lambda}}$

MGF of z

$$M_z(t) = E(e^{tz}) = E\left(e^{t\left(\frac{X-\lambda}{\sqrt{\lambda}}\right)}\right)$$

$$= e^{-\sqrt{\lambda}t} E(e^{tX/\sqrt{\lambda}}) \ ?$$

$$= e^{-t\sqrt{\lambda}} e^{\lambda(e^{t/\sqrt{\lambda}} - 1)}$$

$$= e^{-t\sqrt{\lambda}} \cdot e^{\lambda\left(1 + t/\sqrt{\lambda} + t^2/2\lambda + t^3/6\lambda\sqrt{\lambda} + \dots - 1\right)}$$

$$= e^{-t\sqrt{\lambda}} e^{t\sqrt{\lambda} + \frac{1}{2}t^2 + \frac{t^3}{6\sqrt{\lambda}} + \dots}$$

$$= e^{t^2/2} \text{ as } \lambda \to \infty.$$

which is mgf of $N(0,1)$.

Let $X \sim Bin(n,p)$ and $Z = \dfrac{X-np}{\sqrt{npq}}$

Then as $N \to \infty$, $Z \to N(0,1)$

$$M_Z(t) = \left(E\, e^{t} \cdot e^{t(x-np)/\sqrt{npq}} \right)^n$$

$$= e^{-npt/\sqrt{npq}} \left(q + p\, e^{t/\sqrt{npq}} \right)^n$$

Lognormal distribution:

If $X \sim N(\mu, \sigma^2)$ let us consider $Y = e^X$.

Then Y is said to have log-normal distribution.

$$f_Y(y) = \frac{1}{\sigma y \sqrt{2\pi}}\, e^{\frac{-1}{2\sigma^2}(\log y - \mu)^2}$$

$$E(Y^K) = e^{K\mu + \frac{1}{2}K^2\sigma^2}$$

$$E(Y) = e^{\mu + \sigma^2/2}$$

$$E(Y^2) = e^{2\mu + \sigma^2/2}$$

$$V(X) = e^{2\mu + \sigma^2}\left(e^{\sigma^2} - 1\right)$$

Function of a random variable

Let X be a r.v. defined on space $(\Omega, \mathcal{B}, P)$.

Let $Y = g(x)$ be a real valued fn.

Then Y is also a random variable.

$$F_Y(y) = P(Y \le y)$$

$$= P(g(x) \le y) = P(\{x : g(x) \le y\})$$

$$= P(g^{-1}(-\infty, y])$$

Examples: Let X have cdf $F_X(x)$

Let $Y_1 = |X|$ $\quad Y_2 = aX+b$ $\quad Y_3 = X^2$

$Y_4 = \log_e X$ $\quad Y_5 = e^X$ $\quad Y_6 = \max(X, 0)$

$$F_{Y_1}(y_1) = P(|X| \le y_1) = P(-y_1 \le X \le y_1)$$
$$= F(y_1) - F(-y_1)$$

$$F_{Y_2}(y_2) = P(aX+b \le y_1) = P\left(X \le \frac{y_1-b}{a}\right) = F\left(\frac{y_1-b}{a}\right)$$

$$F_{Y_3}(y_3) = P(X^2 \le y_3) = F(\sqrt{y_3}) - F(-\sqrt{y_3}\,-)$$

$$F_{Y_4}(y_4) = P(\ln X \le y_4) = F(e^{y_4})$$

$$F_{Y_5}(y_5) = F(\ln y_5)$$

$$F_{Y_6}(y_6) = F(\max(x, 0) \le y_6)$$

$$= P\left(\left\{\begin{array}{ll} 0 \le y_6 & \text{for } x \le 0 \\ x \le y_6 & \text{for } x \ge 0 \end{array}\right\} \le y_6\right)$$

$$= 0 \;;\; y_6 \le 0$$
$$= F(y_6) \;;\; y_6 \ge 0$$

Distribution of a function of a r.v. in discrete case:

$$P_X(x_i) = P(X = x_i)$$

$$Y = g(x) \quad g(x_i) = y_i$$

$$P(Y = y_i) = P(g(x) = y_i)$$

$$= \sum P(x_i | g(x_i) = y_i)$$

Let $X \sim Geo(p)$

$$P_X(K) = q^{K-1} p = (1-p)^{K-1} p$$

$$Y = \sin X$$

$$Y = \sin 1, \sin 2, \sin 3, \ldots\ldots$$

$$P(Y) = p \quad qp \quad qp^2 \ldots\ldots$$

$\sin K \neq \sin l$ for $\forall K, l \in \mathbb{N}$ $(K \neq l)$

For discrete, we need to look at the repeatation.

<u>continuos case:</u>

Let X be a continuos r.v. with pdf $f_X(x)$. Let $Y = g(X)$ be a real valued function of X.

case: g is monotonically increasing.

$$F_Y(y) = P(Y \leq y)$$

$$= P(g(x) \leq y)$$

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

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

$$f_Y(y) = \frac{\partial}{\partial y} F_X(g^{-1}(y)) = F_X'(g^{-1}(y)).(g^{-1}(y))'$$

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

~~case:~~ e.g., Let $X = N(\mu, \sigma^2)$ $\quad y = e^x$

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

$-\infty < x < \infty.$

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2} \cdot \frac{1}{y}$$

$$= \underline{\frac{1}{\sigma y\sqrt{2\pi}}} e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2} \qquad 0 < y < \infty.$$

case: g is monotonically decreasing.

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

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

Thus if g is a monotonically differentiable function.

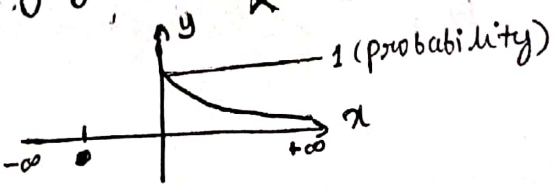
$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy} g^{-1}(y)\right|$$

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

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

$g(y) = e^{-\lambda x}$ $\qquad x = g^{-1}(y) = \frac{-\ln y}{\lambda}$

$\lambda x = -\ln g(y)$



$$f_Y(y) = \lambda e^{\ln y} \cdot \frac{-1}{\lambda y} = \underline{\underline{1}}$$

Let X be a continuos r.v. with pdf $f_X(x)$ and cdf $F_X(x)$.

Let $Y = F_X(x)$

So, F is increasing. $x = F^{-1}(y)$

cdf of Y,

$$G_Y(y) = P(F_X(x) \le y)$$

$$\cancel{P\left(\int_{-\infty}^{x} f_X(x)dx \le y\right)}$$

$$= P(x \le F_X^{-1}(y)) = F_X F_X^{-1}(y) = \underline{\underline{y}}$$

$$G_Y(y) = \begin{cases} 0, & y<0 \\ y; & 0 \le y \le 1 \\ 1, & y>1 \end{cases}$$ so $g_Y(y) = 1 \; ; 0<y<1$.

"probability density of cummulative probability is uniform".

Transformation of many valued function.

<u>Theorem:</u> Let X be a continuos r.v. with pdf $f_X(x)$. Let $Y = g(X)$ be a differentiable function and assume that $g'(x)$ is continuos and non-zero at all but a finite number of values of x. For every real y, let,

$$g(x_K(y)) = y \quad \text{for } k = 1,2 \ldots, n.$$

$$\& \quad g'(x_K(y)) \neq 0$$

$$f_Y(y) = \sum_{K=1}^{n} f(x_K(y)) \left| g'(x_K(y)) \right|^{-1}$$

Scanned by CamScanner

Example: $X \sim N(0,1)$

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

Let $Y = X^2 \rightarrow x = \sqrt{y}$ and $x = -\sqrt{y}$

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-y/2} \cdot \frac{1}{2\sqrt{y}} , & y > 0 \\ 0 ; & y \leq 0 \end{cases}$$

$$= \frac{1}{\sqrt{2}\, \Gamma(1/2)} y^{1/2 - 1} e^{-y/2}$$

$$= \text{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right)$$

Example: $f_X(x) = \frac{1}{2\sigma} e^{-|x|/\sigma} \qquad -\infty < x < \infty$

↳ Laplace distribution.

$y = |x|$?

$f_Y(y) = \frac{1}{\sigma} e^{-y/\sigma} ; \; y > 0$

$\sim \exp(1/\sigma)$

Jointly distributed random variables:

A random vector $\underline{X} = (X_1, \ldots, X_n)$ is a function from Ω to $\mathbb{R}^n$.

First we consider case $n=2$.

(X,Y) is jointly distributed.

$X \to$ Height

$Y \to$ Weight

Probability distribution of (X,Y):

The joint cdf of (X,Y) is given by,

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

properties:

1. $\lim_{y\to\infty} F_{X,Y}(x,y) = F_X(x) \quad \forall x$

   $\lim_{x\to\infty} F_{X,Y}(x,y) = F_Y(y) \quad \forall y$

2. $\lim_{x\to-\infty} F_{X,Y}(x,y) = \lim_{y\to-\infty} F_{X,Y}(x,y) = 0$

4. $F_{X,Y}(x,y)$ is non-decreasing in both x and y.

5. $F_{X,Y}(x,y)$ is continuos from right in both x and y.

(X,Y) discrete

$X \to x_1, x_2, \ldots$

$Y \to y_1, y_2, \ldots$

pmf satisfies:

(I). $0 \le P_{X,Y}(x_i, y_j) = P(X = x_i, Y = y_j) \le 1$

(II). $\sum_{x_i}\sum_{y_j} P(x_i, y_j) = 1$

<u>Example:</u> Suppose there is a car showroom having 10 cars. Five are good, two have defective transmission and ~~two~~ three have defective steering.

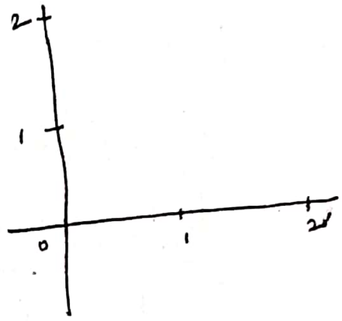
suppose two cars are selected at random.

X → no. of cars with DT

Y → no. of cars with DS.

$X = \{0, 1, 2\}$ $\quad Y = \{0, 1, 2, 3\}$

$P(0,0) = \dfrac{{}^5C_2}{{}^{10}C_2} = \dfrac{10}{45}$



$P(1,0) = \dfrac{{}^5C_1 \cdot {}^2C_1}{{}^{10}C_2} = \dfrac{10}{45}$

$P(2,0) = \dfrac{{}^5C_0 \, {}^2C_2}{{}^{10}C_2} = \dfrac{1}{45}$

$P(0,1) = \dfrac{{}^5C_1 \cdot {}^3C_1}{{}^{10}C_2} = \dfrac{15}{45}$

$P(0,2) = \dfrac{{}^3C_2}{{}^{10}C_2} = \dfrac{3}{45}$

$P(1,1) = \dfrac{6}{{}^{10}C_2} = \dfrac{6}{45}$

Rest all <u>zero</u>

y \ x	0	1	2	
0	10/45	10/45	[illegible]/45	2[illegible]/45
1	15/45	6/45	0	21/45
2	3/45	0	0	3/45
	28/45	16/45	[illegible]/45	

$P(Y=1) = \dfrac{{}^7C_1 \, {}^3C_1}{{}^{10}C_2} = \dfrac{21}{45}$ ✓

$P(X=1) = \dfrac{{}^8C_1 \cdot {}^2C_1}{{}^{10}C_2} = \dfrac{16}{45}$ ✓

Marginal distribution:

From joint pmf, we can find marginal pmf's of X and Y.

$$P_X(x_i) = \sum_{y_j \in Y} P_{X,Y}(x_i, y_j)$$

$$P_Y(y_i) = \sum_{x_i \in X} P_{X,Y}(x_i, y_j)$$

Conditional distributions:

The conditional pmf of X given $Y = y_j$ is given by,

$$P_{X|Y=y_j}(x_i|y_j) = P(X = x_i \mid Y = y_j)$$

$$= \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$$

Similarly the conditional distribution of Y given $X = x_i$ is defined as,

$$P_{Y|X=x_i}(y_j|x_i) = \frac{P(x_i, y_j)}{P(x_i)}, \quad y_j \in Y.$$

In previous Example,

$$P\left(\frac{X}{Y=0}\right) = \begin{cases} \frac{10}{21}; & x = 0 \quad P(0|0) \\ \frac{10}{21}; & x = 1 \quad P(0|0) \\ \frac{1}{21}; & x = 2 \quad P(2|0) \end{cases}$$

$$P_{Y|X=1} = \begin{cases} \frac{10}{16}; & y = 0 \rightarrow P_{Y|X=1}(1|1) \\ \frac{6}{16}; & y = 1 \\ 0; & 0/w \end{cases}$$

Scanned by CamScanner

Incase (X, Y) is jointly continuosly we have joint pdf $f_{X,Y}(x,y)$ satisfying:

(i) $f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2$

(ii) $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$

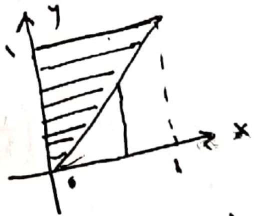
For any set $A \subset \mathbb{R}^2$

$$P((x,y) \in A) = \iint_A f_{X,Y}(x,y)\,dx\,dy$$

Ex: $f_{X,Y}(x,y) = \begin{cases} 10xy^2 & ; 0<x<y<1 \\ 0 & ; o/w \end{cases}$

$$\int_{y=0}^{1}\int_{x=0}^{y} 10xy^2\,dx\,dy = 10\int_0^1 y^2 dy \left(\frac{y^2}{2}\right) = \frac{5}{5}(y^5)_0^1 = 1$$

$\therefore$ Valid pdf.



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

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

$$f_{X|Y=y_0} = \frac{f_{X,Y}(x,y_0)}{f_Y(y_0)}$$

$$f_{Y|X=x_0} = \frac{f_{X,Y}(x_0,y)}{f_X(x_0)}$$

Scanned by CamScanner

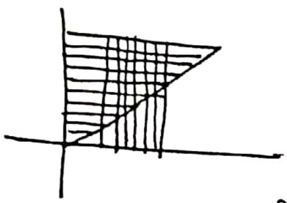
In previous example:

$$f_X(x) = \int_x^1 10xy^2\,dy = \frac{10x}{3}(1-x^3) \quad ; \; 0<x<1$$

$$f_Y(y) = \int_0^y 10xy^2\,dx = 10y^2\left(\frac{y^2}{2}\right) = 5y^4 \quad ; \; 0<y<1$$

$$P\left(\frac{1}{4}<X<\frac{1}{2}\right) = \int_{1/4}^{1/2} f_X(x)\,dx = \frac{5}{3}\left(\frac{3}{16}\right) - \frac{2}{3}\left(\frac{1}{32}\right)\left(\frac{31}{32}\right)$$

$$= \frac{5}{16} - \frac{31}{3\times 512}$$



$$P\left(\frac{1}{3}<Y<\frac{2}{3}\right) = \int_{1/3}^{2/3} f_Y(y) = \frac{31}{243}$$

$$f^{(x,y)}_{X|Y=y} = \frac{f(x,y)}{f(y)} = \frac{10xy^2}{5y^4} = \frac{2x}{y^2} \qquad \begin{matrix} 0<x<y \\ 0<y<1 \end{matrix}$$

$$f^{(x,y)}_{X|y=1/2} = \begin{cases} 8x \; ; & 0<x<1/2 \\ 0 \; ; & o/w \end{cases}$$

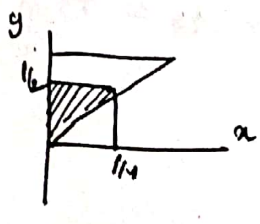
$$P\left(X<\frac{1}{4} \,\middle|\, Y=\frac{1}{2}\right) = 4\left(\frac{1}{16}\right) = 0.25$$

conditional pdf of Y when X = x

$$f^{(y|x)}_{Y|X=x} = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{10xy^2}{\frac{10x}{3}(1-x^3)} = \frac{3y^2}{1-x^3}$$

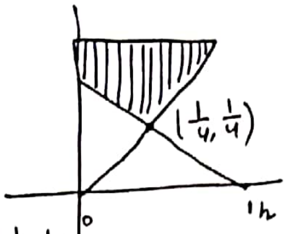
$$x<y<1$$
$$0<x<1$$

$$P\left(Y>\frac{2}{3} \,\middle|\, X=\frac{1}{2}\right) = \frac{24}{7}\int_{2/3}^{1} y^2\,dy = \frac{152}{189}$$



$$P\left(X<\frac{1}{4}, Y<\frac{1}{2}\right) = \int_{x=0}^{1/4}\int_{y=x}^{1/2} 10xy^2$$

$P(X+Y > \frac{1}{2})$



$$= \int_{x=0}^{1/4} \int_{y=\frac{1}{2}-x}^{1} 10\, xy^2\, dx + \int_{x=\frac{1}{4}}^{1/2} \int_{y=x}^{1} 10 xy^2 \,dx\,dy$$

(or) $P(X+Y > \frac{1}{2}) = 1 - P(X+Y \leq 1/2)$

Product moments:

$$\mu'_{r,s} = E(x^r y^s)$$

$$= \begin{cases} \sum\sum_{(x_i, y_j)} x_i^r y_j^s \; P_{X,Y}(x_i, y_j) \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f(x,y)\, dx\, dy \end{cases}$$

provided the corresponding series are integral convergent.

let $E(X) = \mu_X$ $E(Y) = \mu_Y$

$\mu_{r,s} = E(X-\mu_X)^r (Y-\mu_Y)^s = cov(X,Y)$

this is called covariance between (X,Y). at $r=1$, $s=1$.

We define coefficient of correlation between (X,Y)

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{var(X)\, var(Y)}} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

$\sigma_X^2 = Var(X)$

$\sigma_Y^2 = Var(Y)$

Let us consider,

$$U = \frac{X-\mu_x}{\sigma_x}$$

$$V = \frac{Y-\mu_y}{\sigma_y}$$

Let us consider $E(U-V)^2 \geq 0$

$\Rightarrow \quad EU^2 + EV^2 - 2E(UV) \geq 0$

$\Rightarrow \quad E(UV) \leq \dfrac{E(U^2)+E(V^2)}{2} \quad \ldots\ldots (1)$

Let us again consider $E(U+V)^2 \geq 0$

$\Rightarrow \quad E(UV) \geq -\left(\dfrac{E(U^2)+E(V^2)}{2}\right) \quad \ldots\ldots (2)$

combining (1) and (2);

$$-\frac{1}{2}\left(E(U^2)+E(V^2)\right) \leq E(UV) \leq \frac{1}{2}\left(E(U^2)+E(V^2)\right)$$

$$U = \frac{X-\mu_x}{\sigma_x} \; ; \; E(U) = \frac{E(X)-E(X)}{\sigma_x} = 0$$

$$E(U^2) = \frac{E(X-\mu_x)^2}{\sigma_x^2} = \frac{\sigma_x^2}{\sigma_x^2} = 1$$

$$E(V)=0 \; ; \; E(V^2) = 1$$

$$E(UV) = \frac{E(X-\mu_x)(Y-\mu_y)}{\sigma_x \sigma_y} = \frac{cov(X,Y)}{\sigma_x \cdot \sigma_y} = \rho_{x,y}$$

$\therefore \quad \boxed{-1 \leq \rho_{x,y} \leq 1}$

For equality at ± 1 we must have, $P\left(\dfrac{X-\mu_x}{\sigma_x} = \dfrac{Y-\mu_y}{\sigma_y}\right) = 1$

$\Rightarrow P(X = cY + d) = 1$

c & d are real

i.e., X and Y perfectly positively linearly related.

For equality at $\underline{-1}$ $\quad P(U=-V)=1$

$\Rightarrow \quad P(X = cY+d) = 1$

c & d are real & $c<0$

i.e., X and Y are perfectly negatively linearly related.

$\therefore$ coefficient of correlation can be considered as a measure of linear relationship between X and Y.

$U = aX+b$

$V = cY+d$

$$\rho_{U,V} = \frac{cov(U,V)}{\sigma_U \sigma_V} = \frac{E(U-E(U))(V-E(V))}{\sigma_U \sigma_V}$$

$$= \frac{E(aX+b-aE(X)-b)(c\cancel{X}+d-cE(Y)-d)}{\sqrt{E(U-E(U))^2}\sqrt{E(V-E(V))^2}}$$

$$= \frac{ac\, cov(X,Y)}{\sqrt{a^2\sigma_X^2 c^2\sigma_Y^2}}$$

$$= \rho_{X,Y} \cdot \frac{ac}{|ac|}$$

$$= \begin{cases} \rho_{X,Y} \ ; & ac>0 \\ -\rho_{X,Y} \ ; & ac<0 \end{cases}$$

Ex:

$$f_{X,Y}(x,y) = \begin{cases} x+y \\ 0 \end{cases}$$

(Test)

Independent random variable:

We say that random variables X & Y are independent if

$F_{X,Y}(x,y) = F_X(x) F_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$

For discrete distribution,

$$P_{X,Y}(x,y) = P_X(x_i) P_Y(y_j) \quad \forall (x_i, y_j)$$

For continuous distribution,

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$$

If random variables X & Y are independent then,

$$E(X^r Y^s) = E(X^r) E(Y^s)$$

and $\operatorname{cov}(X,Y) = E(X-\mu_X)(Y-\mu_Y)$

$$= E(X-\mu_X)\, E(Y-\mu_Y)$$

$$= \underline{0}$$

If random variables X and Y are independent, then their correlation is zero. $\rho_{X,Y} = 0$

Converse of the above statement is not true, X and Y may be dependent and $\rho_{X,Y}$ may be zero.

Eg:

X \ Y	-1	0	1
0	0	1/3	0
1	1/3	0	1/3

$$\mathrm{cov}(X,Y) = E(X-\mu_X)(Y-\mu_Y)$$
$$= E(XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y)$$
$$= E(XY) - E(X)E(Y)$$
$$= \frac{1}{3}(0) + \frac{1}{3}(-1) + \frac{1}{3}(1) - (0)(0)$$
$$= \underline{\underline{0}}$$

F

$$P_{X,Y}(0,1) = 0 \neq P_X(0)P_Y(1) = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) = \underline{\underline{\frac{1}{9}}}$$

Joint MGF:

$$M_{X,Y}(s,t) = E(e^{sX+tY})$$

Theorem: X and Y are independent

$$\Leftrightarrow M_{X,Y}(s,t) = M_X(s)M_Y(t) \quad \forall (s,t) \in \mathbb{R}^2$$

Theorem: If X and Y are independent and $S = X+Y$

$$M_S(t) = M_X(t)M_Y(t)$$

Bivariate Normal Distribution:

(X,Y) is said to have bivariate normal distribution if it has pdf of the form,

heir

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{\frac{-1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right)}$$

l may

$(x_1, x_2) \in \mathbb{R}^2$

$(\mu_1, \mu_2) \in \mathbb{R}^2$

$(\sigma_1, \sigma_2) \in \mathbb{R}^2_{(+,+)}$

$-1 < \rho < 1$

Scanned by CamScanner

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\,\sigma_1} e^{\frac{-1}{2(1-\rho^2)\sigma_1^2}\left[x-\left\{\mu_1+\frac{\rho\sigma_1}{\sigma_2}(y-\mu_2)\right\}\right]^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_2} e^{-\frac{1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2}$$

So marginal pdf of Y is $N(\mu_2, \sigma_2^2)$.

So conditional distribution of X given $Y=y$ is,

$$N\left(\mu_1 + \frac{\rho\sigma_1}{\sigma_2}(y-\mu_2),\ \sigma_1^2(1-\rho^2)\right)$$

So conditional distribution of Y given $X=x$ is,

$$N\left(\mu_2 + \frac{\rho\sigma_2}{\sigma_1}(x-\mu_1),\ \sigma_2^2(1-\rho^2)\right)$$

Theorem: If (X,Y) has BVN then the ~~magnitude~~ marginals and conditionals are univariate normal the converse is also true.

$$E(X) = \mu_1$$
$$E(Y) = \mu_2$$
$$E(X \mid Y=y) = \mu_1 + \frac{\rho\sigma_1}{\sigma_2}(y-\mu_2)$$
$$V(X \mid Y=y) = \sigma_1^2(1-\rho^2)$$
$$E(Y \mid X=x) = \mu_2 + \frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)$$
$$V(Y \mid X=x) = \sigma_2^2(1-\rho^2)$$

Let (X,Y) be joint distribution of

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} E\{g(x,y) \mid Y=y\}\, f_Y(y)\,dy = E(E(g(x,y) \mid x))$$

• so if $(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

Then $\text{cov}(X,Y) = E(X-\mu_X)(Y-\mu_Y)$

$$= E\left\{ E(X-\mu_X)(Y-\mu_Y) \mid Y \right\}$$

$$= E\left[(Y-\mu_2)\, E\{(x-\mu_1) \mid y\} \right]$$

$$= E(Y-\mu_2)^2 \cdot \frac{\rho\sigma_1}{\sigma_2} = \frac{\rho\sigma_1\sigma_2^2}{\sigma_2}$$

$\therefore$ $\text{cov}(X,Y) = \rho\sigma_1\sigma_2$

$\text{corr}(X,Y) = \frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2} = \rho$

$$MGF = E[e^{sX} M_{Y|X}(t)]$$

$$= E\left[e^{sX} e^{\{\mu_2 + \frac{\rho\sigma_2}{\sigma_1}(X-\mu_1)\} t + \sigma_2^2(1-\rho^2)t^2}\right]$$

$$= e^{\mu_2 t - \frac{\rho\sigma_2}{\sigma_1}\mu_1 t + \sigma_2^2(1-\rho^2)t^2} E\left(e^{(s+\frac{\rho\sigma_2}{\sigma_1})X}\right)$$

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

$$= e^{\mu_1 s + \mu_2 t + \frac{1}{2}\sigma_1^2 s^2 + \frac{1}{2}\sigma_2^2 t^2 + \rho\sigma_1\sigma_2 st}$$

" common class test on 2^{nd} april"

Theorem: Let $(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then X, Y are independent iff $\rho = 0$.

Pf:

Let $\rho = 0$, $M_{X,Y}(s,t) = e^{\mu_1 s + \mu_2 t + \frac{1}{2}\sigma_1^2 s^2 + \frac{1}{2}\sigma_2^2 t^2}$

$= M_X(s) M_Y(t)$

Hence, X and Y are independent.

Theorem: $(X,Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$\Leftrightarrow aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2)$

Pf: $U = aX + bY$

$M_U(t) = E(e^{tU}) = E\{e^{t(aX+bY)}\}$

$= E\{e^{(at)X + (bt)Y}\}$

$= M_{X,Y}(at, bt)$

$= e^{\mu_1 at + \mu_2 bt + \frac{1}{2}\sigma_1^2 a^2 t^2 + \frac{1}{2}\sigma_2^2 b^2 t^2 + \rho\sigma_1\sigma_2 ab t^2}$

$\sim N(a\mu_1 + b\mu_2;\ a^2\sigma_1^2 + b^2\sigma_2^2 + 2\rho\sigma_1\sigma_2 ab)$

Scanned by CamScanner

Random vectors:

$X = (X_1, X_2, \ldots, X_K) : \Omega \rightarrow \mathbb{R}^K$

$F_{\underline{X}}(x) = P(X_1 \le x_1, X_2 \le x_2, \ldots, X_K \le x_K)$

$$\lim_{x_i \to -\infty} F_{\underline{X}}(\underline{x}) = 0 \quad \forall i = 1, \ldots, K$$

$$\lim_{x_i \to \infty} F_X(x) = F(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_K)$$

$F_{\underline{X}}(\underline{x})$ is non-decreasing in each of its arguments and continuous from right in each of its arguments.

If $X_1, X_2 \ldots, X_K$ are jointly discrete, the we have pmf

$P_X(\underline{x}) = p_{X_1, \ldots, X_K}(x_1, \ldots, x_K)$

$p(X_1 = x_1, \ldots X_K = x_K) = P_{\underline{X}}(\underline{x})$

one can deduce all combinations of marginal and conditional pmfs.

If $X_1, X_2, \ldots, X_K$ are jointly continuous, then we have pdf,

$f_{\underline{X}}(x) = f(x_1, x_2 \ldots, x_K)$

The joint mgf $\underline{X} = (X_1, X_2, \ldots, X_K)$ at the point $\underline{t} = (t_1, \ldots t_K)$ is defined as,

$$M_{\underline{X}}(\underline{t}) = E\left(e^{\sum_{i=1}^{K} t_i X_i}\right)$$

are said to be mutually independent if

$$M_{\underline{X}}(\underline{t}) = \prod_{i=1}^{K} M_{X_i}(t_i) \quad \forall\, t \in \mathbb{R}^K$$

Equivalenty. $P_{\underline{X}}(\underline{x}) = \prod_{i=1}^{K} p_{X_i}(x_i)$

Theorem: Let $X_1, X_2, \ldots, X_n$ be independent rv's and $S_n = X_1 + X_2 + \ldots + X_n$ then

$$M_{S_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

Pf: $$M_{S_n}(t) = E(e^{tS_n}) = \prod_{i=1}^{n} E(e^{tX_i})$$

$$= \prod_{i=1}^{n} M_{X_i}(t)$$

Additive properties of some distributions:

1. Binomial:

Let $X_1, X_2, \ldots, X_K$ be independent

$X_i \sim Bin(n_i, p)$

Then $S_n \sim Bin(\Sigma n_i, p)$

Pf: $$M_{S_n}(t) = \prod_{i=1}^{K} M_{X_i}(t) = \prod_{i=1}^{K} (q + pe^t)^{n_i}$$

$$= (q + pe^t)^{\Sigma n_i}$$ which is mgf of $Bin(\Sigma n_i, p)$.

2. Poisson:

Let $X_1, \ldots, X_n$ be independent

$X_i \sim P(\lambda_i)$

Then $S_n \sim P(\Sigma \lambda_i)$

Pf: $$M_{S_n}(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} e^{\lambda(e^t - 1)}$$

$$= P(\Sigma \lambda_i)$$

3. Geometric:

$X_i \sim Geo(p)$ $\quad$ $S_r \sim NB(r, p)$

Pf: $$M_{S_r}(t) = \prod_{i=1}^{r} \left(\frac{pe^t}{1 - qe^t}\right) = \left(\frac{pe^t}{1 - qe^t}\right)^r$$ which is negative Binomial.

Scanned by CamScanner

4. $X \sim NB(r_i, p)$

Then $S_K = NB(\sum r_i, p)$

5. Let $X_1, \ldots, X_r \sim Exp(\lambda)$

Then $S_r \sim Gamma(r, \lambda)$

pf: $M_{S_r}(t) = \prod_{i=1}^{r} M_{X_i}(t) = \prod_{i=1}^{r} \left(\frac{\lambda}{\lambda - t}\right)$

$= \left(\frac{\lambda}{\lambda - t}\right)^r$ which is mgf of $Gamma(r, \lambda)$

6. $X_i \sim Gamma(r_i, \lambda)$ $i = 1, 2, \ldots, K$

Then $S_K \sim Gamma(\sum r_i, \lambda)$

7. linearity property of normal distribution:

let $X_1, \ldots, X_n$ be independently distributed with

$X_i \sim N(\mu_i, \sigma_i^2)$ $i = 1, 2, \ldots, n$

$Y = \sum_{i=1}^{n} (a_i X_i + b_i)$

$M_X(t) = E\, e^{t\sum(a_i X_i + b_i)}$

$= e^{t\sum b_i} \prod_{i=1}^{n} M_{X_i}(at)$

$= e^{t\sum b_i} \left[\prod_{i=1}^{n} e^{\mu_i a_i t + \frac{1}{2}\sigma_i^2 a_i^2 t^2} \right]$

$= e^{t\sum b_i}\, e^{t\sum a_i \mu_i + \frac{1}{2}t^2 \sum a_i^2 \sigma_i^2}$

$= e^{t(\sum(a_i\mu_i + b_i)) + \frac{1}{2}t^2(\sum a_i^2 \sigma^2)}$

$\sim$ mgf of $N(\sum(a_i\mu_i + b_i), \sum a_i^2 \sigma_i^2)$

Ex: Let $X_1, X_2, \ldots, X_n \overset{IND}{\sim} N(\mu, \sigma^2)$

$$Y = \bar{X} = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$$

$$\sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Functions of several random variables:

Ex: Let $X_1, \ldots, X_n \overset{iid}{\sim} Exp(1)$

so the joint pdf of $\underset{\sim}{X} = (X_1, X_2, X_3)$ is

$$f_X(x) = \begin{cases} e^{-(x_1+x_2+x_3)} &, x_1>0 \ x_2>0 \ x_3>0 \\ 0 &, \text{otherwise} \end{cases}$$

$$Y_1 = X_1 + X_2 + X_3$$

$$Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3}$$

$$Y_3 = \frac{X_1}{X_1 + X_2}$$

$$x_1 = y_1 y_2 y_3$$

$$x_2 = y_1 y_2 (1 - y_3)$$

$$x_3 = y_1 (1 - y_2)$$

$$J = \begin{vmatrix} y_2 y_3 & y_1 y_3 & y_1 y_2 \\ y_2(1-y_3) & y_1(1-y_3) & -y_1 y_2 \\ 1-y_2 & -y_1 & 0 \end{vmatrix}$$

$$= -y_1^2 y_2$$

$y_1 > 0;\ 0 < y_2 < 1;\ 0 < y_3 < 1$

The joint pdf of $\underset{\sim}{Y} = (Y_1, Y_2, Y_3)$ is

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \begin{cases} e^{-y_1} y_1^2 y_2 &; y_1 > 0,\ 0 < y_2 < 1,\ 0 < y_3 < 1. \\ 0 &; o/w \end{cases}$$

$$f_{Y_1}(y_1) = \frac{1}{2} e^{-y_1} y_1^2 ;\ y_1 > 0$$

$$f_{Y_2}(y_2) = \begin{cases} 2y_2 , & 0 < y_2 < 1 \\ 0 , & o/w \end{cases}$$

$$f_{Y_3}(y_3) = \begin{cases} 1 , & 0 < y_3 < 1 \\ 0 , & o/w \end{cases}$$

Distribution of order statistics:

Let $X_1, X_2, \dots, X_n \overset{iid}{\sim} f(x)$

$X_{(1)} = \min(X_1, \dots, X_n)$

$X_{(2)} = \text{second} \min(X_1, \dots, X_n)$

$\vdots$

$X_{(n)} = n^{th} \min(X_1, \dots, X_n)$

So we have $n!$ inverses

$x_1 = y_1$

$x_2 = y_2$

$\vdots$

$x_n = y_n$

$$J = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdots & & & \\ \cdots & & 0 & 1 \end{vmatrix}$$

$|J| = 1$

$x_1 = y_2$

$x_2 = y_1$

$x_3 = y_3$

$x_n = y_n$

$\rightarrow |J| = 1$

The joint pdf of $X_1, X_2, \dots, X_n$ is

$$f_{\underline{x}}(\underline{x}) = \prod_{i=1}^{n} f(x_i)$$

The joint pdf of $Y_1, \dots, Y_n$ is

$$f_{\underline{y}}(\underline{y}) = n! \prod_{i=1}^{n} f(y_i) \; ; \quad -\infty < y_1 < y_2 \cdots < y_n < \infty$$

$$f_{Y_1}(y_1) = n(1 - F(y))^{n-1} f(y)$$

$$f_{Y_n}(y_n) = n(F(y))^{n-1} f(y_n)$$

Sampling distributions

Population: A statistical population is a collection of measurements, it could be qualitative or quantitative.

Sample: Subset of population.

Random sample: Each unit of the population has the same probability of getting selected.

We say that the population has a theoritical distribution P_θ. $\theta \subset \Omega$.

Let $X_1, X_2, \ldots, X_n$ be a random sample from this population. That is each has the same distribution P_θ.

$X_1, X_2, \ldots, X_n$ are iid.

If $p(x)$ is pmf of P_θ.

Then the joint pmf of $X_1, X_2, \ldots, X_n$ is

$$\prod_{i=1}^{n} p(x_i)$$

If $f(x)$ is pdf of P_θ.

Then the joint pdf of $X_1, X_2, \ldots, X_n$ is

$$\prod_{i=1}^{n} f(x_i)$$

A function of random sample $T(x_1, x_2, \ldots, x_n)$ is called a statistic.

For example, sample mean $\bar{x}$

sample variance $s^2 = \frac{1}{(n-1)} \Sigma (x_i - \bar{x})^2$

sample range $X_{(n)} - X_{(1)}$.

The distribution of a statistic is called a sampling distribution.

central limit theorem:

Let $X_1, X_2, \ldots, X_n$ be sequence of iid's with mean μ and variance σ^2. Then the limiting distribution is,

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \text{ is } N(0,1) \text{ as } n \to \infty$$

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i$$

In terms of summation, we express the above limit as

$$\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \to Z \sim N(0,1) \text{ as } n \to \infty$$

$X_1, X_2, \ldots, X_n$ are sequence of iid's with mean μ_1 and variance σ_1^2.

$Y_1, \ldots, Y_n$ are sequence of iid's with mean μ_2 and variance σ_2^2.

$$\bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i \quad ; \quad \bar{Y} = \frac{1}{n_2} \sum_{j=1}^{n} Y_j$$

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \longrightarrow N(0,1) \quad \text{as } n_1, n_2 \to \infty.$$

Chi-square distribution:

A r.v. X is said to have a chi square distⁿ on n-degrees of freedom if it has pdf,

$$f(x) = \frac{1}{2^{n/2}\Gamma\frac{n}{2}} e^{-x/2} x^{\frac{n}{2}-1}, \quad x>0$$

$$X \sim \chi^2_n$$

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

$$Y = X^2 \sim \chi^2_1$$

$$M_X(t) = (1-2t)^{-n/2}$$

$$E(X) = n$$

$$V(X) = 2n$$

$X_1, X_2, \dots, X_n$ are independent $N(0,1)$

Then $W = \sum_{i=1}^{n} X_i^2 \sim \chi^2_n$.

$$s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

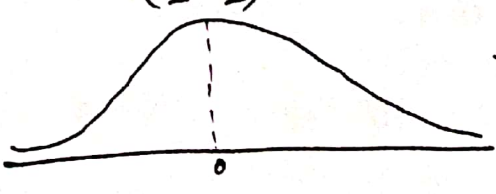
Also $\bar{X}$ and s^2 are independent distributions.

A random variable T is said to have students t distribution of n d.f. of it has pdf,

$$f_T(t) = \frac{1}{\sqrt{n}\, B\left(\frac{n}{2}, \frac{1}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}; \quad t \in \mathbb{R}$$

$T \sim t_n$

$T \to N(0,1)$ as $n \to \infty$



$X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$

$$\frac{\sqrt{n}(\bar{X}-\mu)}{s} \sim t_{n-1}$$

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

$Y \sim \chi^2_n$

Then the distribution $T = \dfrac{X}{\sqrt{\frac{Y}{n}}}$, is t_n

$E(T) = 0$; $V(T) = \dfrac{n}{n-2}$, $n > 2$

$\beta_1 = 0$; $\beta_2 = \dfrac{6}{n-4}$, $n > 4$.

F- distribution:

If $\left.\begin{matrix} X_1 \sim \chi^2_m \\ X_2 \sim \chi^2_n \end{matrix}\right]$ independent

$U = \dfrac{nX_1}{mX_2}$ is said to have F- dist^n on (m,n) if

$$f(u) = \left(\frac{m}{n}\right)^{m/2} \cdot \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \cdot \frac{u^{m/2 - 1}}{\left(1 + \frac{mu}{n}\right)^{(m+n)/2}} \quad ; u > 0$$