

Chapter one

(Probability distributions)

The probability distributions divided into two types which they are :-

1. Discrete probability distⁿ.
2. Continuous Probability distⁿ.

* Discrete Type:-

1- Discrete uniform distribution

the r.v. X is uniformly distⁿ $ud(k)$ if and only if the prob. density function p.d.f. has the following form :-

$$f(x) = P(X=x) = \begin{cases} \frac{1}{k} & x=1, \dots, k \\ 0 & \text{o.w.} \end{cases}$$

To verify that $f(x)$ is a p.d.f., we note that:

1. $f(x) > 0 \quad \forall x \in \{x=1, \dots, k\}$

2. $\sum_x f(x) = 1 \Rightarrow \sum_{x=1}^k \frac{1}{k} \Rightarrow \frac{1}{k} \sum_{x=1}^k 1 \Rightarrow \frac{1}{k} \cdot k = 1$

The c.d.f of X :-

$$F(x) = P(X \leq x) = \sum_{t=1}^x f(t) = \sum_{t=1}^x \frac{1}{k} = \begin{cases} 0 & x < 1 \\ \frac{x}{k} & 1 \leq x \leq k \\ 1 & x \geq k \end{cases}$$

• mean (μ_x) and variance (σ_x^2) of X :-

$$\underbrace{E(X)}_{\mu_x} = \sum_x x \cdot f(x) = \frac{1}{k} \sum_{x=1}^k x = \frac{1}{k} \cdot \frac{k(k+1)}{2} = \boxed{\frac{k+1}{2}}$$

$$\underbrace{V(X)}_{\sigma_x^2} = E(X^2) - (E(X))^2$$

$$E(X^2) = \sum_{x=1}^k x^2 \cdot f(x) = \frac{1}{k} \sum_{x=1}^k x^2 = \frac{1}{k} \cdot \frac{k(k+1)(2k+1)}{6}$$

$$\therefore \sigma_x^2 = V(X) = \boxed{\frac{k^2 - 1}{12}}$$

example (1): let a r.v. $X \sim \text{Ud}(4)$, ① write down a p.d.f, c.d.f of X . ② evaluate $P(1 < X \leq 3)$

sol.

$$\textcircled{1} P(X) = \begin{cases} \frac{1}{4} & X=1, \dots, 4 \\ 0 & \text{o.w.} \end{cases}, F(X) = \begin{cases} 0 & X < 1 \\ \frac{x}{4} & 1 \leq X < 4 \\ 1 & X \geq 4 \end{cases}$$

② using p.d.f. 3

$$P(1 < X \leq 3) = \sum_{x=2}^3 f(x) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

using c.d.f.

$$\begin{aligned} F(X) = F(3) - F(1) &\Rightarrow P(X \leq 3) - P(X \leq 1) \\ &\Rightarrow \frac{3}{4} - \frac{1}{4} = \frac{1}{2} \end{aligned}$$

let a r.v. $X \sim \text{Ud}(N)$, with $\text{var}(X) = 2$, find N : (H.W.)

2. Bernoulli distⁿ.

The r.v. (X) is said to have Bernoulli distⁿ, denoted by $X \sim \text{Ber}(p)$ if X has a p.d.f.:-

$$f(x) = \begin{cases} p^x q^{1-x} & x = 0, 1 \\ 0 & \text{o.w.} \end{cases}, \quad \begin{matrix} p+q=1 \\ q=1-p \\ p=1-q \end{matrix}, \quad 0 \leq p \leq 1$$

Note:- Any random experiment whose outcomes have been classified into two categories "success" and "failure" is called Bernoulli trial. If a r.v. (X) is defined as:

$$X = \begin{cases} 0 & , \text{ if Bernoulli trial results in failure } (q) \\ 1 & , \text{ if Bernoulli trial results in success } (p) \end{cases}$$

To verify that $f(x)$ is a p.d.f., we note that:

1. $f(x) > 0 \quad \forall x \in \{x=0, 1\}$

2. $\sum_x f(x) = 1 \Rightarrow \sum_{x=0}^1 p^x q^{1-x} \Rightarrow p^0 q + p q^0 = p+q=1$

The c.d.f. of X:-

$$F(x) = P(X \leq x) = \sum_{t=0}^x p^t q^{1-t} = \begin{cases} 0 & x < 0 \\ q & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

The m.g.f. of X:-

$$M_X(t) = E e^{tx} = \sum_{x=0}^1 e^{tx} p^x q^{1-x} = q + p e^t, \quad -\infty < t < \infty$$

• Mean and variance of X .

$$\mu_X = E(X) = \sum_{x=0}^1 x \cdot f(x) \Rightarrow \sum_{x=0}^1 x \cdot p^x q^{1-x} = \textcircled{p}$$

$$\sigma_X^2 = V(X) = E(X^2) - (E(X))^2$$

$$E(X^2) = \sum_{x=0}^1 x^2 \cdot f(x) \Rightarrow \sum_{x=0}^1 x^2 \cdot p^x q^{1-x} = p$$

$$\therefore V(X) = p - p^2 = p(1-p) = \textcircled{pq}$$

H.w.: Verify the mean and variance of X by considering:- ① The differentiation of m.g.f. at $t=0$. ② The differentiation of $\ln$ m.g.f. at $t=0$.

Example (2): Let a r.v. $X \sim \text{Ber}(\frac{1}{3})$. Write down the p.d.f, m.g.f of X and μ_X , σ_X^2 .

Sol.: $p = \frac{1}{3}$, $q = \frac{2}{3}$, because $p+q=1$

$$\therefore f(x) = \begin{cases} (\frac{1}{3})^x (\frac{2}{3})^{1-x} & x=0,1 \\ 0 & \text{o.w.} \end{cases}$$

$$\mu_X(t) = \frac{2}{3} + \frac{1}{3} e^t, \quad \mu_X = p = \frac{1}{3}, \quad \sigma_X^2 = pq = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

example (3):- let the r.v. X has m.g.f. $\mu_X(t) = \frac{1}{4} + \frac{3}{4} e^t$
what is the distⁿ of X . Find μ_X and σ_X^2

Sol.: $M_X(t) = \frac{1}{4} + \frac{3}{4} e^t$ is the m.g.f. for $X \sim \text{Ber}(\frac{3}{4})$

$$\therefore \mu_X = E(X) = p = \frac{3}{4}, \quad \sigma_X^2 = V(X) = pq = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$$

3- Binomial distribution

A r.v. X is said to have Binomial distⁿ. denoted by $X \sim b(n, p)$, if X has p.d.f.:

$$f(x) = P(X=x) = \begin{cases} C_x^n p^x q^{n-x} & x=0, 1, 2, \dots, n \\ 0 & \text{o.w.} \end{cases} \quad \begin{matrix} q=1-p \\ 0 < p < 1 \end{matrix}$$

where

n : number of trials or sample size.

X : number of successes, $n-X$ is number of failures.

Remark:

1- the Bernoulli distⁿ. $\text{Ber}(p)$ is a Binomial distⁿ. $b(n, p)$ where $n=1$

2- In general
$$\sum_{x=0}^n C_x^n p^x q^{n-x} = (p+q)^n$$

this is called the Binomial expansion.

Now to check a $f(x)$ is valid p.d.f., we note that:

1. $f(x) > 0$, $\forall x \in \{x=0, \dots, n\}$

2. $\sum_x f(x) = 1 \Rightarrow \sum_{x=0}^n C_x^n p^x q^{n-x} = (p+q)^n = (1)^n = 1$

Binomial expansion.

• Mean and variance of X :-

$$\mu_X \equiv E(X) = \sum x \cdot f(x) = np$$

$$\sigma_X^2 = v(x) = E(X^2) - (E(X))^2 = npq$$

The m.g.f. of X :-

$$\begin{aligned} \mu_X(t) &= E e^{tx} = \sum_{x=0}^n e^{tx} C_x^n p^x q^{n-x} = \sum_{x=0}^n C_x^n (pe^t)^x q^{n-x} \\ &= (q + pe^t)^n, \quad -\infty < t < \infty \end{aligned}$$

* The additive probability in Binomial distⁿ.

Let $X_1 \sim b(n_1, p)$ independent of $X_2 \sim b(n_2, p)$

The $Y = X_1 + X_2 \sim b(n_1 + n_2, p)$

example(4):- let the r.v. $X \sim b(7, \frac{1}{2})$. write down the p.d.f. and find μ_X and σ_X^2 . Evaluate

$$\Pr(X \leq 2), \Pr(X > 2), \Pr(X = 5)$$

Sol. $n = 7, p = \frac{1}{2}$

$$f(x) = P(X=x) = \begin{cases} C_x^7 \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{7-x} & x=0,1,\dots,7 \\ 0 & \text{o.w.} \end{cases}$$

$$\mu_x = np = 7 \cdot \frac{1}{2} = \frac{7}{2}, \quad \sigma_x^2 = npq = 7 - \frac{1}{2} \cdot \frac{1}{2} = \frac{7}{4}$$

$$P(X \leq 2) = \sum_{x=0}^2 f(x) = f(0) + f(1) + f(2)$$

$$\Rightarrow C_0^7 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^7 + C_1^7 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^6 + C_2^7 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^5 = 0.2265.$$

$$P(X > 2) = 1 - P(X \leq 2) = 1 - 0.2265 = 0.77344.$$

$$P(X=5) = C_5^7 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^2 = 0.16466$$

H.w.: let the r.v. $X \sim b(7, p)$ such that

$$P(X=3) = P(X=4), \text{ Find } \mu_x \text{ and } \sigma_x^2.$$

Sol. of H.w.:

$$f(3) = f(4) \Rightarrow \cancel{C_3^7} \cancel{p^3} q^4 = \cancel{C_3^7} p^4 \cancel{q}$$

$$\Rightarrow q = p \Rightarrow 1-p = p \Rightarrow 1 = 2p \Rightarrow p = \frac{1}{2}$$

$$\mu_x = np = \frac{7}{2}, \quad \sigma_x^2 = npq = \frac{7}{4}$$

4 - poisson distribution

A r.v. X is said to have poisson distⁿ denoted by $X \sim P_0(\lambda)$, if it has p.d.f.

$$f(x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & x = 0, 1, 2, 3, \dots \\ 0 & \text{o.w.} \end{cases} \quad \text{where } \lambda > 0$$

Note: $e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$

To check that $f(x)$ is a p.d.f, we note that:

1. $f(x) > 0 \quad \forall x \in \{x = 0, 1, 2, \dots\}$

2. $\sum_x f(x) = 1 \Rightarrow \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \Rightarrow e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \frac{e^{-\lambda}}{e^{-\lambda}} = 1$

the m.g.f. of X :

$$\begin{aligned} \mu_x(t) &= E e^{tx} = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda} \lambda^x}{x!} \Rightarrow e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\ &= \frac{e^{-\lambda}}{e^{-\lambda}} = e^{\lambda(e^t - 1)} \end{aligned}$$

• Mean and variance of X :

$$\mu_x = E(X) = \sigma_x^2 = V(X) = \boxed{\lambda}$$

or $\psi_x(t) = \ln \mu_x(t) = \lambda(e^t - 1)$

$$\dot{\psi}_x(t) = \lambda e^t \Big|_{t=0} = \underline{\underline{\lambda}} \rightarrow E(X)$$

(8)

$$\dot{\psi}_x(t) = \lambda e^t \Big|_{t=0} = \underline{\underline{\lambda}} \rightarrow v(x)$$

H.w. verify the mean and variance of X by considering
 1. the differentiation of m.g.f. at $t=0$.
 2. the direct expectation approach.

example(5): let a r.v. $X \sim P_0(2)$. write down the p.d.f. and the m.g.f. of X , find μ_x and σ_x^2 .

Evaluate $P(X \leq 1)$, $P(X \geq 1)$, $P(2 \leq X \leq 5)$

Sol. $\lambda = 2$

$$f(x) = \begin{cases} \frac{e^{-2} 2^x}{x!} & x=0,1,2,\dots \\ 0 & o.w. \end{cases}, \quad \mu_x(t) = e^{2(e^t - 1)}$$

$$\mu_x = \sigma_x^2 = 2$$

$$\begin{aligned} P(X \leq 1) &= \sum_{x=0}^1 f(x) = f(0) + f(1) \\ &= \frac{e^{-2} 2^0}{0!} + \frac{e^{-2} 2^1}{1!} = 0.406006 \end{aligned}$$

$$\begin{aligned} P(X \geq 1) &= 1 - P(X < 1) = 1 - P(X=0) \\ &= 1 - \frac{e^{-2} 2^0}{0!} = 0.864665 \end{aligned}$$

H.w.
 $P(2 \leq X \leq 5)$

Example (6): let a r.v. X has m.g.f. $\mu_X(t) = e^{4(e^t - 1)}$

what is the distⁿ of X , Find μ_X and σ_X^2 . Evaluate

$$P(2 \leq X \leq 7)$$

Sol:

$\mu_X(t) = e^{4(e^t - 1)}$ is for a r.v. $X \sim \text{Po}(4)$

$$\mu_X = \sigma_X^2 = 4$$

$$\begin{aligned} P(2 \leq X \leq 7) &= \sum_{x=2}^7 f(x) = f(2) + f(3) + f(4) + f(5) + f(6) + f(7) \\ &= \frac{e^{-4} 4^2}{2!} + \frac{e^{-4} 4^3}{3!} + \frac{e^{-4} 4^4}{4!} + \frac{e^{-4} 4^5}{5!} + \frac{e^{-4} 4^6}{6!} + \frac{e^{-4} 4^7}{7!} = ?? \end{aligned}$$

5. Geometric distⁿ.

the r.v. X has a geometric distⁿ $X \sim \text{Geo}(p)$, iff

$$f(x) = \begin{cases} p q^x & x=0, 1, 2, \dots \\ 0 & \text{o.w.} \end{cases} \quad \text{or} \quad f(x) = \begin{cases} p q^{x-1} & x=1, 2, \dots \\ 0 & \text{o.w.} \end{cases}$$

To verify that $f(x)$ is a p.d.f, we note that.

$$1 - f(x) > 0, \quad \forall x \in \{x=0, \dots\}$$

$$\sum_x f(x) = 1 \Rightarrow \sum_{x=0}^{\infty} p q^x = p \cdot [1 + q + q^2 + \dots]$$

$$= p \cdot \frac{1}{1-q} = \frac{p}{p} = 1$$

(10)

The c.d.f. of X :-

$$F(x) = P(X \leq x) = \sum_{t=0}^x P(t) = p \sum_{t=0}^x q^t$$

$$\Rightarrow P[1+q+q^2+\dots+q^x] = \cancel{p} \cdot \frac{1-q^{x+1}}{\cancel{1-q}} = \begin{cases} 0 & x < 0 \\ 1-q^{x+1} & 0 \leq x < \infty \\ 1 & x = \infty \end{cases}$$

the m.g.f. of X :-

$$\begin{aligned} M_x(t) &= E e^{tx} = \sum_{x=0}^{\infty} e^{tx} f(x) = p \sum_{x=0}^{\infty} (qe^t)^x \\ &= p \cdot \frac{1}{1-qe^t} \end{aligned}$$

• Mean and variance of X :-

$$M_x = E(X) = \frac{q}{p}, \quad \sigma_x^2 = V(X) = \frac{q}{p^2}$$

$$\text{or } \psi_x(t) = \ln M_x(t) = \ln p - \ln(1-qe^t)$$

$$\psi'_x(t) = \left[\frac{1}{1-qe^t} - qe^t \right] \bigg|_{t=0} = \frac{q}{1-q} = \left(\frac{q}{p} \right)$$

$$\psi''_x(t) = \frac{(1-qe^t) \cdot qe^t + (qe^t) \cdot (qe^t)}{(1-qe^t)^2} \bigg|_{t=0} = \frac{(1-q)q + q^2}{(1-q)^2}$$

$$= \frac{q - \cancel{q^2} + \cancel{q^2}}{(1-q)^2} = \left(\frac{q}{p^2} \right)$$

(11)

Example (7):- let the r.v. $X \sim \text{Geo}(\frac{2}{3})$, write down the p.d.f, c.d.f., m.g.f. of X , Evaluate $P(X \geq 3)$ and $P(X \geq 5 | X \geq 2)$

Sol. $p = \frac{2}{3}$, $q = \frac{1}{3}$

$$f(x) = \begin{cases} (\frac{2}{3})(\frac{1}{3})^x & x=0,1,2,\dots \\ 0 & \text{o.w.} \end{cases}, F(x) = \begin{cases} 0 & x < 0 \\ 1 - (\frac{1}{3})^{x+1} & 0 \leq x < \infty \\ 1 & x = \infty \end{cases}$$

$$\mu_x(t) = \frac{\frac{2}{3}}{1 - (\frac{1}{3})e^t} = \frac{2}{3} \cdot \frac{1}{3 - e^t} = \frac{2}{3 - e^t}$$

$$\begin{aligned} P(X \geq 3) &= 1 - P(X < 3) = 1 - [f(0) + f(1) + f(2)] \\ &= 1 - \frac{2}{3} \left(1 + \frac{1}{3} + (\frac{1}{3})^2 \right) = \frac{1}{27} \end{aligned}$$

$$\begin{aligned} P(X \geq 5 | X \geq 2) &= \frac{P(X \geq 5 \cap X \geq 2)}{P(X \geq 2)} \\ &= \frac{P(X \geq 5)}{P(X \geq 2)} = \frac{1 - P(X \leq 4)}{1 - P(X \leq 1)} = \frac{(\frac{1}{3})^5}{(\frac{1}{3})^2} \\ &= \frac{1}{27} \end{aligned}$$

Exercises for Discrete distributions

- ①. if $X \sim \text{unif}(N)$, find the value N such that
$$P(X \leq \mu_X) = 0.56$$
- ②. let $X \sim b(n, p)$, if $\mu_X = 5$, $\sigma_X^2 = 4$ find
 n, p and Evaluate $P(X > \mu_X)$
- ③. if $X \sim b(n, p)$, find $E\left(\frac{X - np}{\sqrt{npq}}\right)$
- ④. let $\mu_X(t) = (0.3 + 0.7e^t)^6$ be the m.g.f. of r.v. X .
what is the distⁿ. of X , Find μ_X and σ_X^2
- ⑤. Let a r.v. $X \sim b(2, p)$ & $Y \sim b(4, p)$, if
 $P(X \geq 1) = \frac{5}{9}$. Find $P(Y \geq 1)$
- ⑥. if $X \sim \text{Po}(\lambda)$, such that $P(X=0) = P(X=1)$
find μ_X , $\mu_X(t)$, σ_X^2 , $P(X > 2)$
- ⑦. let $X \sim \text{Geo}\left(\frac{2}{3}\right)$, find $P(X \leq 1 | X \leq 2)$

* continuous Types:-

1- Uniform Distribution:-

A continuous r.v. X is said to have a uniform distⁿ, denoted $X \sim U(a, b)$, if X has p.d.f:-

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{o.w.} \end{cases} \quad \begin{matrix} -\infty < a, b < \infty \\ b > a \end{matrix} \quad (a, b) \text{ are parameters}$$

To check that $f(x)$ is a p.d.f., we note that:-

1. $f(x) > 0$, $\forall x \in (a, b)$

2. $\int_x f(x) dx = 1 \Rightarrow \int_a^b \frac{1}{b-a} dx = \frac{1}{b-a} x \Big|_a^b = \frac{b-a}{b-a} = 1.$

The c.d.f. of X :- x

$$F(x) = P(X \leq x) = \int_{t=a}^x f(t) dt = \frac{1}{b-a} t \Big|_a^x = \frac{x-a}{b-a}$$

$$\therefore F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

The m.g.f. of X :-

$$\begin{aligned} \mu_x(t) &= E e^{tx} = \int_a^b e^{tx} \cdot \frac{1}{b-a} = \frac{1}{b-a} \cdot \frac{1}{t} e^{tx} \Big|_a^b \\ &= \frac{(e^{tb} - e^{ta})}{t(b-a)} \end{aligned}$$

* mean and variance of X :-

$$\mu_X = E(X) = \int_a^b x \cdot f(x) dx = \frac{1}{(b-a)} \cdot \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)}$$

$$= \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}$$

$$\sigma_X^2 = V(X) = E(X^2) - (E(X))^2$$

$$E(X^2) = \int_a^b x^2 \cdot f(x) dx = \frac{1}{(b-a)} \cdot \frac{x^3}{3} \Big|_a^b = \frac{b^3 - a^3}{3(b-a)}$$

$$= \frac{(b-a)(a^2 + b^2 + ab)}{3(b-a)} = \frac{a^2 + b^2 + ab}{3}$$

$$\therefore \sigma_X^2 = \frac{a^2 + b^2 + ab}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

H.W.:- Prove that $E(X)^m = \frac{b^{m+1} - a^{m+1}}{(m+1)(b-a)}$

Note:- if $a=0$ & $b=1$, Then the r.v. X is uniform distⁿ over interval $(0,1)$, That is $X \sim U(0,1)$ with p.d.f.

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{o.w.} \end{cases}$$

$$\mu_X = \frac{1}{2}, \quad \sigma_X^2 = \frac{1}{12}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

$$\mu_X(t) = \frac{e^t - 1}{t}$$

Example (2): Let a r.v. $X \sim U(2, 5)$, write down the p.d.f., c.d.f and M.g.f. of X , Find μ_X and σ_X^2 , Evaluate $\Pr(3 \leq X \leq 4)$

Sol:
$$f(x) = \begin{cases} \frac{1}{3} & 2 \leq x \leq 5 \\ 0 & \text{o.w.} \end{cases}, F(x) = \begin{cases} 0 & x < 2 \\ \frac{x-2}{3} & 2 \leq x \leq 5 \\ 1 & x > 5 \end{cases}$$

$$\mu_X(t) = \frac{e^{5t} - e^{2t}}{3t}, \mu_X = E(X) = \frac{7}{2}, \sigma_X^2 = \frac{3}{4}$$

$$\Pr(3 \leq X \leq 4) = \int_3^4 f(x) dx = \frac{1}{3} x \Big|_3^4 = \frac{1}{3} (4-3) = \frac{1}{3}$$

or
$$\Pr(3 \leq X \leq 4) = F(4) - F(3) = \frac{4-2}{3} - \frac{3-2}{3} = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$$

example (9): Let a r.v. $X \sim U(a, b)$ with $\mu_X = 1$ and $\sigma_X^2 = \frac{4}{3}$, find $\Pr(X \leq 0)$

Sol:
$$\mu_X = \frac{a+b}{2} = 1 \Rightarrow a+b = 2 \dots \textcircled{1}$$

$$\sigma_X^2 = \frac{(b-a)^2}{12} = \frac{4}{3} \Rightarrow (b-a)^2 = 16 \Rightarrow b-a = \pm 4 \dots \textcircled{2}$$

consider
$$\begin{array}{r} b+a=2 \\ b-a=-4 \\ \hline 2b=-2 \end{array} \Rightarrow b=-1 \text{ substitution in eq (1)}$$

$$a - 1 = 2 \Rightarrow a = 3 \text{ impossible because } b > a$$

$$\text{So } b + \cancel{a} = 2$$

$$\frac{b - \cancel{a} = 4}{2b = 6} \Rightarrow b = 3 \text{ substitution in eq(1)}$$

$$a + 3 = 2 \Rightarrow a = -1$$

$$\therefore X \sim U(-1, 3)$$

$$f(x) = \begin{cases} \frac{1}{4} & -1 < x < 3 \\ 0 & \text{o.w.} \end{cases}$$

$$P(X < 0) = \int_{-1}^0 f(x) = \frac{1}{4} x \Big|_{-1}^0 \Rightarrow \frac{1}{4}(0 + 1) = \frac{1}{4}$$

2. Exponential Distribution:-

A r.v. X is said to have an exponential dist.ⁿ denoted by $X \sim \exp(\lambda)$, if X has p.d.f.:-

$$f(x) = P(X=x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & 0 < x < \infty, \lambda > 0 \\ 0 & \text{o.w.} \end{cases}$$

To check that $f(x)$ is valid p.d.f., we note that:-

$$1. f(x) > 0, \forall x \in (0, \infty)$$

$$2. \int_x f(x) dx = 1 \Rightarrow \int_0^{\infty} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx \Rightarrow -e^{-\frac{x}{\lambda}} \Big|_0^{\infty} \\ (17) \Rightarrow -[e^{-\infty} - e^0] = 1$$

The c.d.f. of X :-

$$F(x) = P(X \leq x) = \int_0^x f(t) dt = \int_0^x \frac{1}{\lambda} e^{-\frac{t}{\lambda}} dt$$
$$\Rightarrow -e^{-\frac{t}{\lambda}} \Big|_0^x = - \left[e^{-\frac{x}{\lambda}} - e^0 \right] = \begin{cases} 1 - e^{-\frac{x}{\lambda}} & x \leq \infty \\ 1 & 0 \leq x < \infty \\ & \lambda = \infty \end{cases}$$

The m.g.f. of X :-

$$M_x(t) = E e^{tx} = \int_0^{\infty} e^{tx} \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx \Rightarrow \frac{1}{\lambda} \int_0^{\infty} e^{-\frac{x}{\lambda}(1-\lambda t)} dx$$
$$= \frac{1}{\cancel{\lambda}} \cdot \left(-\frac{\cancel{\lambda}}{1-\lambda t} \cdot e^{-\frac{x}{\lambda}(1-\lambda t)} \right) \Big|_0^{\infty} = \frac{1}{1-\lambda t} (e^0 - e^{\infty})$$
$$= \frac{1}{1-\lambda t}, \quad t < \frac{1}{\lambda}$$

* Mean and variance of X :-

$$\psi_x^{(1)}(t) = \ln M_x(t) = -\ln(1-\lambda t)$$

$$\psi_x^{(1)}(t) = \frac{\lambda}{1-\lambda t} \Big|_{t=0} = \lambda = M_x' = E(x)$$

$$\psi_x^{(2)}(t) = \frac{\lambda^2}{(1-\lambda t)^2} \Big|_{t=0} \Rightarrow \lambda^2 = \sigma_x^2 = v(x)$$

H.w.: let $X \sim \text{exp}(\lambda)$, using a direct expectation to find μ_X, σ_X^2 .

Note: when $\lambda=1$, then a r.v $X \sim \text{exp}(1)$, we say that X has Standard Exponential distⁿ with p.d.f.:

$$f(x) = \begin{cases} e^{-x} & 0 < x < \infty \\ 0 & \text{o.w.} \end{cases} \quad \therefore \mu_X = \sigma_X^2 = 1$$

example (10): let the r.v. $X \sim \text{exp}(\lambda)$. if $P(X \leq 1) = P(X > 1)$
Find μ_X and σ_X^2 .

Sol.

$$P(X \leq 1) = P(X > 1)$$

$$P(X \leq 1) = 1 - P(X \leq 1) \Rightarrow 2P(X \leq 1) = 1$$

$$\Rightarrow P(X \leq 1) = \frac{1}{2}$$

$$\Rightarrow F(1) = \frac{1}{2} \Rightarrow 1 - e^{-\frac{1}{\lambda}} = \frac{1}{2} \Rightarrow e^{-\frac{1}{\lambda}} = 1 - \frac{1}{2}$$

$$\Rightarrow e^{-\frac{1}{\lambda}} = \frac{1}{2} \Rightarrow -\frac{1}{\lambda} = \ln\left(\frac{1}{2}\right)$$

$$\Rightarrow \lambda = \frac{1}{\ln(2)} \quad \therefore \mu_X = \frac{1}{\ln(2)}, \quad \sigma_X^2 = \left(\frac{1}{\ln(2)}\right)^2$$

3. Normal distribution (Gaussian distⁿ)

let r.v. X has a Normal distⁿ, denoted by $X \sim N(\mu, \sigma^2)$ with parameters μ, σ^2 and a p.d.f. has following form:-

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} & -\infty < x, \mu < \infty \\ & \sigma^2 > 0 \\ 0 & \text{o.w.} \end{cases}$$

the parameter μ represents a location parameter & σ^2 represents a shape parameter.

The m.g.f. of X :-

$$M_X(t) = E e^{tX} = \int_{-\infty}^{\infty} e^{tX} \cdot f(x) = e^{\mu t + \frac{1}{2} t^2 \sigma^2}$$

* Mean & variance of X :-

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

$$\psi_X'(t) = \mu + t \sigma^2 \Big|_{t=0} = \mu = E(X)$$

$$\psi_X''(t) = \sigma^2 \Big|_{t=0} = \sigma^2 = v(X)$$

Note:- if $\mu=0$, $\sigma^2=1$, we say that a r.v. X has standard normal distribution that is

$X \sim N(0, 1)$, where the p.d.f & m.g.f. of X in

This case are:

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} & -\infty < x < \infty \\ 0 & \text{o.w.} \end{cases}$$

$$M_x(t) = e^{\frac{1}{2}t^2}$$

* Relation Between $N(\mu, \sigma^2)$ and $N(0, 1)$

Theorem (1):- If a r.v. $X \sim N(\mu, \sigma^2)$, then a r.v.

$$Y = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

Proof:- c.d.f. Technique

let $G(Y) = P(Y \leq y)$ be the c.d.f. of r.v. Y .

$$G(Y) = P\left(\frac{X - \mu}{\sigma} \leq y\right) = P(X \leq \sigma y + \mu)$$

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

$$\text{Let } t = \frac{x - \mu}{\sigma} \Rightarrow x = \sigma t + \mu \Rightarrow dx = \sigma dt$$

$$\therefore G(Y) = \int_{-\infty}^y \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \quad (21)$$

and the p.d.f. of Y is $g(y) = G(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$

$$\therefore Y \sim N(0, 1)$$

* using m.g.f. technique

$$M_Y(t) = E e^{tY} = \int_{-\infty}^{\infty} e^{t \left(\frac{X-\mu}{\sigma} \right)} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2} (X-\mu)^2} dX$$

$$\Rightarrow e^{-\frac{t\mu}{\sigma}} \int_{-\infty}^{\infty} e^{\frac{tx}{\sigma}} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dX$$

represent m.g.f. of $\left(\frac{X}{\sigma} \right)$

$$\therefore e^{-\frac{t\mu}{\sigma}} \cdot e^{\frac{t\mu}{\sigma} + \frac{1}{2} t^2 \cancel{\sigma^2}}$$

$$\Rightarrow \cancel{e^{-\frac{t\mu}{\sigma}}} + \cancel{\frac{t\mu}{\sigma}} + \frac{1}{2} t^2 = e^{\frac{1}{2} t^2} \text{ is m.g.f. of } Y \sim N(0, 1)$$

Notes:

Let a r.v. $X \sim N(\mu, \sigma^2)$ and suppose we wish to compute $P(a < X < b)$, By theorem(1): If $X \sim N(\mu, \sigma^2)$, then $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right)$$

$$= P\left(\frac{a-\mu}{\sigma} < Y < \frac{b-\mu}{\sigma}\right)$$

$$= P\left(Y < \frac{b-\mu}{\sigma}\right) - P\left(Y < \frac{a-\mu}{\sigma}\right)$$

$$\begin{aligned}
 P(-a < Y < a) &= P(Y < a) - P(Y < -a) \\
 &= P(Y < a) - P(Y \geq a) \\
 &= P(Y < a) - (1 - P(Y < a)) \\
 &= 2 P(Y < a) - 1.
 \end{aligned}$$

Example (11): Let a r.v. $X \sim N(2, 25)$. Find $P(0 < X < 10)$.

Sol.

$$\because X \sim N(\mu, \sigma^2) \quad , \mu = 2 \quad , \sigma^2 = 25$$

$$\because X \sim N(2, 25) \quad , Y = \frac{X - \mu}{\sigma} = \frac{X - 2}{5} \sim N(0, 1)$$

$$\begin{aligned}
 P(0 < X < 10) &= P\left(\frac{0-2}{5} < Y < \frac{10-2}{5}\right) \\
 &= P\left(-\frac{2}{5} < Y < \frac{8}{5}\right) = P(-0.4 < Y < 1.6)
 \end{aligned}$$

$$\Rightarrow P(Y < 1.6) - P(Y < -0.4)$$

$$\Rightarrow P(Y < 1.6) - 1 + P(Y < 0.4)$$

من جداول التوزيع الطبيعي القياسي
من جداول التوزيع الطبيعي القياسي

$$\Rightarrow 0.945 - 1 + 0.655 = \underline{\underline{0.6}}$$

41. Gamma distribution

The r.v. X has a Gamma distⁿ with parameters (α, β) , denoted by $X \sim G(\alpha, \beta)$ iff:-

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\frac{x}{\beta}} & 0 < x < \infty \\ 0 & \text{o.w.} \end{cases} \quad \alpha, \beta > 0$$

$$\text{or} \quad f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & \text{o.w.} \end{cases} \quad \alpha, \beta > 0$$

Remark: $\Gamma(\alpha)$ is called a Gamma function which is defined as:-

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

$$\text{if } \alpha = 1, \text{ then } \Gamma(1) = \int_0^{\infty} e^{-x} dx = 1$$

if $\alpha > 1$, then

$$\Gamma(\alpha) = (\alpha-1)(\alpha-2) \Gamma(\alpha-2) = (\alpha-1)!$$

Note: It has been shown that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

To check that $f(x)$ is a p.d.f, we note that:-

1. $f(x) > 0$, $\forall x \in (0, \infty)$

2. $\int_x f(x) dx = 1$ (H.w. prove that)

Note: The Gamma distⁿ. reduce to exponential distⁿ.
when $\alpha = 1$

The m.g.f. of X :-

$$M_X(t) = E e^{tX} = \int_0^{\infty} e^{tx} \cdot f(x) dx = \frac{1}{(1-\beta t)^\alpha}, (t < \frac{1}{\beta})$$

* Mean & variance of X

$$\psi_X(t) = \ln M_X(t) = -\alpha \ln(1-\beta t)$$

$$\psi_X'(t) = \frac{\alpha \beta}{1-\beta t} \Big|_{t=0} = \alpha \beta = E(X) = \mu$$

$$\psi_X''(t) = \frac{\alpha \beta^2}{(1-\beta t)^2} \Big|_{t=0} = \alpha \beta^2 = v(X) = \sigma_X^2$$

H.w.: let $X \sim G(\alpha, \beta)$. Find μ_X, σ_X^2 by using direct expectation technique.

example(12): let the r.v. $X \sim G(3, 2)$. write down the p.d.f. and m.g.f. of X . Find μ, σ^2 . Evaluate $P(12.6 < X < 16.8)$

Sol.

$$P(x) = \begin{cases} \frac{1}{\sqrt[3]{3} \cdot 2^3} x^{3-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & \text{o.w.} \end{cases}$$

$$\mu_x(t) = (1 - 2t)^{-3}, \quad \mu_x = \alpha \beta = 6$$

$$\sigma_x^2 = \alpha \beta^2 = 12$$

$$P(12.6 < X < 16.8) = \int_{12.6}^{16.8} P(x) dx$$

$$= \frac{1}{16} \int_{12.6}^{16.8} x^2 e^{-\frac{x}{2}} dx = \underline{\underline{0.04}}$$

example(13): let a r.v. X has m.g.f. $\mu_x(t) = (1 - 3t)^{-5}$, what is the distⁿ. of X . Find μ_x & σ_x^2

Sol: the given m.g.f. for r.v. $X \sim G(5, 3)$, where $\alpha = 5, \beta = 3$

$$\mu_x = \alpha\beta = 15, \quad \sigma_x^2 = \alpha\beta^2 = 45$$

example (14): find the constant k so that given function is a p.d.f. $f(x) = \begin{cases} k x^5 e^{-3x} & x > 0 \\ 0 & \text{o.w.} \end{cases}$

Sol: $\int_0^{\infty} f(x) dx = 1 \Rightarrow k \int_0^{\infty} x^5 e^{-3x} dx = 1$

$$\Rightarrow k \cdot \frac{\Gamma_6}{3^6} \underbrace{\int_0^{\infty} \frac{3^6}{\Gamma_6} x^{6-1} e^{-3x} dx}_1 = 1$$

$$\therefore k \cdot \frac{\Gamma_6}{3^6} = 1 \Rightarrow k = \boxed{\frac{3^6}{5!}}$$

H.w: let $X \sim G(2, 2)$, find $\Pr(X, 5)$

5. Beta distribution

the r.v. X has a Beta distribution. $X \sim \text{Be}(\alpha, \beta)$ with parameter α, β , i.e.,

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{o.w.} \end{cases}$$

$\alpha, \beta > 0$

Note: when $\alpha = \beta = 1$, the Beta distⁿ reduce to uniform distⁿ. that is $Be(1,1) \equiv U(0,1)$

to check that $f(x)$ is a p.d.f., we note that:

1. $f(x) > 0 \quad \forall x \in (0,1)$

2. $\int_x f(x) dx = 1 \Rightarrow$ (H.w.) prove that,

mean and variance of X :-

$$\mu_x = E(X) = \int_{x=0}^1 x \cdot f(x) dx$$

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

$$\sigma_x^2 = v(x) = E(X^2) - (E(X))^2$$

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

$$\therefore \sigma_x^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad (28)$$

H.w: let $X \sim \text{Be}(\alpha, \beta)$, Find $E(X^r)$

example (15): let a r.v. $X \sim \text{Be}(3, 5)$, write down the p.d.f. of X . find μ_X and σ_X^2 , and evaluate $P(0.2 < X < 0.4)$

Sol.

$$f(x) = \begin{cases} \frac{\Gamma(3) \Gamma(5)}{\Gamma(3+5)} x^2 (1-x)^4 & 0 < x < 1 \\ 0 & \text{o.w.} \end{cases}$$

$$\mu_X = \frac{\alpha}{\alpha + \beta} = \frac{3}{8}, \quad \sigma_X^2 = \frac{\alpha \beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} = \frac{5}{192}$$

$$P(0.2 < X < 0.4) = \int_{0.2}^{0.4} f(x) dx = 105 \int_{0.2}^{0.4} x^2 (1-x)^4 dx$$

example (16): find the constant c in the following p.d.f. $f(x) = \begin{cases} c x^{19} (1-x)^{29} & 0 < x < 1 \\ 0 & \text{o.w.} \end{cases}$

Sol.

$$\int_x f(x) dx = 1 \Rightarrow c \int_0^1 x^{19} (1-x)^{29} dx = 1$$

$$\Rightarrow c \cdot \frac{\Gamma(20) \Gamma(30)}{\Gamma(50)} \int_0^1 \frac{\Gamma(50)}{\Gamma(20) \Gamma(30)} x^{19} (1-x)^{29} dx = 1$$

(29)

$$\Rightarrow C = \frac{\sqrt{s_0}}{\sqrt{s_0} \sqrt{s_0}}$$

6. Chi-Square distribution (χ^2)

consider a special case of gamma distⁿ. in which
 $\alpha = \frac{r}{2}$, $\beta = 2$, where r is integer.

a r.v. $X \sim \chi^2(r)$, i.f.f.:-

$$f(x) = \begin{cases} \frac{1}{\Gamma(\frac{r}{2}) 2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & \text{o.w.} \end{cases}$$

r is degrees of freedom

Notes:-

1. $\chi^2(r) \equiv G(\frac{r}{2}, 2)$

2. This distⁿ. plays extremely important role in both theoretical and applied statistics.

The m.g.f. of X :-

$$M_X(t) = E e^{tX} = (1-2t)^{-\frac{r}{2}}, \quad t < \frac{1}{2}$$

Mean and variance of X .

Since $X^2(r) \equiv G(\frac{r}{2}, 2)$

$$\mu_X = E(X) = \left(\frac{r}{2}\right)(2) = \boxed{r}$$

$$\sigma_X^2 = V(X) = \left(\frac{r}{2}\right)(4) = \boxed{2r}$$

H.w.: let $X \sim X^2(r)$. Find μ_X, σ_X^2 by using

1. m.g.f. of X .

2. direct expectation technique.

example (17): let $X \sim X^2(4)$, write down the p.d.f., m.g.f. of X , find μ_X, σ_X^2 . Evaluate $P(0.4845 \leq X \leq 11.1)$

Sol. $r=4$

$$f(x) = \begin{cases} \frac{1}{4} x e^{-\frac{x}{2}} & x > 0 \\ 0 & \text{o.w.} \end{cases}, \quad \mu_X(t) = (1-2t)^{-2}$$

$$\mu_X = 4, \quad \sigma_X^2 = 8$$

$$\begin{aligned} P(0.4845 \leq X \leq 11.1) &= \underbrace{P(X \leq 11.1)}_{\text{میں جدول توزیع مربع کی}} - \underbrace{P(X \leq 0.4845)}_{\text{میں جدول توزیع مربع کی}} \\ &= 0.975 - 0.025 = 0.95 \end{aligned}$$

Example (18): Find μ and σ^2 of r.v. X whose p.d.f. is

$$f(x) = \begin{cases} \frac{1}{16} x^2 e^{-\frac{x}{2}} & 0 < x < \infty \\ 0 & \text{o.w.} \end{cases}, \text{ Evaluate } P(12.6 < X < 14.4)$$

Sol. the given p.d.f. for r.v. $X \sim \chi^2(6)$, so $r = 6$
then $\mu = r = 6$, $\sigma^2 = 2r = 12$

$$\begin{aligned} P(12.6 < X < 14.4) &= \underbrace{P(X < 14.4)}_{\chi^2 \text{ جدول}} - \underbrace{P(X < 12.6)}_{\chi^2 \text{ جدول}} \\ &= 0.975 - 0.95 = 0.025 \end{aligned}$$

H.w.: let a r.v. X has m.g.f. $M_X(t) = (1-2t)^{-8}$
what is the dist. of X . Find μ_X , σ_X^2 and
Evaluate $Pr(7.96 < X < 26.3)$

Exercices for continuous distributions

- ① let a r.v. $X \sim U(1, 2)$. Find the constant c which that $P(X > c + \mu_x) = \frac{1}{3}$
- ② if $X \sim G(\alpha, \frac{1}{\beta})$, find $E(\frac{1}{X})$
- ③ if a r.v. $X \sim \text{Be}(\alpha^2, \beta^T)$. Find $E(1-X)^2$
- ④ find the constant R in the following p.d.f's
 - a. $f(x) = \begin{cases} R x^4 e^{-5x} & x > 0 \\ 0 & \text{o.v.} \end{cases}$
 - b. $f(x) = \begin{cases} R x^{21} (1-x)^{51} & 0 < x < 1 \\ 0 & \text{o.w.} \end{cases}$
- ⑤ if $X \sim \exp(\lambda)$, find the parameter λ when the c.d.f. of X equal to $(\frac{1}{2})$.
- ⑥ consider you have the following p.d.f.
$$f(x) = \begin{cases} c e^{-\frac{1}{2}(x^2 - 8x)} & -\infty < x < \infty \\ 0 & \text{o.w.} \end{cases}$$
Find (c, μ_x, σ_x^2) (33)

Section Two

"Marginal and Conditional Distribution" And Stochastic Independent

Notation

For two r.v.'s X_1 and X_2 , we shall call

$f(x_1, x_2)$ a joint p.d.f of X_1 and X_2

$F(x_1, x_2) = \Pr(X_1 \leq x_1, X_2 \leq x_2)$ a joint c.d.f of X_1 and X_2

$M(t_1, t_2) = E(e^{t_1 X_1 + t_2 X_2})$ a joint m.g.f of X_1 and X_2

In general, when we have n r.v.'s $X_1, X_2, \dots, X_n$, we shall call

$f(x_1, x_2, \dots, x_n)$ a joint p.d.f of $X_1, X_2, \dots, X_n$

$F(x_1, x_2, \dots, x_n) = \Pr(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$ a joint c.d.f of $X_1, X_2, \dots, X_n$

$M(t_1, t_2, \dots, t_n) = E(e^{t_1 X_1 + t_2 X_2 + \dots + t_n X_n})$ a joint m.g.f of $X_1, X_2, \dots, X_n$.

Marginal Distributions

"In 2-D case"

Defn. Let $f(x_1, x_2)$ be the joint p.d.f of r.v.'s X_1 and X_2 . We define the marginal p.d.f of X_1 and the marginal p.d.f of X_2 as

$$f_1(x_1) = \begin{cases} \sum_{x_2} f(x_1, x_2) & , \text{ for disc. case} \\ \int_{x_2} f(x_1, x_2) dx_2 & , \text{ for cont. case} \end{cases}$$

$$f_2(x_2) = \begin{cases} \sum_{x_1} f(x_1, x_2) & , \text{ disc} \\ \int_{x_1} f(x_1, x_2) dx_1 & , \text{ cont.} \end{cases}$$

Note

suppose, we require to evaluate

$$1. \Pr(a \leq X_1 \leq b) = \begin{cases} \sum_{x_1=a}^b f_1(x_1) = \sum_{x_1=a}^b \sum_{x_2} f(x_1, x_2) & , \text{ disc} \\ \int_a^b f_1(x_1) dx_1 = \int_a^b \int_{x_2} f(x_1, x_2) dx_2 dx_1 & , \text{ cont.} \end{cases}$$

$$2. P(c \leq X_2 \leq d) = \begin{cases} \sum_{x_2=c}^d f_2(x_2) = \sum_{x_2=c}^d \sum_{x_1} f(x_1, x_2) & \text{disc} \\ \int_c^d f_2(x_2) dx_2 = \int_c^d \int_{x_1} f(x_1, x_2) dx_1 dx_2, & \text{cont.} \end{cases}$$

$$3. E[u(X_1)] = \begin{cases} \sum_{x_1} u(x_1) f_1(x_1) = \sum_{x_1} \sum_{x_2} u(x_1) f(x_1, x_2) & \text{disc} \\ \int_{x_1} u(x_1) f_1(x_1) dx_1 = \int_{x_1} \int_{x_2} u(x_1) f(x_1, x_2) dx_2 dx_1, & \text{cont.} \end{cases}$$

$$4. E[u(X_2)] = \begin{cases} \sum_{x_2} u(x_2) f_2(x_2) = \sum_{x_2} \sum_{x_1} u(x_2) f(x_1, x_2) & \text{disc} \\ \int_{x_2} u(x_2) f_2(x_2) dx_2 = \int_{x_2} \int_{x_1} u(x_2) f(x_1, x_2) dx_1 dx_2, & \text{cont.} \end{cases}$$

Example (1) Let the joint p.d.f of r.v.s X and Y be

$$(i) f(x, y) = \begin{cases} \frac{1}{21}(x+y), & x=1, 2, 3; y=1, 2 \\ 0 & \text{e.w} \end{cases}$$

$$(ii) f(x, y) = \begin{cases} e^{-(x+y)}, & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{e.w} \end{cases}$$

(a) Find the marginal p.d.f of X and the marginal p.d.f of Y

(b) Find $E(X)$, $\text{Var}(X)$, $E(Y)$, $\text{Var}(Y)$, $E(XY)$

Solution (i)

(a) The marginal p.d.f of X is $f_1(x) = \sum_y f(x, y)$

$$\Rightarrow f_1(x) = \sum_{y=1}^2 \frac{1}{21}(x+y) = \frac{1}{21}[(x+1) + (x+2)] = \frac{1}{21}(2x+3), x=1, 2, 3$$

= 0 e.w

The marginal p.d.f of Y is $f_2(y) = \sum_x f(x, y)$

$$\Rightarrow f_2(y) = \sum_{x=1}^3 \frac{1}{21}(x+y) = \frac{1}{21}[(1+y) + (2+y) + (3+y)] = \frac{1}{7}(y+2), y=1, 2$$

= 0 e.w

(3)

$$\mu_1 = E(X) = \sum_x x f_1(x) = \frac{1}{21} \sum_{x=1}^3 x(2x+3) = \frac{46}{21}$$

$$E(X^2) = \sum_x x^2 f_1(x) = \frac{1}{21} \sum_{x=1}^3 x^2(2x+3) = \frac{114}{21}$$

$$\text{Var}(X) = \sigma_1^2 = E(X^2) - \mu_1^2 = \frac{114}{21} - \left(\frac{46}{21}\right)^2 = \frac{278}{441}$$

$$\mu_2 = E(Y) = \sum_y y f_2(y) = \frac{1}{7} \sum_{y=1}^2 y(y+2) = \frac{11}{7}$$

$$E(Y^2) = \sum_y y^2 f_2(y) = \frac{1}{7} \sum_{y=1}^2 y^2(y+2) = \frac{19}{7}$$

$$\text{Var}(Y) = \sigma_2^2 = E(Y^2) - \mu_2^2 = \frac{19}{7} - \left(\frac{11}{7}\right)^2 = \frac{12}{49}$$

$$\begin{aligned} E(XY) &= \sum_x \sum_y xy f(x,y) = \frac{1}{21} \sum_{x=1}^3 \sum_{y=1}^2 xy(x+y) = \frac{1}{21} \sum_{x=1}^3 [x(x+1) + 2x(x+2)] \\ &= \frac{1}{21} \sum_{x=1}^3 (3x^2 + 5x) = \frac{1}{21} (8 + 22 + 42) = \frac{24}{7} \end{aligned}$$

Solution (ii)

$$\begin{aligned} (a) \quad f_1(x) &= \int_y f(x,y) dy = \int_0^\infty e^{-(x+y)} dy = e^{-x} \int_0^\infty e^{-y} dy = -e^{-x} (e^{-y}) \Big|_0^\infty \\ &= -e^{-x} (0 - 1) = e^{-x}, \quad 0 < x < \infty \\ &= 0, \quad \text{e.w.} \end{aligned}$$

Similarly

$$\begin{aligned} f_2(y) &= e^{-y}, \quad 0 < y < \infty \\ &= 0, \quad \text{e.w.} \end{aligned}$$

$$\mu_1 = E(X) = \int_x x f_1(x) dx = \int_0^\infty x e^{-x} dx = 1 = E(Y) = \mu_2$$

$$E(X^2) = \int_x x^2 f_1(x) dx = \int_0^\infty x^2 e^{-x} dx = 2 = E(Y^2)$$

$$\text{Var}(X) = \sigma_1^2 = E(X^2) - \mu_1^2 = 2 - (1)^2 = 1 = \text{Var}(Y) = \sigma_2^2$$

$$E(XY) = \int_x \int_y xy f(x,y) dx dy = \int_0^\infty \int_0^\infty xy e^{-(x+y)} dx dy$$

$$= \left(\int_0^{\infty} x e^{-x} dx \right) \left(\int_0^{\infty} y e^{-y} dy \right) = \left(\int_0^{\infty} x e^{-x} dx \right)^2 = (1)^2 = 1$$

Defn Let $f(x_1, x_2)$ and $F(x_1, x_2) = \Pr(X_1 \leq x_1, X_2 \leq x_2)$ be the joint p.d.f and the joint c.d.f of r.v.s X_1 and X_2 . Then the marginal c.d.f of X_1 and the marginal c.d.f of X_2 are respectively

$$F_1(x_1) = F(x_1, \infty) \text{ and } F_2(x_2) = F(\infty, x_2)$$

Note If $f_1(x_1)$ and $f_2(x_2)$ are the marginal p.d.f's, then the marginal c.d.f's $F_1(x_1)$ and $F_2(x_2)$ can be obtained

as :

$$F_1(x_1) = \begin{cases} \sum_{t_1=-\infty}^{x_1} f_1(t_1) & , \text{ disc} \\ \int_{-\infty}^{x_1} f_1(t_1) dt_1 & , \text{ cont.} \end{cases}$$

$$F_2(x_2) = \begin{cases} \sum_{t_2=-\infty}^{x_2} f_2(t_2) & , \text{ disc} \\ \int_{-\infty}^{x_2} f_2(t_2) dt_2 & , \text{ cont.} \end{cases}$$

Example (2) Back to Example (1) and find the joint c.d.f of X and Y . Also find the marginal c.d.f's.

(i) We have $f(x, y) = \frac{1}{2!}(x+y)$, $x=1, 2, 3$, $y=1, 2$
 $= 0$, e.w

$$\begin{aligned} F(x, y) &= \sum_{t=-\infty}^x \sum_{s=-\infty}^y f(t, s) = \sum_{t=1}^x \sum_{s=1}^y \frac{1}{2!}(t+s) = \frac{1}{2!} \sum_{t=1}^x [(t+1) + (t+2) + \dots + (t+y)] \\ &= \frac{1}{2!} \sum_{t=1}^x \left[yt + \frac{y(y+1)}{2} \right] = \frac{1}{2!} \left[\frac{y x(x+1)}{2} + \frac{x y(y+1)}{2} \right] = \frac{1}{4!} xy(x+y+2) \end{aligned}$$

$$F(x, y) = \begin{cases} 0 & , x < 1, y < 1 \\ \frac{1}{42} xy(x+y+2) & , 1 \leq x < 3, 1 \leq y < 2 \\ 1 & , x \geq 3, y \geq 2 \end{cases}$$

The marginal c.d.f of X and the marginal c.d.f of Y are :-

$$F_1(x) = F(x, 2) = \begin{cases} 0 & , x < 1 \\ \frac{1}{21} x(x+4) & , 1 \leq x < 3 \\ 1 & , x \geq 3 \end{cases}$$

$$F_2(y) = F(3, y) = \begin{cases} 0 & , y < 1 \\ \frac{1}{14} y(y+5) & , 1 \leq y < 2 \\ 1 & , y \geq 2 \end{cases}$$

$$(ii) f(x, y) = e^{-(x+y)} , 0 < x < \infty, 0 < y < \infty$$

= 0 , e.w

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(t, s) dt ds = \int_0^x \int_0^y e^{-(t+s)} ds dt = \left(\int_0^x e^{-t} dt \right) \left(\int_0^y e^{-s} ds \right)$$

$$= \left(-e^{-t} \Big|_0^x \right) \left(-e^{-s} \Big|_0^y \right) = \begin{cases} (1 - e^{-x})(1 - e^{-y}) & , x \geq 0, y \geq 0 \\ 0 & , 0 < x < \infty, 0 < y < \infty \\ 1 & , x = y = \infty \end{cases}$$

$$F_1(x) = F(x, \infty) = \begin{cases} 0 & , x < 0 \\ 1 - e^{-x} & , 0 \leq x < \infty \\ 1 & , x = \infty \end{cases}$$

$$F_2(y) = F(\infty, y) = \begin{cases} 0 & , y < 0 \\ 1 - e^{-y} & , 0 \leq y < \infty \\ 1 & , y = \infty \end{cases}$$

The Covariance and The Correlation Coefficient

Let the r.v's X and Y have means μ_1, μ_2 and variances σ_1^2, σ_2^2 respectively. The mathematical expectation $E[(X - \mu_1)(Y - \mu_2)]$ is called the covariance of X and Y , and denoted by $\text{cov}(X, Y)$.

Note: $\text{cov}(X, Y) = E[(X - \mu_1)(Y - \mu_2)] = E[XY - \mu_1 Y - \mu_2 X + \mu_1 \mu_2]$
 $= E(XY) - \mu_1 E(Y) - \mu_2 E(X) + \mu_1 \mu_2$
 $= E(XY) - \mu_1 \mu_2 - \mu_2 \mu_1 + \mu_1 \mu_2 = E(XY) - \mu_1 \mu_2$

Defn The correlation coefficient between two r.v's X and Y , denoted by $\rho = \rho(X, Y)$, is defined to be

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{E(XY) - \mu_1 \mu_2}{\sigma_1 \sigma_2}$$

Note: $-1 \leq \rho \leq 1$

Example (3) Let the joint p.d.f of r.v's X and Y be

$$f(x, y) = x + y, \quad 0 < x < 1, \quad 0 < y < 1$$

$= 0$, e.w

Compute the correlation coefficient between X and Y .

Solution

the marginal p.d.f of X is

$$f_1(x) = \int_y f(x, y) dy = \int_{y=0}^1 (x + y) dy = xy + \frac{1}{2} y^2 \Big|_0^1 = x + \frac{1}{2}, \quad 0 < x < 1$$

$= 0$, e.w

Similarly, the marginal p.d.f of Y is

$$f_2(y) = y + \frac{1}{2}, \quad 0 < y < 1$$

$= 0$, e.w

$$\mu_1 = E(X) = \int_x x f_1(x) dx = \int_0^1 x(x + \frac{1}{2}) dx = \frac{x^3}{3} + \frac{x^2}{4} \Big|_0^1 = \frac{7}{12} = E(Y) = \mu_2$$

$$E(X^2) = \int_x x^2 f_1(x) dx = \int_0^1 x^2(x + \frac{1}{2}) dx = \frac{x^4}{4} + \frac{x^3}{6} \Big|_0^1 = \frac{5}{12}$$

$$\sigma_1^2 = \text{Var}(X) = E(X^2) - \mu_1^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144} = \text{Var}(Y) = \sigma_2^2$$

$$E(XY) = \int_x \int_y xy f(x, y) dx dy = \int_0^1 \int_0^1 xy(x + y) dy dx = \int_0^1 \int_0^1 (x^2 y + xy^2) dy dx$$

$$= \int_0^1 \left(\frac{x^2 y^2}{2} + \frac{xy^3}{3} \right) \Big|_0^1 dx = \int_0^1 \left(\frac{x^2}{2} + \frac{x}{3} \right) dx = \frac{x^3}{6} + \frac{x^2}{6} \Big|_0^1 = \frac{1}{3}$$

$$\text{Cov}(X, Y) = E(XY) - \mu_1 \mu_2 = \frac{1}{3} - \left(\frac{7}{12}\right)\left(\frac{7}{12}\right) = -\frac{1}{144}$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144}} \sqrt{\frac{11}{144}}} = -\frac{1}{11}$$

Joint Moment Generating Function "In two-D case"

Let $f(x, y)$ be the joint p.d.f of r.v's X, Y . The mathematical expectation $E(e^{t_1 X + t_2 Y})$, if it exists for $-h_1 < t_1 < h_1, -h_2 < t_2 < h_2$, is called the joint m.g.f of X and Y , and denoted by $M(t_1, t_2)$.

That is

$$M(t_1, t_2) = E(e^{t_1 X + t_2 Y}) = \begin{cases} \sum_x \sum_y e^{t_1 x + t_2 y} f(x, y) & \text{for disc. case} \\ \int_x \int_y e^{t_1 x + t_2 y} f(x, y) dx dy & \text{for cont. case} \end{cases}$$

If we set $t_1 = t_2 = 0$, we have $M(0, 0) = 1$

Notes

(i) The marginal m.g.f of X and the marginal m.g.f of Y can be obtained as

$$M(t_1, 0) = E(e^{t_1 X}) = M_1(t_1)$$

$$M(0, t_2) = E(e^{t_2 Y}) = M_2(t_2)$$

(ii) If we differentiate $M(t_1, t_2)$ partially k times w.r. to t_1 and m times w.r. to t_2 , we have

$$\frac{\partial^{k+m} M(t_1, t_2)}{\partial t_2^m \partial t_1^k} = \begin{cases} \sum_x \sum_y x^k y^m e^{t_1 x + t_2 y} f(x, y) & \text{for disc case} \\ \int_x \int_y x^k y^m e^{t_1 x + t_2 y} f(x, y) dx dy & \text{for cont. case} \end{cases}$$

set $t_1 = t_2 = 0$, we have

$$\left. \frac{\partial^{k+m} M(t_1, t_2)}{\partial t_2^m \partial t_1^k} \right|_{t_1=t_2=0} = \begin{cases} \sum_x \sum_y x^k y^m f(x, y) \\ \int_x \int_y x^k y^m f(x, y) dx dy \end{cases} \rightarrow E(X^k Y^m)$$

In particular

$$\frac{\partial M(t_1, t_2)}{\partial t_1} \Big|_{t_1=t_2=0} = \frac{\partial M(0,0)}{\partial t_1} = E(X) = \mu_1$$

$$\frac{\partial M(t_1, t_2)}{\partial t_2} \Big|_{t_1=t_2=0} = \frac{\partial M(0,0)}{\partial t_2} = E(Y) = \mu_2$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_1^2} \Big|_{t_1=t_2=0} = \frac{\partial^2 M(0,0)}{\partial t_1^2} = E(X^2) \Rightarrow \text{Var}(X) = \sigma_1^2 = E(X^2) - \mu_1^2$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_2^2} \Big|_{t_1=t_2=0} = \frac{\partial^2 M(0,0)}{\partial t_2^2} = E(Y^2) \Rightarrow \text{Var}(Y) = \sigma_2^2 = E(Y^2) - \mu_2^2$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_2 \partial t_1} \Big|_{t_1=t_2=0} = \frac{\partial^2 M(0,0)}{\partial t_2 \partial t_1} = E(XY)$$

Then

$$\text{Cov}(X, Y) = E(XY) - \mu_1 \mu_2 \quad \text{and} \quad \rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2}$$

(iii) If we set $\psi(t_1, t_2) = \ln M(t_1, t_2)$, then

$$\frac{\partial \psi(t_1, t_2)}{\partial t_i} = \frac{\frac{\partial M(t_1, t_2)}{\partial t_i}}{M(t_1, t_2)}, \quad i=1, 2$$

$$\Rightarrow \frac{\partial \psi(0,0)}{\partial t_i} = \frac{\frac{\partial M(0,0)}{\partial t_i}}{M(0,0)} = \frac{\mu_i}{1} = \mu_i, \quad i=1, 2$$

$$\frac{\partial^2 \psi(t_1, t_2)}{\partial t_i^2} = \frac{M(t_1, t_2) \frac{\partial^2 M(t_1, t_2)}{\partial t_i^2} - \frac{\partial M(t_1, t_2)}{\partial t_i} \cdot \frac{\partial M(t_1, t_2)}{\partial t_i}}{[M(t_1, t_2)]^2}, \quad i=1, 2$$

$$\frac{\partial^2 \psi(0,0)}{\partial t_i^2} = \frac{M(0,0) \frac{\partial^2 M(0,0)}{\partial t_i^2} - \frac{\partial M(0,0)}{\partial t_i} \cdot \frac{\partial M(0,0)}{\partial t_i}}{[M(0,0)]^2}$$

$$= \frac{(1) E(X_i^2) - E(X_i) \cdot E(X_i)}{(1)^2} = \sigma_i^2, \quad i=1, 2$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_2 \partial t_1} = \frac{M(t_1, t_2) \frac{\partial^3 M(t_1, t_2)}{\partial t_2 \partial t_1} - \frac{\partial M(t_1, t_2)}{\partial t_1} \cdot \frac{\partial M(t_1, t_2)}{\partial t_2}}{[M(t_1, t_2)]^2}$$

$$\frac{\partial^2 \gamma(0,0)}{\partial t_2 \partial t_1} = \frac{(1) E(XY) - E(X) \cdot E(Y)}{(1)^2} = \text{cov}(X, Y)$$

Example (4) Let the joint p.d.f of r.v X and Y be

$$f(x, y) = e^{-y}, 0 < x < y < \infty$$

$$= 0, \text{ e.w.}$$

- (1) Find the marginal p.d.f of X and the marginal p.d.f of Y .
- (2) Find $\mu_1 = E(X)$, $\sigma_1^2 = \text{var}(X)$, $\mu_2 = E(Y)$, $\sigma_2^2 = \text{var}(Y)$, $E(XY)$, $\text{cov}(X, Y)$, and $\rho(X, Y)$ by using direct expectation approach.
- (3) Find the joint c.d.f of X and Y .
- (4) Find the marginal c.d.f of X and the marginal c.d.f of Y .
- (5) Find the joint m.g.f of X and Y .
- (6) Find the marginal m.g.f of X and the marginal m.g.f of Y .
- (7) Use the joint m.g.f and find $E(X)$, $\text{var}(X)$, $E(Y)$, $\text{var}(Y)$, $\text{cov}(X, Y)$ and $\rho(X, Y)$.

Solution

(1) The marginal p.d.f of X is $f_1(x) = \int_y f(x, y) dy = \int_{y=x}^{\infty} e^{-y} dy$

$$f_1(x) = -e^{-y} \Big|_x^{\infty} = -(e^{-\infty} - e^{-x}) = e^{-x}, 0 < x < \infty$$

$$= 0, \text{ e.w.}$$

The marginal p.d.f of Y is $f_2(y) = \int_x f(x, y) dx = \int_{x=0}^y e^{-y} dx = e^{-y} x \Big|_0^y$

$$f_2(y) = e^{-y}(y-0) = y e^{-y}, 0 < y < \infty$$

$$= 0, \text{ e.w.}$$

(2) $\mu_1 = E(X) = \int_x x f_1(x) dx = \int_0^{\infty} x e^{-x} dx = 1$

$$E(X^2) = \int_x x^2 f_1(x) dx = \int_0^{\infty} x^2 e^{-x} dx = 2$$

$$\Rightarrow \sigma_1^2 = \text{var}(X) = E(X^2) - \mu_1^2 = 2 - (1)^2 = 1$$

$$\mu_2 = E(Y) = \int_y y f_2(y) dy = \int_0^{\infty} y \cdot y \bar{e}^y dy = \int_0^{\infty} y^2 \bar{e}^y dy = 2$$

$$E(Y^2) = \int_y y^2 f_2(y) dy = \int_0^{\infty} y^2 \cdot y \bar{e}^y dy = \int_0^{\infty} y^3 \bar{e}^y dy = 6$$

$$\Rightarrow \sigma_2^2 = \text{Var}(Y) = E(Y^2) - \mu_2^2 = 6 - (2)^2 = 2$$

$$E(XY) = \int_x \int_y xy f(x,y) dx dy = \int_{x=0}^{\infty} \int_{y=x}^{\infty} xy \bar{e}^y dy dx = \text{or } \int_{y=0}^{\infty} \int_{x=0}^y xy \bar{e}^y dx dy$$

$$= \int_{y=0}^{\infty} y \bar{e}^y \left(\frac{x^2}{2} \right) \Big|_0^y dy = \frac{1}{2} \int_0^{\infty} y^3 \bar{e}^y dy = \frac{6}{2} = 3$$

$$\text{Cov}(X, Y) = E(XY) - \mu_1 \mu_2 = 3 - (1)(2) = 1$$

$$\rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{1}{\sqrt{1} \sqrt{2}} = \frac{1}{\sqrt{2}}$$

(3) The joint c.d.f of X and Y is

$$\begin{aligned} F(x, y) &= \Pr(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(t, s) dt ds = \int_{t=0}^x \int_{s=t}^y \bar{e}^s ds dt \\ &= \int_{t=0}^x \left(-\bar{e}^s \right) \Big|_t^y dt = \int_{t=0}^x (\bar{e}^t - \bar{e}^y) dt = -\bar{e}^t - \bar{e}^y t \Big|_0^x \\ &= (-\bar{e}^x - x \bar{e}^y) - (-1 - 0) = \begin{cases} 1 - \bar{e}^x - x \bar{e}^y, & 0 < x < y < \infty \\ 1, & x = y, y = \infty \end{cases} \end{aligned}$$

(4) The marginal c.d.f of X and the marginal c.d.f of Y are:-

$$F_1(x) = F(x, \infty) = \begin{cases} 1 - \bar{e}^x, & 0 < x < \infty \\ 1, & x = \infty \end{cases}, F_2(y) = F(y, y) = \begin{cases} 1 - \bar{e}^y - y \bar{e}^y, & 0 < y < \infty \\ 1, & y = \infty \end{cases}$$

(5) The joint m.g.f of X and Y is

$$M(t_1, t_2) = E(e^{t_1 X + t_2 Y}) = \int_0^\infty \int_0^\infty e^{t_1 x + t_2 y} f(x, y) dx dy = \int_0^\infty \int_0^\infty e^{t_1 x + t_2 y} e^{-y} dy dx$$

$$= \int_0^\infty e^{t_1 x} \left\{ \int_0^\infty e^{-(1-t_2)y} dy \right\} dx = \int_0^\infty e^{t_1 x} \left(-\frac{1}{1-t_2} e^{-(1-t_2)y} \right) \Big|_0^\infty dx$$

$$= -\frac{1}{1-t_2} \int_0^\infty e^{t_1 x} [0 - e^{-(1-t_2)x}] dx = \frac{1}{1-t_2} \int_0^\infty e^{-(1-t_1-t_2)x} dx$$

$$= -\frac{1}{(1-t_2)(1-t_1-t_2)} e^{-(1-t_1-t_2)x} \Big|_0^\infty = -\frac{1}{(1-t_2)(1-t_1-t_2)} (0 - 1)$$

$$\therefore M(t_1, t_2) = \frac{1}{(1-t_2)(1-t_1-t_2)} = (1-t_2)^{-1} (1-t_1-t_2)^{-1}, \quad t_2 < 1 \text{ and } t_1 + t_2 < 1$$

$$\text{check: } M(0, 0) = 1$$

(6) The marginal m.g.f of X and the marginal m.g.f of Y are:-

$$M_1(t_1) = M(t_1, 0) = \frac{1}{(1-t_1)} = (1-t_1)^{-1}, \quad M_2(t_2) = M(0, t_2) = \frac{1}{(1-t_2)^2} = (1-t_2)^{-2}$$

$$(7) \frac{\partial M(t_1, t_2)}{\partial t_1} = (1-t_2)^{-1} (1-t_1-t_2)^{-2} \Rightarrow \frac{\partial M(0, 0)}{\partial t_1} = 1 = E(X) = \mu_1$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_1^2} = 2(1-t_2)^{-1} (1-t_1-t_2)^{-3} \Rightarrow \frac{\partial^2 M(0, 0)}{\partial t_1^2} = 2 = E(X^2)$$

$$\Rightarrow \sigma_1^2 = \text{Var}(X) = E(X^2) - \mu_1^2 = 2 - (1)^2 = 1$$

$$\frac{\partial M(t_1, t_2)}{\partial t_2} = (1-t_2)^{-1} (1-t_1-t_2)^{-2} + (1-t_1-t_2)^{-1} (1-t_2)^{-2} \Rightarrow \frac{\partial M(0, 0)}{\partial t_2} = 1 + 1 = 2 = E(Y)$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_2^2} = 2(1-t_2)^{-1} (1-t_1-t_2)^{-3} + (1-t_1-t_2)^{-2} (1-t_2)^{-2} + 2(1-t_1-t_2)^{-1} (1-t_2)^{-3} + (1-t_2)^{-2} (1-t_1-t_2)^{-2}$$

and
defn
Let
the marg

$$\Rightarrow \frac{\partial^2 M(0,0)}{\partial t_2^2} = 2+1+2+1 = 6 = E(Y^2)$$

$$\Rightarrow \sigma_2^2 = \text{Var}(Y) = E(Y^2) - \mu_2^2 = 6 - (2)^2 = 2$$

$$\frac{\partial^2 M(t_1, t_2)}{\partial t_2 \partial t_1} = \frac{\partial^2 M(t_1, t_2)}{\partial t_1 \partial t_2} = 2(1-t_2)^{-1}(1-t_1-t_2)^{-3} + (1-t_1-t_2)^{-2}(1-t_2)^{-2}$$

$$\Rightarrow \frac{\partial^2 \psi(0,0)}{\partial t_2 \partial t_1} = 2+1 = 3 = E(XY)$$

$$\Rightarrow \text{Cov}(X, Y) = E(XY) - \mu_1 \mu_2 = 3 - (1)(2) = 1$$

$$\Rightarrow \rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{1}{\sqrt{1} \sqrt{2}} = \frac{1}{\sqrt{2}}$$

OR

$$\psi(t_1, t_2) = \ln M(t_1, t_2) = -\ln(1-t_2) - \ln(1-t_1-t_2)$$

$$\frac{\partial \psi(t_1, t_2)}{\partial t_1} = -\frac{(-1)}{1-t_1-t_2} = (1-t_1-t_2)^{-1} \Rightarrow \frac{\partial \psi(0,0)}{\partial t_1} = 1 = E(X) = \mu_1$$

$$\frac{\partial^2 \psi(t_1, t_2)}{\partial t_1^2} = (1-t_1-t_2)^{-2} \Rightarrow \frac{\partial^2 \psi(0,0)}{\partial t_1^2} = 1 = \text{Var}(X) = \sigma_1^2$$

$$\frac{\partial \psi(t_1, t_2)}{\partial t_2} = -\frac{(-1)}{(1-t_2)} - \frac{(-1)}{1-t_1-t_2} = (1-t_2)^{-1} + (1-t_1-t_2)^{-1}$$

$$\Rightarrow \frac{\partial \psi(0,0)}{\partial t_2} = 1+1 = 2 = E(Y) = \mu_2$$

$$\frac{\partial^2 \psi(t_1, t_2)}{\partial t_2^2} = (1-t_2)^{-2} + (1-t_1-t_2)^{-2} \Rightarrow \frac{\partial^2 \psi(0,0)}{\partial t_2^2} = 1+1 = 2 = \text{Var}(Y) = \sigma_2^2$$

$$\frac{\partial^2 \psi(t_1, t_2)}{\partial t_2 \partial t_1} = \frac{\partial^2 \psi(t_1, t_2)}{\partial t_1 \partial t_2} = (1-t_1-t_2)^{-2} \Rightarrow \frac{\partial^2 \psi(0,0)}{\partial t_2 \partial t_1} = 1 = \text{Cov}(X, Y)$$

$$\Rightarrow \rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{1}{\sqrt{1} \sqrt{2}} = \frac{1}{\sqrt{2}}$$

Conditional Distribution

"In 2-D space"

Defn

Let $f_1(x)$, $f_2(y)$, and $f(x,y)$ be respectively the marginal p.d.f of X , the marginal p.d.f of Y , and the joint p.d.f of X and Y . We define the conditional p.d.f of X given $Y=y$ as

$$f_1(x|y) = \frac{f(x,y)}{f_2(y)}, \quad f_2(y) \neq 0$$

and the conditional p.d.f of Y given $X=x$ as

$$f_2(y|x) = \frac{f(x,y)}{f_1(x)}, \quad f_1(x) \neq 0$$

To show that $f_1(x|y)$ is valid p.d.f, we note that

(i) $f_1(x|y) \geq 0$, because $f(x,y) \geq 0$ and $f_2(y) > 0$

$$(ii) \sum_x f_1(x|y) = \sum_x \frac{f(x,y)}{f_2(y)} = \frac{1}{f_2(y)} \sum_x f(x,y) = \frac{f_2(y)}{f_2(y)} = 1$$

or

$$\int_x f_1(x|y) dx = \int_x \frac{f(x,y)}{f_2(y)} dx = \frac{1}{f_2(y)} \int_x f(x,y) dx = \frac{f_2(y)}{f_2(y)} = 1$$

Notes

$$(1) \Pr(a \leq X \leq b | Y=y) = \begin{cases} \sum_{x=a}^b f_1(x|y) & , \text{ for disc. case} \\ \int_a^b f_1(x|y) dx & , \text{ for cont. case} \end{cases}$$

$$\Pr(c \leq Y \leq d | X=x) = \begin{cases} \sum_{y=c}^d f_2(y|x) & , \text{ disc} \\ \int_c^d f_2(y|x) dy & , \text{ cont.} \end{cases}$$

(2) If $u(X)$ is a function of X alone, then the conditional expectation of $u(X)$ given $Y=y$ is

$$E[u(X) | Y=y] = \begin{cases} \sum_x u(x) f_1(x|y) & , \text{ disc} \\ \int_x u(x) f_1(x|y) dx & , \text{ cont.} \end{cases}$$

If $u(Y)$ is a function of Y alone, then the conditional ~~amp~~ of $u(Y)$ given $X=x$ is

$$E[u(Y)|X=x] = \begin{cases} \sum_y u(y) f_2(y|x) & , \text{ disc} \\ \int u(y) f_2(y|x) dy & , \text{ cont.} \end{cases}$$

In particular the conditional mean and the conditional variance of X given $Y=y$

$$E[X|Y=y] = \mu_{x|y} = \begin{cases} \sum_x x f_1(x|y) & , \text{ disc} \\ \int x f_1(x|y) dx & , \text{ cont.} \end{cases}$$

$$E[X^2|Y=y] = \begin{cases} \sum_x x^2 f_1(x|y) & , \text{ disc} \\ \int x^2 f_1(x|y) dx & , \text{ cont.} \end{cases}$$

and

$$\text{Var}(X|Y=y) = \sigma_{x|y}^2 = E[X^2|Y=y] - \mu_{x|y}^2$$

Similarly the conditional mean and the conditional variance of Y given $X=x$

$$E[Y|X=x] = \mu_{y|x} = \begin{cases} \sum_y y f_2(y|x) & , \text{ disc} \\ \int y f_2(y|x) dy & , \text{ cont.} \end{cases}$$

$$E[Y^2|X=x] = \begin{cases} \sum_y y^2 f_2(y|x) & , \text{ disc} \\ \int y^2 f_2(y|x) dy & , \text{ cont.} \end{cases}$$

and

$$\text{Var}(Y|X=x) = \sigma_{y|x}^2 = E[Y^2|X=x] - \mu_{y|x}^2$$

Example (5) Back to Example (4) and find

- (1) The conditional p.d.f of X given $Y=y$
- (2) The conditional p.d.f of Y given $X=x$
- (3) $E(X|Y=y)$ and $Var(X|Y=y)$
- (4) $E(Y|X=x)$ and $Var(Y|X=x)$, (5) $P(1 \leq X \leq 2)$ and $P(1 \leq X \leq 2 | Y=2)$

Solution

We have $f(x,y) = e^{-y}$, $0 < x < y < \infty$

$= 0$, e.w

$$f_1(x) = e^{-x}, 0 < x < \infty, f_2(y) = y e^{-y}, 0 < y < \infty$$

$= 0$, e.w

$= 0$, e.w

- (1) The conditional p.d.f of X given $Y=y$ is

$$f_1(x|y) = \frac{f(x,y)}{f_2(y)} = \frac{e^{-y}}{y e^{-y}} = \frac{1}{y}, 0 < x < y \text{ for any } 0 < y < \infty$$

$= 0$, e.w

- (2) The conditional p.d.f of Y given $X=x$ is

$$f_2(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, x < y < \infty \text{ for any } 0 < x < \infty$$

$= 0$, e.w

$$(3) \mu_{x|y} = E(X|Y=y) = \int_x x f_1(x|y) dx = \int_{x=0}^y x \frac{1}{y} dx = \frac{1}{y} \left(\frac{x^2}{2} \right)_0^y = \frac{y}{2}$$

$$E(X^2|Y=y) = \int_x x^2 f_1(x|y) dx = \int_{x=0}^y x^2 \frac{1}{y} dx = \frac{1}{y} \left(\frac{x^3}{3} \right)_0^y = \frac{y^2}{3}$$

$$\sigma_{x|y}^2 = Var(X|Y=y) = E(X^2|Y=y) - \mu_{x|y}^2 = \frac{y^2}{3} - \left(\frac{y}{2} \right)^2 = \frac{y^2}{12}$$

$$(4) \mu_{y|x} = E(Y|X=x) = \int_y y f_2(y|x) dy = \int_{y=x}^{\infty} y e^{-(y-x)} dy = e^x \int_x^{\infty} y e^{-y} dy$$

$$= e^x \left(-y e^{-y} - e^{-y} \right) \Big|_x^{\infty} = e^x (x e^{-x} + e^{-x}) = x+1$$

$$E(Y^2|X=x) = \int_y y^2 f_2(y|x) dy = \int_{y=x}^{\infty} y^2 e^{-(y-x)} dy = e^x \int_x^{\infty} y^2 e^{-y} dy$$

$$= e^x \left(-y^2 e^{-y} - 2y e^{-y} - 2e^{-y} \right) \Big|_x^{\infty} = e^x (x^2 e^{-x} + 2x e^{-x} + 2e^{-x}) = x^2 + 2x + 2$$

$$\sigma_{Y|X}^2 = \text{Var}(Y|X=x) = E(Y^2|X=x) - \mu_{Y|X}^2 = x^2 + 2x + 2 - (x+1)^2$$

$$= x^2 + 2x + 2 - x^2 - 2x - 1 = 1$$

$$(5) \Pr(1 \leq X \leq 2) = \int_{x=1}^2 f_1(x) dx = \int_1^2 \bar{e}^x dx = -\bar{e}^x \Big|_1^2 = \bar{e}^1 - \bar{e}^2$$

$$\Pr(1 \leq X \leq 2 | Y=2) = \int_{x=1}^2 f_1(x|Y=2) dx = \int_1^2 \frac{1}{2} dx = \frac{x}{2} \Big|_1^2 = \frac{1}{2}(2-1) = \frac{1}{2}$$

Example (6) Let the r.v. X has p.d.f $f(x)$ and c.d.f $F(x)$.
 Define $f(x|X > x_0) = \frac{f(x)}{1-F(x_0)}$, $x_0 < x < \infty$, x_0 fixed number
 $= 0$, e.w., $F(x_0) \neq 1$

(a) Show that $f(x|X > x_0)$ be a conditional p.d.f of X given $X > x_0$

(b) Let $f(x) = \bar{e}^x$, $0 < x < \infty$

$= 0$, e.w. Compute $\Pr(X > 2 | X > 1)$

Solution

(a) (i) $f(x|X > x_0) \geq 0$, yes, because $f(x) \geq 0$ and $1-F(x_0) > 0$

$$(ii) \int_{x_0}^{\infty} f(x|X > x_0) dx = \int_{x_0}^{\infty} \frac{f(x)}{1-F(x_0)} dx = \frac{1}{1-F(x_0)} \int_{x_0}^{\infty} f(x) dx$$

$$= \frac{1}{1-F(x_0)} F(x) \Big|_{x_0}^{\infty} = \frac{F(\infty) - F(x_0)}{1-F(x_0)} = \frac{1-F(x_0)}{1-F(x_0)} = 1$$

$$(b) f(x) = \bar{e}^x, 0 < x < \infty \Rightarrow F(x) = \int_{-\infty}^x f(t) dt = \int_0^x \bar{e}^{-t} dt = -\bar{e}^{-t} \Big|_0^x = 1 - \bar{e}^{-x}$$

$$= 0, \text{ e.w.}$$

$$f(x|X > 1) = \frac{f(x)}{1-F(1)} = \frac{\bar{e}^x}{1-1+\bar{e}^1} = e \cdot \bar{e}^x, 1 < x < \infty$$

$$= 0, \text{ e.w.}$$

$$\Pr(X > 2 | X > 1) = \int_2^{\infty} f(x|X > 1) dx = \int_2^{\infty} e \cdot \bar{e}^x dx = -e(\bar{e}^x) \Big|_2^{\infty}$$

$$= -e(0 - \bar{e}^2) = \bar{e}^1$$

Example (2), we have

$$F(x, y) = \frac{1}{42} xy(x+y+2), \quad F_1(x) = \frac{1}{21} x(x+4), \quad F_2(y) = \frac{1}{14} y(y+5)$$

$$F_1(x) \cdot F_2(y) = \frac{1}{294} xy(x+4)(y+5) \neq \frac{1}{42} xy(x+y+2) = F(x, y) \\ \Rightarrow X \text{ and } Y \text{ are stochastically dependent}$$

(iii) $F(x, y) = (1 - e^{-x}) / (1 - e^{-y})$, $F_1(x) = 1 - e^{-x}$, $F_2(y) = 1 - e^{-y}$

$$F_1(x) \cdot F_2(y) = (1 - e^{-x})(1 - e^{-y}) = F(x, y) \Rightarrow X \text{ and } Y \text{ are stochastically independent.}$$

In Example (3), we have

$$f(x, y) = x + y, \quad f_1(x) = x + \frac{1}{2}, \quad f_2(y) = y + \frac{1}{2}$$

$$f_1(x) \cdot f_2(y) = (x + \frac{1}{2})(y + \frac{1}{2}) \neq x + y = f(x, y) \Rightarrow X \text{ and } Y \text{ are stochastically dependent.}$$

In Example (4), we have

$$f(x, y) = e^{-y}, \quad f_1(x) = e^{-x}, \quad f_2(y) = y e^{-y}$$

$$f_1(x) \cdot f_2(y) = e^{-x} \cdot y e^{-y} = y e^{-(x+y)} \neq e^{-y} = f(x, y) \Rightarrow X \text{ and } Y \text{ are stochastically dependent.}$$

Theorem (1) Let $f(x_1, x_2)$ be the joint p.d.f of r.v.'s X_1 and X_2 . The r.v.'s X_1 and X_2 are stochastically independent iff $f(x_1, x_2)$ can be written as a product of a non-negative function of x_1 alone and a non-negative function of x_2 alone, that is

$$f(x_1, x_2) = g(x_1) \cdot h(x_2) \text{ where } g(x_1) > 0 \text{ and } h(x_2) > 0$$

proof

$\Rightarrow$ Let the r.v.'s X_1 and X_2 are stochastically independent, then

$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$, where $f_1(x_1)$ is the marginal p.d.f of r.v. X_1 which is a function of x_1 alone and $f_2(x_2)$ is the marginal p.d.f of X_2 which is a function of x_2 alone.

Thus, the condition $f(x_1, x_2) = g(x_1) h(x_2)$ is fulfilled.

$\Leftarrow$ Let $f(x_1, x_2) = g(x_1) h(x_2)$, where $g(x_1) > 0$ is a function of x_1 alone and $h(x_2) > 0$ is a function of x_2 alone.

Consider the cont. case

$$f_1(x_1) = \int_{x_2} f(x_1, x_2) dx_2 = \int_{x_2} g(x_1) h(x_2) dx_2 = g(x_1) \int_{x_2} h(x_2) dx_2 = c_1 g(x_1)$$

and

$$f_2(x_2) = \int_{x_1} f(x_1, x_2) dx_1 = \int_{x_1} g(x_1) h(x_2) dx_1 = h(x_2) \int_{x_1} g(x_1) dx_1 = c_2 h(x_2)$$

$$\text{but } 1 = \int_{x_1} \int_{x_2} f(x_1, x_2) dx_1 dx_2 = \int_{x_1} \int_{x_2} g(x_1) h(x_2) dx_1 dx_2$$

$$= \left(\int_{x_1} g(x_1) dx_1 \right) \cdot \left(\int_{x_2} h(x_2) dx_2 \right) = c_1 c_2$$

$$\therefore f(x_1, x_2) = g(x_1) \cdot h(x_2) = c_1 g(x_1) \cdot c_2 h(x_2) = f_1(x_1) \cdot f_2(x_2)$$

$\therefore$ The r.v.'s X_1 and X_2 are st. indep.

Theorem (2) If the r.v.'s X_1 and X_2 are st. indep., then

$$(1) \Pr(a \leq X_1 \leq b, c \leq X_2 \leq d) = \Pr(a \leq X_1 \leq b) \cdot \Pr(c \leq X_2 \leq d)$$

$$(2) E[u(X_1) \cdot v(X_2)] = E[u(X_1)] \cdot E[v(X_2)], \text{ where } u(x_1) \text{ is a function of } x_1 \text{ alone and } v(x_2) \text{ is a function of } x_2 \text{ alone.}$$

proof

Consider the cont. case

$$\begin{aligned} (1) \Pr(a \leq X_1 \leq b, c \leq X_2 \leq d) &= \int_a^b \int_c^d f(x_1, x_2) dx_1 dx_2 \\ &= \int_a^b \int_c^d f_1(x_1) \cdot f_2(x_2) dx_1 dx_2 \\ &= \left(\int_a^b f_1(x_1) dx_1 \right) \cdot \left(\int_c^d f_2(x_2) dx_2 \right) \\ &= \Pr(a \leq X_1 \leq b) \cdot \Pr(c \leq X_2 \leq d) \end{aligned}$$

$$\begin{aligned} (2) E[u(X_1) \cdot v(X_2)] &= \int_{x_1} \int_{x_2} u(x_1) \cdot v(x_2) f(x_1, x_2) dx_1 dx_2 \\ &= \int_{x_1} \int_{x_2} u(x_1) \cdot v(x_2) f_1(x_1) \cdot f_2(x_2) dx_1 dx_2 \\ &= \left(\int_{x_1} u(x_1) f_1(x_1) dx_1 \right) \cdot \left(\int_{x_2} v(x_2) f_2(x_2) dx_2 \right) \\ &= E[u(X_1)] \cdot E[v(X_2)] \end{aligned}$$

If the r.v.'s X and Y are stoch. indep., then

$$E(XY) = E(X) \cdot E(Y) = \mu_1 \mu_2$$

$$Cov(X, Y) = E(XY) - \mu_1 \mu_2 = E(X) \cdot E(Y) - \mu_1 \mu_2 = \mu_1 \mu_2 - \mu_1 \mu_2 = 0$$

$\Rightarrow \rho = \rho(X, Y) = 0$, but the converse is not true. That is if $\rho = 0$ does not imply that X and Y are stoch. indep.

Theorem (4) Let $f(x_1, x_2)$ be the joint p.d.f of r.v.'s X_1 and X_2 with marginals p.d.f's $f_1(x_1)$ and $f_2(x_2)$ and let $M(t_1, t_2)$ be the joint m.g.f of r.v.'s X_1 and X_2 with marginals m.g.f's $M_1(t_1)$ and $M_2(t_2)$. The r.v.'s X_1 and X_2 are stoch. indep. iff $M(t_1, t_2) = M_1(t_1) M_2(t_2)$

Proof

$\Rightarrow$ Let the r.v.'s X_1 and X_2 be stoch. indep., then

$$M(t_1, t_2) = E(e^{t_1 X_1 + t_2 X_2}) = E(e^{t_1 X_1} \cdot e^{t_2 X_2}) \stackrel{\text{by Theorem (3), part (2)}}{=} E(e^{t_1 X_1}) \cdot E(e^{t_2 X_2}) = M_1(t_1) \cdot M_2(t_2)$$

$\Leftarrow$ Let $M(t_1, t_2) = M_1(t_1) \cdot M_2(t_2)$ and consider the cont. case

$$M_1(t_1) = E(e^{t_1 X_1}) = \int_{x_1} e^{t_1 x_1} f_1(x_1) dx_1 \text{ and } M_2(t_2) = E(e^{t_2 X_2}) = \int_{x_2} e^{t_2 x_2} f_2(x_2) dx_2$$

So

$$\begin{aligned} M_1(t_1) \cdot M_2(t_2) &= \left(\int_{x_1} e^{t_1 x_1} f_1(x_1) dx_1 \right) \left(\int_{x_2} e^{t_2 x_2} f_2(x_2) dx_2 \right) \\ &= \iint_{x_1, x_2} e^{t_1 x_1 + t_2 x_2} f_1(x_1) f_2(x_2) dx_1 dx_2 \end{aligned}$$

But

$$M(t_1, t_2) = E(e^{t_1 X_1 + t_2 X_2}) = \iint_{x_1, x_2} e^{t_1 x_1 + t_2 x_2} f(x_1, x_2) dx_1 dx_2$$

and $M(t_1, t_2) = M_1(t_1) \cdot M_2(t_2)$, then

$$\iint_{x_1, x_2} e^{t_1 x_1 + t_2 x_2} f_1(x_1) f_2(x_2) dx_1 dx_2 = \iint_{x_1, x_2} e^{t_1 x_1 + t_2 x_2} f(x_1, x_2) dx_1 dx_2$$

that implies $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$

$\therefore$ the r.v.'s X_1 and X_2 are stoch. indep.

Example (8) Back to Example (4), we have
 $M(t_1, t_2) = (1-t_2)^{-1} (1-t_1-t_2)^{-1}$, $M_1(t_1) = (1-t_1)^{-1}$, $M_2(t_2) = (1-t_2)^{-1}$

$$M_1(t_1) \cdot M_2(t_2) = (1-t_1)^{-1} (1-t_2)^{-1} \neq (1-t_1-t_2)^{-1} = M(t_1, t_2)$$

$\therefore$ the r.v.'s X and Y are stochastically dependent.

Example (9) Let the joint m.g.f of r.v.'s X and Y be $M(t_1, t_2) = \frac{e^{t_1^2}}{1-2t_2}$

- (1) Find the marginal m.g.f of X and the marginal m.g.f of Y .
- (2) Test whether or not X and Y are stochastically independent.
- (3) Find the correlation coefficient between X and Y .

Solution

(1) The marginal m.g.f of X and the marginal p.d.f of Y are

$$M_1(t_1) = M(t_1, 0) = e^{t_1^2}, \quad M_2(t_2) = M(0, t_2) = (1-2t_2)^{-1}$$

$$(2) M_1(t_1) \cdot M_2(t_2) = \frac{e^{t_1^2}}{1-2t_2} \neq M(t_1, t_2) \Rightarrow X \text{ and } Y \text{ are stochastically dependent.}$$

$$(3) \text{ Let } \psi(t_1) = \ln M_1(t_1) = t_1^2 \Rightarrow \frac{\partial \psi(t_1)}{\partial t_1} = 2t_1 \Rightarrow \left. \frac{\partial \psi(t_1)}{\partial t_1} \right|_{t_1=0} = 0 = E(X)$$

$$\frac{\partial^2 \psi(t_1)}{\partial t_1^2} = 2 \Rightarrow \left. \frac{\partial^2 \psi(t_1)}{\partial t_1^2} \right|_{t_1=0} = 2 = \text{Var}(X) = \sigma_1^2$$

$$\text{Let } \psi(t_2) = \ln M_2(t_2) = -\ln(1-2t_2) \Rightarrow \frac{\partial \psi(t_2)}{\partial t_2} = \frac{2}{1-2t_2} = 2(1-2t_2)^{-1}$$

$$\Rightarrow \left. \frac{\partial \psi(t_2)}{\partial t_2} \right|_{t_2=0} = 2 = E(Y) = \mu_2, \quad \frac{\partial^2 \psi(t_2)}{\partial t_2^2} = 4(1-2t_2)^{-2}$$

$$\left. \frac{\partial^2 \psi(t_2)}{\partial t_2^2} \right|_{t_2=0} = 4 = \text{Var}(Y) = \sigma_2^2$$

$$\text{Let } \psi(t_1, t_2) = \ln M(t_1, t_2) = t_1^2 (1-2t_2)^{-1} \Rightarrow \ln(1-2t_2)$$

$$\frac{\partial \psi(t_1, t_2)}{\partial t_1} = 2t_1 (1-2t_2)^{-1} \Rightarrow \frac{\partial^2 \psi(t_1, t_2)}{\partial t_2 \partial t_1} = 4t_1 (1-2t_2)^{-2}$$

$$\left. \frac{\partial^2 \psi(t_1, t_2)}{\partial t_2 \partial t_1} \right|_{t_1=t_2=0} = 0 = \text{Cov}(X, Y) \Rightarrow \rho = \frac{\text{Cov}(X, Y)}{\sigma_1 \sigma_2} = \frac{0}{2\sqrt{2}} = 0$$

Sampling Theory:-

Defn. (Statistic):-

A function of one or more r.v's which is not depend on any unknown parameter is called a statistic

For example: If the r.v $X \sim N(\mu, \sigma^2)$, we may ask If $Y = \frac{X - \mu}{\sigma}$ is a statistic $\Rightarrow Y$ is not a statistic unless μ and σ are known but if $Y = \sum_{i=1}^n X_i$ then Y is a statistic.

Defn. (Random Sample $\equiv$ r.s)

Let $X_1, X_2, \dots, X_n$ be sto. indep. r.v's each of which has the same p.d.f $f(x)$, that is $f_i(x_i) = f(x_i)$ $i=1, 2, \dots, n$, So the joint p.d.f of these r.v's $f(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n) = \prod_{i=1}^n f(x_i)$ then the r.v's $X_1, X_2, \dots, X_n$ are said to constitute a random sample of size n from a distn. whose p.d.f $f(x)$

Example (1):

Let X_1, X_2, X_3 be sto. indep. r.v's with $X_1 \sim \text{Exp}(\lambda)$, $X_2 \sim N(\mu, \sigma^2)$, and $X_3 \sim \text{Be}(\alpha, \beta)$. Then the joint p.d.f of X_1, X_2 and X_3 is

$$f(x_1, x_2, x_3) = \frac{1}{\lambda} e^{-\frac{x_1}{\lambda}} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_2 - \mu}{\sigma}\right)^2} \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x_3^{\alpha-1} (1-x_3)^{\beta-1}$$

$$= 0, \quad \text{e.w} \quad \begin{aligned} 0 &< x_1 < \infty \\ -\infty &< x_2 < \infty \\ 0 &< x_3 < 1 \end{aligned}$$

Example (2):

Let X_1, X_2, X_3 be sto. indep. r.v's with $X_i \sim N(\mu_i, \sigma_i^2)$, $i=1, 2, 3$. Then the joint p.d.f of these r.v's is

$$f(x_1, x_2, x_3) = \prod_{i=1}^3 f_i(x_i) = \prod_{i=1}^3 \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\left(\frac{x_i - \mu_i}{\sigma_i}\right)^2} = \frac{1}{(2\pi)^{3/2} (\sigma_1 \sigma_2 \sigma_3)} e^{-\frac{1}{2} \sum_{i=1}^3 \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2}$$
$$-\infty < x_i < \infty, i=1, 2, 3$$

Example (3):

Let X_1, X_2, X_3 be stoch. indep. r.v.'s with $X_i \sim N(\mu, \sigma^2)$ $i=1, 2, 3$. Then the joint p.d.f of these r.v.'s is

$$f(x_1, x_2, x_3) = \prod_{i=1}^3 f(x_i) = \prod_{i=1}^3 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

$$= (2\pi)^{-3/2} \sigma^{-3} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^3 (x_i - \mu)^2}$$

$i=1, 2, 3, \quad -\infty < x_i < \infty$

Note: The r.v.'s X_1, X_2, X_3 in Example (2) are indep. but not represent a random sample while the r.v.'s X_1, X_2, X_3 in Example (3) represent a random sample of size 3 from $N(\mu, \sigma^2)$.

Example (4):

Let $X_1, X_2, \dots, X_n$ be a r.s. of size n from

- a) $Ber(p)$ b) $Geo(p)$ c) $P(\lambda)$ d) $U(a, b)$
 e) ~~$U(a, b)$~~ f) $Exp(\lambda)$ g) $Be(\alpha, \beta)$ h) ~~$U(a, b)$~~
 h) $N(0, 1)$ i) $N(\mu, \sigma^2)$

Write the joint p.d.f in each case

Sample Mean and Sample Variance

Defn:

Let $X_1, X_2, \dots, X_n$ be a r.s. of size n from any distn. (Disc. or Cont.), we define

1) The statistic $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ is called the sample mean

2) The statistic $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ is called the sample Variance

Transformation (change) Variables P.d.f Technique:-

Continuous Type:-

Let $X_1, X_2, \dots, X_n$ be n r.v.'s of cont. type defined on the n -dimensional space with joint p.d.f $f(x_1, x_2, \dots, x_n)$. Let the r.v.'s $Y_i = u_i(X_1, X_2, \dots, X_n)$, $i = 1, 2, \dots, K$ be functions of $X_1, X_2, \dots, X_n$ ($K \leq n$) suppose the joint p.d.f of $Y_1, Y_2, \dots, Y_K$ is required.

If $K < n$, we introduce an additional new r.v.'s $Y_{k+1} = u_{k+1}(X_1, X_2, \dots, X_n)$, $Y_{k+2} = u_{k+2}(X_1, X_2, \dots, X_n)$, $\dots$, $Y_n = u_n(X_1, X_2, \dots, X_n)$ so that the functions $y_i = u_i(x_1, x_2, \dots, x_n)$, $i = 1, 2, \dots, n$ define one-to-one transformation that maps the space A of $\{X_i\}$ onto the space B of $\{Y_i\}$ with inverse transforms $x_i = w_i(y_1, y_2, \dots, y_n)$, $i = 1, 2, \dots, n$ and the Jacobian $n \times n$ determinant of the 1st partial derivatives

$$J = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(y_1, y_2, \dots, y_n)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \dots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \dots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \dots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} \neq 0$$

Then the joint p.d.f of $Y_1, Y_2, \dots, Y_n$ is

$$g(y_1, y_2, \dots, y_n) = \int [w_1(y_1, y_2, \dots, y_n), w_2(y_1, y_2, \dots, y_n), \dots, w_n(y_1, y_2, \dots, y_n)] |J| \cdot f(x_1, x_2, \dots, x_n) dy_1 \dots dy_n$$

and the joint p.d.f of $Y_1, Y_2, \dots, Y_K$ is

$$g^*(y_1, y_2, \dots, y_K) = \int_{y_{k+2}} \int_{y_{k+1}} \dots \int_{y_n} g(y_1, y_2, \dots, y_n) dy_n \dots dy_{k+2} dy_{k+1}$$

Example (1):

Let the r.v X has p.d.f $f(x) = 3x^2$, $0 < x < 1$
 $= 0$, e.w

Determine the p.d.f of r.v $Y = X^3$ and evaluate $P_r(\frac{1}{2} < Y < \frac{3}{4})$

Sol:

~~Let the r.v X has p.d.f~~

The function $y = x^3$: $A = \{x: 0 < x < 1\} \rightarrow B = \{y: 0 < y < 1\}$
 with inverse $x = y^{1/3}$ and $J = \frac{dx}{dy} = \frac{1}{3} y^{-2/3}$. Then the
 p.d.f of Y is

$$g(y) = f(y^{1/3}) |J| = 3(y^{1/3})^2 \cdot \frac{1}{3} y^{-2/3} = 1$$

$$\therefore g(y) = 1, \quad 0 < y < 1$$

$$= 0, \quad \text{e.w}$$

Example (2):

Let the r.v $X \sim U(0, 1)$. Find and name the distr. of
 r.v $Y = -2 \ln X$.

Sol:

The function $y = -2 \ln x$: $A = \{x: 0 < x < 1\} \rightarrow B = \{y: 0 < y < \infty\}$
 with inverse $x = e^{-\frac{1}{2}y}$ and

$$J = \frac{dx}{dy} = -\frac{1}{2} e^{-\frac{1}{2}y}$$

$$g(y) = f(e^{-\frac{1}{2}y}) \cdot |J| = (1) \cdot \frac{1}{2} e^{-\frac{1}{2}y} = \frac{1}{2} e^{-\frac{y}{2}}, \quad 0 < y < \infty$$

$$= 0, \quad \text{e.w}$$

That is $Y \sim \text{Exp}(2)$

Example (3)

Let the r.v. X has p.d.f $f(x) = \frac{1}{\pi} (1+x^2)^{-1}$, $-\infty < x < \infty$
e.w

Find the p.d.f of r.v. $Y = X^2$

Sol:

The function $y = x^2$: $\mathcal{A} = \{x: -\infty < x < \infty\} \xrightarrow[\text{on to}]{\text{not (1-1)}}$
 $\mathcal{B} = \{y: 0 < y < \infty\}$. So we write the S-S $\mathcal{A}$ as a union of two mutually exclusive subspaces, say $\mathcal{A}_1$ and $\mathcal{A}_2$. That is $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$ and $\mathcal{A}_1 \cap \mathcal{A}_2 = \emptyset$ where
 $\mathcal{A}_1 = \{x: -\infty < x < 0\}$ and $\mathcal{A}_2 = \{x: 0 \leq x < \infty\}$
Now the function $y = x^2$ define (1-1) transformation that maps of $\mathcal{A}_1$ and $\mathcal{A}_2$ onto $\mathcal{B}$ with inverses

In $\mathcal{A}_1$

$$x = -\sqrt{y} \text{ and } J_1 = \frac{dx}{dy} = -\frac{1}{2\sqrt{y}}$$

In $\mathcal{A}_2$

$$x = \sqrt{y} \text{ and } J_2 = \frac{dx}{dy} = \frac{1}{2\sqrt{y}}$$

then the p.d.f of Y is

$$g(y) = f(\sqrt{y}) |J_1| + f(\sqrt{y}) |J_2|$$
$$= \frac{1}{\pi} (1+y)^{-1} \frac{1}{2\sqrt{y}} + \frac{1}{\pi} (1+y)^{-1} \frac{1}{2\sqrt{y}}$$

$$= \frac{1}{\pi} \frac{1}{(1+y)\sqrt{y}}, \quad 0 < y < \infty$$

= 0, e.w

Example (4):

Let X_1, X_2 be a.r.s of size 2 from $N(0, 1)$. Define the r.v's $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. Find the joint p.d.f of X_1 and Y_2 and show that they are sto. indep. Name the distn. of Y_1 and of Y_2 .

Sol:

The joint p.d.f of X_1 and X_2 is

$$f(x_1, x_2) = f(x_1) \cdot f(x_2) = \frac{1}{2\pi} \exp \left[-\frac{1}{2} (x_1^2 + x_2^2) \right], \quad -\infty < x_1, x_2 < \infty$$

The functions $y_1 = x_1 + x_2$, $y_2 = x_1 - x_2$: $\mathcal{A} = \{(x_1, x_2) : -\infty < x_i < \infty, i=1, 2\} \rightarrow \mathcal{B} = \{(y_1, y_2) : -\infty < y_i < \infty, i=1, 2\}$ with inverses $x_1 = \frac{1}{2}(y_1 + y_2)$, $x_2 = \frac{1}{2}(y_1 - y_2)$ and

$$J = \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2}$$

$$x_1^2 + x_2^2 = \frac{1}{4}(y_1 + y_2)^2 + \frac{1}{4}(y_1 - y_2)^2 = \frac{1}{4}[y_1^2 + 2y_1y_2 + y_2^2 + y_1^2 - 2y_1y_2 + y_2^2]$$

$= \frac{1}{4}(2y_1^2 + 2y_2^2) = \frac{1}{2}(y_1^2 + y_2^2)$. Then the joint p.d.f of Y_1 and Y_2 is

$$g(y_1, y_2) = f\left[\frac{1}{2}(y_1 + y_2), \frac{1}{2}(y_1 - y_2)\right] |J| = \frac{1}{4\pi} e^{-\frac{1}{4}(y_1^2 + y_2^2)}, \quad -\infty < y_i < \infty, i=1, 2$$

The marginal p.d.f of Y_1 is $g_1(y_1) = \int_{y_2} g(y_1, y_2) dy_2$

$$\begin{aligned} g_1(y_1) &= \frac{1}{4} e^{-\frac{1}{4}y_1^2} \int_{-\infty}^{\infty} e^{-\frac{1}{4}y_2^2} dy_2 \\ &= \frac{e^{-\frac{1}{4}y_1^2}}{4\pi} \sqrt{2\pi} \sqrt{2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} \sqrt{2}} e^{-\frac{1}{2} \frac{y_2^2}{2}} dy_2 \\ &= \frac{1}{\sqrt{2\pi} \sqrt{2}} e^{-\frac{1}{2} \frac{y_1^2}{2}}, \quad -\infty < y_1 < \infty \end{aligned}$$

That is $Y_1 \sim N(0, 2)$

Similarly the p.d.f of Y_2 is $g_2(y_2) = \frac{1}{\sqrt{2\pi} \sqrt{2}} e^{-\frac{1}{2} \frac{y_2^2}{2}}, \quad -\infty < y_2 < \infty$

$$g_1(y_1) \cdot g_2(y_2) = \frac{1}{\sqrt{2\pi} \sqrt{2}} e^{-\frac{1}{2} \frac{y_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi} \sqrt{2}} e^{-\frac{1}{2} \frac{y_2^2}{2}} = \frac{1}{4\pi} e^{-\frac{1}{4}(y_1^2 + y_2^2)} = g(y_1, y_2)$$

Therefore Y_1 and Y_2 are stochastically independent.

Example (6):

Let X_1 and X_2 be two indep. r.v's with $X_1 \sim N(0, 1)$ and $X_2 \sim \chi^2_{(2)}$. Find the distn. of r.v $Y = \frac{X_1}{\sqrt{\frac{X_2}{2}}}$

Sol:

$$X_1 \sim N(0, 1) \Rightarrow f_1(x_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2}, \quad -\infty < x_1 < \infty$$

$$X_2 \sim \chi^2_{(2)} \Rightarrow f_2(x_2) = \frac{1}{2} e^{-\frac{x_2}{2}}, \quad 0 < x_2 < \infty$$

Since X_1 and X_2 are stoch. indep., then the joint p.d.f of X_1 and X_2 is

$$f(x_1, x_2) = f_1(x_1) f_2(x_2) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}(x_1^2 + x_2)} \quad \begin{matrix} -\infty < x_1 < \infty \\ 0 < x_2 < \infty \end{matrix}$$

with $Y = \frac{X_1}{\sqrt{\frac{X_2}{2}}}$, set $Z = X_2$

The functions $y = \frac{x_1}{\sqrt{\frac{x_2}{2}}}$, $z = x_2$: $\mathcal{A} = \{(x_1, x_2) : -\infty < x_1 < \infty, 0 < x_2 < \infty\}$

$\longrightarrow \mathcal{B} = \{(y, z) : -\infty < y < \infty, 0 < z < \infty\}$ with inverses

$x_1 = \frac{y\sqrt{z}}{\sqrt{2}}$, $x_2 = z$ and

$$J = \frac{\partial(x_1, x_2)}{\partial(y, z)} = \begin{vmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{vmatrix} = \begin{vmatrix} \frac{\sqrt{z}}{\sqrt{2}} & -\frac{y}{2\sqrt{2}\sqrt{z}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{z}{2}}$$

Then the joint p.d.f of Y and Z is

$$\begin{aligned} g(y, z) &= f\left(\frac{y\sqrt{z}}{\sqrt{2}}, z\right) |J| \\ &= \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{zy^2}{2} + z\right)} \cdot \sqrt{\frac{z}{2}} \end{aligned}$$

$$= \frac{1}{4\sqrt{\pi}} z^{\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{y^2}{2}+1\right)z}, \quad -\infty < y < \infty, \quad 0 < z < \infty$$

$= 0$? e.w
 The marginal p.d.f of Y is

$$g_1(y) = \int_z g(y, z) dz = \frac{1}{4\sqrt{\pi}} \int_0^{\infty} z^{\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{y^2}{2}+1\right)z} dz$$

$$\text{Set } w = \left(\frac{y^2}{2}+1\right)z \Rightarrow dw = \left(\frac{y^2}{2}+1\right) dz$$

$$g_1(y) = \frac{1}{4\sqrt{\pi}} \int_0^{\infty} \frac{w^{\frac{1}{2}}}{\left(\frac{y^2}{2}+1\right)^{\frac{1}{2}}} e^{-\frac{w}{2}} \cdot \frac{dw}{\left(\frac{y^2}{2}+1\right)}$$

$$= \frac{1}{4\sqrt{\pi}} \left(\frac{y^2}{2}+1\right)^{-\frac{3}{2}} \int_0^{\infty} w^{\frac{1}{2}} e^{-\frac{w}{2}} dw$$

$$= \frac{1}{4\sqrt{\pi}} \left(\frac{y^2}{2}+1\right)^{-\frac{3}{2}} \Gamma\left(\frac{3}{2}\right) \cdot \frac{3}{2} \int_0^1 \frac{1}{\Gamma\left(\frac{3}{2}\right) 2^{\frac{3}{2}}} w^{\frac{3}{2}-1} e^{-\frac{w}{2}} dw$$

$$= \frac{1}{4\sqrt{\pi}} \left(\frac{y^2}{2}+1\right)^{-\frac{3}{2}} \frac{1}{2} \Gamma\left(\frac{1}{2}\right) \cdot 2\sqrt{2}$$

$$= \frac{1}{2\sqrt{2}} \left(\frac{y^2}{2}+1\right)^{-\frac{3}{2}}, \quad -\infty < y < \infty$$

$$= 0 \quad , \quad \text{e.w}$$

The t -distn. (Student-distn.)

In 1908 Gosset derive the equation of t -distn. for small samples. He published his work secretly under the name "student".

Theorem (Gosset Theorem):

Let X and Y be sto. indep. r.v's with $X \sim N(0, 1)$ and $Y \sim \chi^2(r)$. Then the r.v $T = \frac{X}{\sqrt{\frac{Y}{r}}}$ has p.d.f

$$g(t) = \frac{\Gamma\left(\frac{r+2}{2}\right)}{\sqrt{\pi r} \Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+2}{2}}, \quad -\infty < t < \infty, r > 0$$

and T is said to have a t -distn. with r degrees of freedom and denoted by $T \sim t(r)$

Proof:-

Given $X \sim N(0, 1)$ then the p.d.f of X is

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < \infty$$

also $Y \sim \chi^2(r)$ then the p.d.f of Y is

$$f_2(y) = \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} y^{\frac{r}{2}-1} e^{-\frac{y}{2}}, \quad 0 < y < \infty$$

Since X and Y are sto. indep. then the joint p.d.f of X and Y is

$$f(x, y) = f_1(x) f_2(y) = \frac{1}{\sqrt{2\pi} \Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} y^{\frac{r}{2}-1} e^{-\frac{1}{2}(x^2 + y)}, \quad \begin{matrix} -\infty < x < \infty \\ 0 < y < \infty \end{matrix}$$

We have the transformation $T = \frac{X}{\sqrt{\frac{Y}{r}}}$, Let $W = Y$

Now, the function $t = \frac{x}{\sqrt{\frac{y}{r}}}$ and $w = y$ define a one-to-one

transformation that maps the space $A = \{(x, y) : -\infty < x < \infty, 0 < y < \infty\}$ into the space $B = \{(t, w) : -\infty < t < \infty, 0 < w < \infty\}$ with inverse transform $x = t \frac{\sqrt{w}}{\sqrt{r}}$, $y = w$ and the Jacobian of transformation is

~~$$J = \frac{\partial(x, y)}{\partial(t, w)} = \begin{vmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{\sqrt{w}}{\sqrt{r}} & \frac{t}{2\sqrt{r}\sqrt{w}} \\ 0 & 1 \end{vmatrix} = \frac{\sqrt{w}}{\sqrt{r}}$$~~

$$J = \frac{\partial(x, y)}{\partial(t, w)} = \begin{vmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{\sqrt{w}}{\sqrt{r}} & \frac{t}{2\sqrt{r}\sqrt{w}} \\ 0 & 1 \end{vmatrix} = \frac{\sqrt{w}}{\sqrt{r}}$$

Then the joint p.d.f of T and W is given by

$$g(t, w) = f\left(t \frac{\sqrt{w}}{\sqrt{r}}, w\right) |J| = \frac{1}{\sqrt{2\pi} \Gamma\left(\frac{r}{2}\right) 2^{r/2}} w^{r/2-1} e^{-\frac{1}{2}\left(\frac{wt^2}{r} + w\right)} \frac{\sqrt{w}}{\sqrt{r}}$$

$$= \frac{1}{\sqrt{2\pi r} \Gamma\left(\frac{r}{2}\right) 2^{r/2}} w^{\frac{r+1}{2}-1} e^{-\frac{1}{2}\left(1 + \frac{t^2}{r}\right)w}, -\infty < t < \infty, 0 < w < \infty$$

$$= 0, \text{ e.w.}$$

The marginal p.d.f of T is

$$g_1(t) = \int_w g(t, w) dw = \frac{1}{\sqrt{2\pi r} \Gamma\left(\frac{r}{2}\right) 2^{r/2}} \int_{w=0}^{\infty} w^{\frac{r+1}{2}-1} e^{-\frac{1}{2}\left(1 + \frac{t^2}{r}\right)w} dw$$

$$\text{let } z = w\left(1 + \frac{t^2}{r}\right) \Rightarrow dz = \left(1 + \frac{t^2}{r}\right) dw$$

$$g_1(t) = \frac{1}{\sqrt{2\pi r} \Gamma\left(\frac{r}{2}\right) 2^{r/2}} \int_{z=0}^{\infty} \left(\frac{z}{1 + \frac{t^2}{r}}\right)^{\frac{r+1}{2}-1} e^{-\frac{z}{2}} \left(1 + \frac{t^2}{r}\right)^{-1} dz$$

$$= \frac{1}{\sqrt{2\pi r} \Gamma\left(\frac{r}{2}\right) 2^{r/2}} \frac{\Gamma\left(\frac{r+1}{2}\right) 2^{\frac{r+1}{2}}}{\left(1 + \frac{t^2}{r}\right)^{\frac{r+1}{2}-1+1}} \int_0^{\infty} \frac{1}{\Gamma\left(\frac{r+1}{2}\right) 2^{\frac{r+1}{2}}} z^{\frac{r+1}{2}-1} e^{-\frac{z}{2}} dz$$

$$= \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r} \Gamma\left(\frac{r}{2}\right)} \left(1 + \frac{t^2}{r}\right)^{-\frac{r+1}{2}}, -\infty < t < \infty$$

The c.d.f of T is

$G_1(t) = \Pr(T \leq t) = \int_{-\infty}^t g_1(s) ds$. This integral does not have a simple form.

Some approximate value of $\Pr(T \leq t)$ for selected values of t and r are given in table IV of Appendix B.

The m.g.f of t -distn does not exist. To find the mean and variance of T can be found as follows:

We have $T = \frac{X}{\sqrt{\frac{Y}{r}}}$ where X and Y are sto. indep. with $X \sim N(0, 1)$ and $Y \sim \chi^2(r)$

$$E(T) = E\left(X \cdot \frac{1}{\sqrt{\frac{Y}{r}}}\right) = E(X) \cdot E\left(\frac{1}{\sqrt{\frac{Y}{r}}}\right) = 0$$

$$E(T^2) = E\left(X^2 \cdot \frac{r}{Y}\right) = E(X^2) \cdot r E\left(\frac{1}{Y}\right) = r E\left(\frac{1}{Y}\right)$$

$$E\left(\frac{1}{Y}\right) = \int_0^{\infty} \frac{1}{y} \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{r/2}} y^{\frac{r}{2}-1} e^{-\frac{y}{2}} dy$$

$$= \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{r/2}} \frac{1}{\Gamma\left(\frac{r-2}{2}\right) 2^{\frac{r-2}{2}}} \int_0^{\infty} \frac{1}{\Gamma\left(\frac{r-2}{2}\right) 2^{\frac{r-2}{2}}} y^{\frac{r-2}{2}-1} e^{-\frac{y}{2}} dy$$

$$= \frac{\Gamma\left(\frac{r-2}{2}\right)}{2 \Gamma\left(\frac{r}{2}\right)} = \frac{\Gamma\left(\frac{r}{2}-1\right)}{2 \Gamma\left(\frac{r}{2}\right)} = \frac{\Gamma\left(\frac{r}{2}-1\right)}{2\left(\frac{r}{2}-1\right) \Gamma\left(\frac{r}{2}-1\right)} = \frac{1}{r-2}, \quad r > 2$$

$$\text{Var}(T) = E(T^2) - \{E(T)\}^2 = \frac{r}{r-2} - 0 = \frac{r}{r-2}, \quad r \neq 2$$

Example (1)

Let the r.v $X \sim t(10)$. Find $\Pr(|X| > 2.228)$

Sol:

$$\begin{aligned}\Pr(|X| > 2.228) &= 1 - \Pr(|X| \leq 2.228) \\ &= 1 - \Pr(-2.228 \leq X \leq 2.228) \\ &= 1 - [2\Pr(X \leq 2.228) - 1] \\ &= 2 - 2\Pr(X \leq 2.228) \\ &= 2[1 - \Pr(X \leq 2.228)] = 2(1 - 0.975) \\ &= 2(0.025) = 0.05\end{aligned}$$

The F-Distn. (Fisher Theorem)

Let X and Y be stoch. indep. r.v's with $X \sim \chi^2_{(r)}$ and $Y \sim \chi^2_{(s)}$. Then the r.v $F = \frac{X/r}{Y/s} = \frac{sX}{rY}$ has p.d.f

$$g(f) = \frac{\Gamma(\frac{r+s}{2}) (\frac{r}{s})^{r/2}}{\Gamma(\frac{r}{2}) \Gamma(\frac{s}{2})} \frac{f^{r/2-1}}{(1 + \frac{r}{s}f)^{\frac{r+s}{2}}}, \quad 0 < f < \infty$$

and F is said to have F-distn with (r, s) dof's and denoted by $F \sim F(r, s)$

Proof:

Since X and Y are indep. r.v's, then the joint p.d.f is

$$\begin{aligned}f(x, y) &= \frac{1}{\Gamma(\frac{r}{2}) 2^{r/2}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}} \cdot \frac{1}{\Gamma(\frac{s}{2}) 2^{s/2}} y^{\frac{s}{2}-1} e^{-\frac{y}{2}} \\ &= \frac{1}{\Gamma(\frac{r}{2}) \Gamma(\frac{s}{2}) 2^{\frac{r+s}{2}}} x^{\frac{r}{2}-1} y^{\frac{s}{2}-1} e^{-\frac{1}{2}(x+y)}, \quad 0 < x < \infty, 0 < y < \infty\end{aligned}$$

with $F = \frac{sX}{rY}$, Set $W = Y$

The functions $f = \frac{rX}{sY}$, $w = Y$ $\mathcal{A} = \{(x, y) : 0 < x < \infty, 0 < y < \infty\}$
 $\rightarrow \mathcal{B} = \{(f, w) : 0 < f < \infty, 0 < w < \infty\}$ with inverses

$$x = \frac{r}{s} f w, \quad y = w \text{ and}$$

$$J = \frac{\partial(x, y)}{\partial(f, w)} = \begin{vmatrix} \frac{\partial x}{\partial f} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial f} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{r}{s} w & \frac{r}{s} f \\ 0 & 1 \end{vmatrix} = \frac{r}{s} w$$

Then the joint p.d.f of F and W is

$$g(f, w) = f\left(\frac{r}{s} f w, w\right) |J| = \frac{1}{\Gamma\left(\frac{r}{2}\right) \Gamma\left(\frac{s}{2}\right) 2^{\frac{r+s}{2}}} \left(\frac{r}{s} f w\right)^{\frac{r}{2}-1} w^{\frac{s}{2}-1} e^{-\frac{1}{2}\left(\frac{r}{s} f + 1\right) w} \cdot \frac{r}{s} w, \quad 0 < f < \infty, \quad 0 < w < \infty$$

The marginal p.d.f of F is

$$g_1(f) = \int_w g(f, w) dw = \frac{\left(\frac{r}{s}\right)^{\frac{r}{2}} f^{\frac{r}{2}-1}}{\Gamma\left(\frac{r}{2}\right) \Gamma\left(\frac{s}{2}\right) 2^{\frac{r+s}{2}}} \int_{w=0}^{\infty} w^{\frac{r+s}{2}-1} e^{-\frac{1}{2}\left(\frac{r}{s} f + 1\right) w} dw$$

$$\text{Set } z = \left(\frac{r}{s} f + 1\right) w \Rightarrow dz = \left(\frac{r}{s} f + 1\right) dw$$

$$g_1(f) = \frac{\left(\frac{r}{s}\right)^{\frac{r}{2}} f^{\frac{r}{2}-1}}{\Gamma\left(\frac{r}{2}\right) \Gamma\left(\frac{s}{2}\right) 2^{\frac{r+s}{2}}} \int_{z=0}^{\infty} \left(\frac{z}{\frac{r}{s} f + 1}\right)^{\frac{r+s}{2}-1} e^{-\frac{z}{\left(\frac{r}{s} f + 1\right)}} \frac{dz}{\left(\frac{r}{s} f + 1\right)}$$

$$= \frac{\left(\frac{r}{s}\right)^{\frac{r}{2}} f^{\frac{r}{2}-1} \Gamma\left(\frac{r+s}{2}\right) 2^{\frac{r+s}{2}}}{\Gamma\left(\frac{r}{2}\right) \Gamma\left(\frac{s}{2}\right) 2^{\frac{r+s}{2}} \left(\frac{r}{s} f + 1\right)^{\frac{r+s}{2}}} \int_{z=0}^{\infty} z^{\frac{r+s}{2}-1} e^{-\frac{z}{\left(\frac{r}{s} f + 1\right)}} dz$$

$$= \frac{\Gamma\left(\frac{r+s}{2}\right) \left(\frac{r}{s}\right)^{\frac{r}{2}} f^{\frac{r}{2}-1}}{\Gamma\left(\frac{r}{2}\right) \Gamma\left(\frac{s}{2}\right) \left(1 + \frac{r}{s} f\right)^{\frac{r+s}{2}}}, \quad 0 < f < \infty$$

$$= 0, \quad e.w$$

The c.d.f of F is $G_1(f) = \int_0^f g(w) dw = \Pr(F \leq f)$. This integral does not have simple form. Some approximate value of $\Pr(F \leq f)$ are given in table V of appendix A-7 for selected value of r, s and f

Example (2)

If the r.v $X \sim F(r, s)$. What is the distn of r.v $Y = \frac{1}{X}$

Sol:

Since $X \sim F(r, s)$ that mean there are two indep. r.v's say U and V such that $U \sim \chi^2(r)$ and $V \sim \chi^2(s)$ and

$$X = \frac{U/r}{V/s} \sim F(r, s)$$

Now

$$Y = \frac{1}{X} = \frac{1}{\frac{U/r}{V/s}} = \frac{V/s}{U/r} \sim F(s, r)$$

Important Note:

If $X \sim F(r, s)$ then $Y = \frac{1}{X} \sim F(s, r)$

$$\Pr(X \leq a) = \Pr\left(\frac{1}{X} \geq \frac{1}{a}\right) = 1 - \Pr\left(\frac{1}{X} \leq \frac{1}{a}\right) = 1 - \Pr\left(Y \leq \frac{1}{a}\right)$$

Example (3)

Let the r.v $X \sim F(5, 10)$. Find the values of a and b so that $\Pr(X \leq a) = 0.05$ and $\Pr(X \leq b) = 0.95$

Sol:

$\Pr(X \leq b) = 0.95$. From F-distn. table with $r=5$, $s=10$ we have $b=3.33$

$$0.05 = \Pr(X \leq a) = \Pr\left(\frac{1}{X} \geq \frac{1}{a}\right) = 1 - \Pr\left(\frac{1}{X} \leq \frac{1}{a}\right) = 1 - \Pr\left(Y \leq \frac{1}{a}\right)$$

where $Y \sim F(10, 5) \Rightarrow \Pr\left(Y \leq \frac{1}{a}\right) = 1 - 0.05 = 0.95$

From F-distn table with $r=10$, $s=5$, we have $\frac{1}{a} = 0.474$
 $\Rightarrow a = 2.11$

The Distribution of $\bar{X}$ and $\frac{(n-1)S^2}{\sigma^2}$ (Normal Case)

Let $X_1, X_2, \dots, X_n$ be a r.s. of size n from $N(\mu, \sigma^2)$ and suppose that we wish to find the distn. of $\bar{X}$ and $\frac{(n-1)S^2}{\sigma^2}$

Distrn. of $\bar{X}$:

~~For~~ For Normal case, the sample mean $\bar{X} = \frac{1}{n} \sum X_i \sim N(\mu, \frac{\sigma^2}{n})$

Example (1)

Let $X_1, X_2, \dots, X_n$ be a r.s. of size $n=25$ from $N(75, 100)$. Find $\Pr(71 < \bar{X} < 79)$

Sol:

We have $n=25, \mu=75, \sigma^2=100$

each $X_i \sim N(75, 100) \Rightarrow \bar{X} \sim N(75, \frac{100}{25}) = N(75, 4)$

$$\Rightarrow Y = \frac{\bar{X} - 75}{2} \sim N(0, 1)$$

$$\begin{aligned} \Pr(71 < \bar{X} < 79) &= \Pr\left(\frac{71-75}{2} < \frac{\bar{X}-75}{2} < \frac{79-75}{2}\right) = \Pr(-2 < Y < 2) \\ &= 2\Pr(Y < 2) - 1 = 2(0.977) - 1 = 0.954 \end{aligned}$$

Distrn. of $\frac{(n-1)S^2}{\sigma^2}$:

$$\begin{aligned} \sum_{i=1}^n (X_i - \mu)^2 &= \sum_{i=1}^n [(X_i - \bar{X}) + (\bar{X} - \mu)]^2 = \sum_{i=1}^n [(X_i - \bar{X})^2 + 2(X_i - \bar{X})(\bar{X} - \mu) + (\bar{X} - \mu)^2] \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + 2(\bar{X} - \mu) \sum_{i=1}^n (X_i - \bar{X}) + n(\bar{X} - \mu)^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + 2(\bar{X} - \mu)(n\bar{X} - n\bar{X}) + n(\bar{X} - \mu)^2 \\ &= \sum_{i=1}^n (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2 = (n-1)S^2 + n(\bar{X} - \mu)^2 \end{aligned}$$

$$\Rightarrow (n-1)S^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2$$

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 - n\left(\frac{\bar{X} - \mu}{\sigma}\right)^2$$

(26)

Since each $X_i \sim N(\mu, \sigma^2) \Rightarrow \frac{X_i - \mu}{\sigma} \sim N(0, 1)$

$$\Rightarrow \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_{(1)}^2 \Rightarrow \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi_{(n)}^2$$

also $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$

$$\Rightarrow n\left(\frac{\bar{X} - \mu}{\sigma}\right)^2 \sim \chi_{(1)}^2$$

Note: For normal case only we shall accept that $\bar{X}$ and S^2 are sto. indep.

Then $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{(n-1)}^2$

Example (2)

Let $X_1, X_2, \dots, X_n$ be a r.s of size $n=6$ from $N(\mu, 10)$

Find $\Pr(2.3 < S^2 < 22.2)$

Sol: $Y = \frac{(n-1)S^2}{\sigma^2} = \frac{5S^2}{10} = \frac{S^2}{2} \sim \chi_{(5)}^2$

$$\begin{aligned} \Pr(2.3 < S^2 < 22.2) &= \Pr\left(\frac{2.3}{2} < \frac{S^2}{2} < \frac{22.2}{2}\right) \\ &= \Pr(1.15 < Y < 11.1) \\ &= \Pr(Y < 11.1) - \Pr(Y < 1.15) \\ &= 0.95 - 0.05 = 0.9 \end{aligned}$$

Mean and Variance of $\bar{X}$ and S^2 (Any distn.)

Let $X_1, X_2, \dots, X_n$ be a r.s of size n from any distn. (discrete or continuous) having mean μ and variance σ^2

Mean and Variance of $\bar{X}$

$$\bar{X} = \frac{1}{n} \sum X_i = \frac{1}{n} (X_1 + X_2 + \dots + X_n)$$

$$\begin{aligned} E(\bar{X}) &= \frac{1}{n} (EX_1 + EX_2 + \dots + EX_n) \\ &= \frac{1}{n} (\mu + \mu + \mu + \dots + \mu) = \frac{1}{n} (n\mu) = \mu \end{aligned}$$

$$\text{Var}(\bar{X}) = \frac{1}{n^2} [\text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n)]$$

$$= \frac{1}{n^2} (\sigma^2 + \sigma^2 + \dots + \sigma^2) = \frac{n}{n^2} \sigma^2 = \frac{1}{n} \sigma^2$$

Mean and variance of S^2

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

$$(n-1)E(S^2) = \sum_{i=1}^n E(X_i^2) - nE(\bar{X}^2) = \sum (\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)$$

$$= n(\sigma^2 + \mu^2) - \sigma^2 - n\mu^2 = n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2$$

$$= n\sigma^2 - \sigma^2 = (n-1)\sigma^2$$

So

$$E(S^2) = \sigma^2$$

We can show that

$$\text{Var}(S^2) = \frac{1}{n} \left(\mu_4 - \frac{n-3}{n-1} \sigma^4 \right) \text{ where } \mu_r = E[(X - \mu)^r]$$

For examples

1) If the sample from $G(\alpha, \beta)$, then $\mu = \alpha\beta$, $\sigma^2 = \alpha\beta^2$

$$\text{So } E(\bar{X}) = \mu = \alpha\beta \text{ and } \text{Var}(\bar{X}) = \frac{\sigma^2}{n} = \frac{\alpha\beta^2}{n}$$

and

$$E(S^2) = \sigma^2 = \alpha\beta^2$$

$$\text{Var}(S^2) = \frac{1}{n} \left(\mu_4 - \frac{n-3}{n-1} \sigma^4 \right) \text{ where } \mu_4 = E[(X - \mu)^4]$$

$$\mu_4 = E(X^4 - 4\mu X^3 + 6\mu^2 X^2 - 4\mu^3 X + \mu^4)$$

$$= E(X^4) - 4\mu E(X^3) + 6\mu^2 E(X^2) - 4\mu^3 E(X) + \mu^4$$

$$= E(X^4) - 4\mu E(X^3) + 6\mu^2 E(X^2) - 3\mu^4$$

Order statistics

الإحصاءات المرتبة

Def: let $x_1, x_2, \dots, x_n$ denote a random sample from distribution of the continuous type having p.d.f $f(x; \theta)$ which is positive provided $a < x < b$ let y_1 be the smaller of then x_i , y_2 the next x_i in order of magnitude, $\dots$, y_n the largest x_i .

that is $\underbrace{y_1}_{x_{\min}} < y_2 < \dots < \underbrace{y_n}_{x_{\max}}$

نوف (X) تكون دالة متغير عشوائي مستمر به دالة كثافة احتمالية $f(x; \theta)$ ودالة تجميعية $F(x)$ وليكن $(x_1, x_2, \dots, x_n)$ تمثل عينة عشوائية ذات حجم (n) مسحوبة من $(f(x; \theta))$ ولتكن $[x_1, x_2, \dots, x_n]$ تمثل مجموعة الإحصاءات المرتبة المشتقة من العائلة (X) كما سبيل المثال:

$$y_1 = x_{\min} = \min [x_1, x_2, \dots, x_n] = x_{(1)}$$

$$y_n = x_{\max} = \max [x_1, x_2, \dots, x_n] = x_{(n)}$$

Example 1: let $X_1, X_2, \dots, X_n$ be a.s. from a distribution with p.d.f. $f(x; \theta)$, let $Y_1 < Y_2 < Y_3 \dots < Y_n$ be the order statistic find the p.d.f. of Y_1 and that of Y_n , i.e. $g(Y_1), g(Y_n)$

Sol.

$$G(Y_1) = \Pr(Y_1 \leq y_1)$$

$$= 1 - \Pr(Y_1 > y_1)$$

$$= 1 - \Pr(X_{\min} > y_1)$$

or $X_{(1)}$

$$= 1 - \Pr(\text{all } x_i > y_1)$$

$$= 1 - \Pr(X_1 > y_1 \text{ and } X_2 > y_1, \dots, X_n > y_1)$$

$$= 1 - \prod_{i=1}^n \Pr(X_i > y_1)$$

$$= 1 - \prod_{i=1}^n (1 - \Pr(X_i \leq y_1))$$

$$= 1 - \prod_{i=1}^n (1 - F(y_1))$$

$$G(y_1) = 1 - (1 - F(y_1))^n$$

$$\therefore g(y_1) = G'(y_1) = -n(1 - F(y_1))^{n-1} \times -f(y_1)$$

(2)

$$G(y_n) = \Pr(Y_n \leq y_n)$$

$$= \Pr(X_{\max} \leq y_n) = \Pr(\text{all } X_i \leq y_n)$$

$$= \prod_{i=1}^n \Pr(X_i \leq y_n)$$

$$= \prod_{i=1}^n F(y_n) = (F(y_n))^n$$

$$\therefore g(y_n) = G'(y_n) = n (F(y_n))^{n-1} \cdot f(y_n)$$

The general formula for any order statistic y_r ; $r = 1, 2, \dots, n$ is given by

$$g(y_r; \theta) = \frac{n!}{(r-1)! (n-r)!} [F(y_r)]^{r-1} [1 - F(y_r)]^{n-r} f(y_r)$$

The general formula the joint p.d.f. of the order statistic (y_r) and (y_s) , $1 \leq r \leq s \leq n$ is given by :

$$g(y_r, y_s; \theta) = \frac{n!}{(r-1)! (s-r-1)! (n-s)!} \cdot [F(y_r)]^{r-1}$$

$$\cdot f(y_r) \cdot [F(y_s) - F(y_r)]^{s-r-1}$$

$$\cdot f(y_s) \cdot [1 - F(y_s)]^{n-s}$$

(3)

Example 2: Four measurements are taken on a random variable having the p.d.f.:-

$$f(x) = \frac{1}{3} \exp\left(-\frac{x}{3}\right), \quad x > 0$$

Find $\Pr(X_4 \leq 5.1)$ and $\Pr(Y_4 \leq 5.1)$ where Y_4 is a fourth order statistic (the largest order statistic)

Sol.

$$\Pr(X_4 \leq 5.1) = \int_0^{5.1} f(x_4) dx_4$$

$$\therefore f(x_4) = \frac{1}{3} \exp\left(-\frac{x_4}{3}\right)$$

$$\Rightarrow \Pr(X_4 \leq 5.1) = \int_0^{5.1} \frac{1}{3} e^{-\frac{x_4}{3}} dx_4$$

$$\Rightarrow -e^{-\frac{x_4}{3}} \Big|_0^{5.1} \Rightarrow -e^{-\frac{5.1}{3}} + 1 = 1 - e^{-\frac{5.1}{3}}$$

$$\Pr(Y_4 \leq 5.1) = \int_0^{5.1} g(y_4) dy_4$$

or $X_{(4)}$
or $X_{\max}$

$$g(y_r) = \frac{n!}{(r-1)!(n-r)!} [F(y_r)]^{r-1} \cdot [1 - F(y_r)]^{n-r} \cdot f(y_r)$$

$$g(y_4) = \frac{4!}{3! 0!} \left[1 - e^{-\frac{y_4}{3}} \right]^3 \cdot \underbrace{\left[1 - \left(1 - e^{-\frac{y_4}{3}} \right) \right]^0}_1 \cdot \left(\frac{1}{3} e^{-\frac{y_4}{3}} \right)$$

$$\Rightarrow g(y_4) = \frac{4}{3} e^{-\frac{y_4}{3}} \left[1 - e^{-\frac{y_4}{3}} \right]^3$$

$$\therefore \Pr(Y_4 \leq 5-1) = \frac{4}{3} \int_0^{5-1} e^{-\frac{y_4}{3}} \left[1 - e^{-\frac{y_4}{3}} \right]^3 dy_4$$

$$= \frac{4}{3} \int_0^{5-1} e^{-\frac{y_4}{3}} \left[1 - e^{-\frac{y_4}{3}} \right] \left[1 - e^{-\frac{y_4}{3}} \right]^2 dy_4$$

$$\Rightarrow \frac{4}{3} \int_0^{5-1} \left(e^{-\frac{y_4}{3}} - e^{-\frac{2}{3} y_4} \right) \left(1 - 2e^{-\frac{y_4}{3}} + e^{-\frac{2}{3} y_4} \right) dy_4$$

$$\Rightarrow \frac{4}{3} \int_0^{5-1} \left[e^{-\frac{y_4}{3}} - 2e^{-\frac{2}{3} y_4} + e^{-\frac{5}{3} y_4} - e^{-\frac{2}{3} y_4} + 2e^{-\frac{5}{3} y_4} - e^{-\frac{4}{3} y_4} \right] dy_4$$

$$\Rightarrow \frac{4}{3} \left[\int_0^{5-1} e^{-\frac{y_4}{3}} dy_4 - 3 \int_0^{5-1} e^{-\frac{2}{3} y_4} dy_4 + 3 \int_0^{5-1} e^{-\frac{5}{3} y_4} dy_4 - \int_0^{5-1} e^{-\frac{4}{3} y_4} dy_4 \right]$$

example 3: let $Y_1 < Y_2 < Y_3 < Y_4 < Y_5$ be the order statistic of a r.s. size 5 from the uniform distribution having p.d.f

$$f(x; \theta) = \frac{1}{\theta} \quad , \quad 0 \leq x \leq \theta \quad , \quad 0 < \theta < \infty$$

Find ① $E(2Y_3)$

② $g(Y_3 | Y_5)$

Sol:

① $E(2Y_3) = 2E(Y_3) = 2 \cdot \int \cdot Y_3 \cdot g(Y_3) dY_3$

$$\therefore g(Y_r) = \frac{n!}{(r-1)! (n-r)!} [F(Y_r)]^{r-1} [1-F(Y_r)]^{n-r} \cdot f(Y_r)$$

$$\therefore f(X; \theta) = \frac{1}{\theta} \quad , \quad F(X) = \frac{X}{\theta}$$

$$\therefore f(Y_r) = \frac{1}{\theta} \quad , \quad F(Y_r) = \frac{Y_r}{\theta}$$

$$\therefore g(Y_3) = \frac{5!}{2! 2!} \left[\frac{Y_3}{\theta} \right]^2 \left[1 - \frac{Y_3}{\theta} \right]^2 \cdot \frac{1}{\theta}$$

$$\therefore g(Y_3) = \begin{cases} \frac{30}{\theta^3} Y_3^2 \left[1 - \frac{Y_3}{\theta} \right]^2 & 0 \leq Y_3 \leq \theta \\ 0 & \text{o.w.} \end{cases}$$

$$E(y_3) = \frac{60}{\theta^3} \int_0^{\theta} y_3^3 \left[1 - \frac{y_3}{\theta} \right]^2 dy_3$$

$$= \frac{60}{\theta^3} \int_0^{\theta} y_3^3 \left[1 - 2 \frac{y_3}{\theta} + \frac{y_3^2}{\theta^2} \right] dy_3$$

$$\Rightarrow \frac{60}{\theta^3} \int_0^{\theta} \left[y_3^3 - \frac{2}{\theta} y_3^4 + \frac{1}{\theta^2} y_3^5 \right] dy_3$$

$$\Rightarrow \frac{60}{\theta^3} \left[\frac{y_3^4}{4} \Big|_0^{\theta} - \frac{2}{\theta} \frac{y_3^5}{5} \Big|_0^{\theta} + \frac{1}{\theta^2} \frac{y_3^6}{6} \Big|_0^{\theta} \right]$$

$$\Rightarrow \frac{60}{\theta^3} \left[\frac{\theta^4}{4} - \frac{2}{\theta} \frac{\theta^5}{5} + \frac{1}{\theta^2} \frac{\theta^6}{6} \right]$$

$$\Rightarrow \frac{60}{4} \theta - \frac{120}{5} \theta + \frac{60}{6} \theta$$

$$\Rightarrow \theta \left[\frac{60}{4} - \frac{120}{5} + \frac{60}{6} \right] = \theta$$

15 - 24 + 10

$$\textcircled{2} g(y_3 | y_5) = \frac{g(y_3, y_5)}{g(y_5)}$$

$$\therefore g(y_5) = \frac{5!}{4! \cdot 0!} [1 - F(y_5)]^0 [F(y_5)]^4 \cdot f(y_5)$$

$$g(y_5) = 5 \cdot \left[1 - \frac{y_5}{\theta} \right]^0 \left[\frac{y_5}{\theta} \right]^4 \cdot \frac{1}{\theta}$$

$$= \frac{5}{\theta^5} y_5^4$$

$$\therefore g(y_5) = \begin{cases} \frac{5}{\theta^5} y_5^4 & 0 < y_5 < \theta \\ 0 & \text{o.w.} \end{cases}$$

$$g(y_r, y_s) = \frac{n!}{(r-1)! (s-r-1)! (n-s)!} (F(y_r))^{r-1} f(y_r)$$

$$\cdot (F(y_s) - F(y_r))^{s-r-1} \cdot f(y_s) \cdot (1 - F(y_s))^{n-s}$$

$$\Rightarrow g(y_3, y_5) = \frac{5!}{2! 1! 0!} \cdot \left(\frac{y_3}{\theta} \right)^2 \cdot \frac{1}{\theta} \cdot \frac{1}{\theta}$$

$$\cdot \left(\frac{y_5}{\theta} - \frac{y_3}{\theta} \right)^1 \cdot \underbrace{\left(1 - \frac{y_5}{\theta} \right)^0}_1$$

$$= \frac{60}{\theta^5} y_3^2 (y_5 - y_3)$$

$$\therefore g(y_3, y_5) = \begin{cases} \frac{60}{\theta^5} y_3^2 (y_5 - y_3) & 0 < y_3 < y_5 < \theta \\ 0 & \text{o.w.} \end{cases}$$

$$g(y_3 | y_5) = \frac{g(y_3, y_5)}{g(y_5)}$$

$$\Rightarrow \frac{\frac{60}{0.5} y_3^2 (y_5 - y_3)}{\frac{5}{0.5} y_5^4} = g(y_3 | y_5)$$

H.w.: ① $E(y_3 | y_5)$ ② $V(y_5)$

H.w.: let $Y_1 < Y_2 < Y_3$ be the order statistic of a r.s. of size 3 from the uniform distribution, find the variance of $(\frac{4}{3} Y_3)$

Confidence Interval For The Mean μ

Normal Case

Let $X_1, X_2, \dots, X_n$ be a r.s of size n from $N(\mu, \sigma^2)$. Suppose we require c.I for the mean μ with prob. $1-\alpha$.

There are two cases :

Case (i) When σ^2 is known

For normal case, we know that the sample mean $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$
 $\Rightarrow Z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$. Hence, we can find from $N(0, 1)$ distn table two n.o.s, say $\pm z_{1-\frac{\alpha}{2}}$ such that $Pr(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}) = 1-\alpha$

Now

$$-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}} \Rightarrow -z_{1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} < z_{1-\frac{\alpha}{2}} \Rightarrow \bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$$

$\therefore$ The $100(1-\alpha)\%$ c.I for μ is $(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$

Note In practice, if $x_1, x_2, \dots, x_n$ are the observed values of r.v's $X_1, X_2, \dots, X_n$, then the $100(1-\alpha)\%$ c.I for μ is $(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$ where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Example (4) Given the sample 3.3, -0.3, -0.6, -0.9 from $N(\mu, 9)$. Find the 90% c.I for μ .

Solution we have $n=4$, $\sigma=3$, $\bar{x} = \frac{1}{4}(3.3 - 0.3 - 0.6 - 0.9) = 0.375$

$$1-\alpha = 0.9 \Rightarrow \alpha = 0.1 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow 1 - \frac{\alpha}{2} = 0.95$$

From $N(0, 1)$ distn table, we have $z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.645$

$\therefore$ The 90% c.I for μ is $(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$

$$\equiv (0.375 - (\frac{3}{2})(1.645), 0.375 + (\frac{3}{2})(1.645)) \equiv (-2.093, 2.843)$$

Example (5) An observed sample of size 40 from $N(\mu, 10)$ yield

$\bar{x} = 7.164$. Find the 80% c.I for μ .

Solution we have $n=40$, $\sigma = \sqrt{10}$, $\bar{x} = 7.164$

$$1-\alpha = 0.8 \Rightarrow \alpha = 0.2 \Rightarrow \frac{\alpha}{2} = 0.1 \Rightarrow 1 - \frac{\alpha}{2} = 0.9$$

(15)

From $N(0,1)$ distn table, we have $z_{1-\frac{\alpha}{2}} = z_{\frac{\alpha}{2}} = 1.282$

The 80% C.I for μ is $(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$

$$= (7.164 - \frac{\sqrt{10}}{\sqrt{40}} (1.282), 7.164 + \frac{\sqrt{10}}{\sqrt{40}} (1.282)) = (6.523, 7.805)$$

Case (ii) When σ^2 is unknown

For normal case, we know the following

$$(1) \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1)$$

$$(2) \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \quad (3) \bar{X} \text{ and } S^2 \text{ are sto. indep.}$$

$$\text{Then the r.v } T = \frac{\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}}{\sqrt{\frac{\frac{(n-1)S^2}{\sigma^2}}{n-1}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t_{(n-1)}$$

Hence, we can find from t -distn table with $(n-1)$ dof two no.s, say $\pm t_{1-\frac{\alpha}{2}}$ such that $Pr(-t_{1-\frac{\alpha}{2}} < T < t_{1-\frac{\alpha}{2}}) = 1-\alpha$

$$\text{Now } -t_{1-\frac{\alpha}{2}} < T < t_{1-\frac{\alpha}{2}} \Rightarrow -t_{1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{X} - \mu)}{S} < t_{1-\frac{\alpha}{2}} \Rightarrow \bar{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}} < \mu < \bar{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}$$

$\therefore$ The $100(1-\alpha)\%$ C.I for μ is $(\bar{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}, \bar{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}})$

Note In practice, if $x_1, x_2, \dots, x_n$ are the observed values of r.v's $X_1, X_2, \dots, X_n$, then the $100(1-\alpha)\%$ C.I for μ is $(\bar{x} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}, \bar{x} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}})$

$$\text{where } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \text{ and } S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right]$$

Example (6) Back to Example (4) and suppose that the sample 3.3, -0.3, -0.6, -0.9 from $N(\mu, \sigma^2)$. What is the 90% C.I for μ .

Solution We have $n=4$, σ unknown, $\bar{x} = 0.375$

$$S^2 = \frac{1}{3} \left[(3.3)^2 + (-0.3)^2 + (-0.6)^2 + (-0.9)^2 - (4)(0.375)^2 \right]$$

$$= \frac{1}{3} (12.15 - 0.5625) = 3.8625 \Rightarrow S = 1.965$$

(11)

$$1-\alpha = 0.9 \Rightarrow \alpha = 0.1 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow 1-\frac{\alpha}{2} = 0.95$$

From t-distn table with 3 dof, we have $t_{1-\frac{\alpha}{2}} = t_{0.95} = 2.353$

The 90% c.I for μ is $(\bar{x} - \frac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}}, \bar{x} + \frac{s}{\sqrt{n}} t_{1-\frac{\alpha}{2}})$

$$= (0.375 - \frac{1.962}{2}(2.353), 0.375 + \frac{1.962}{2}(2.353)) = (-1.937, 2.687)$$

Non Normal Case (Large Samples)

For (disc. or cont.) large samples size ($n \geq 30$) selected from non-normal distns, we can find (with help of C.L.T) an approximate c.I for μ because most distns has limiting normal distn as $n \rightarrow \infty$.

Let $X_1, X_2, \dots, X_n$ be a r.s of size n ($n \geq 30$) from any distn that has mean μ , variance σ^2 , and m.g.f $M(t)$ exist for $-h < t < h$ and suppose we require a c.I for μ with prob. $1-\alpha$.

There are two cases

Case (i) When σ^2 is known and $n \geq 30$ (n large)

According to C.L.T, we have the r.v $Z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \stackrel{\text{app}}{\sim} N(0,1)$
 so, we can find from $N(0,1)$ distn table two nos, say $\pm z_{1-\frac{\alpha}{2}}$ such that $Pr(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}) = 1-\alpha$.

Now

$$-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}} \Rightarrow z_{1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} < z_{1-\frac{\alpha}{2}} \Rightarrow \bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$$

$\therefore$ As given in the case (i) of normal case, the approximate $100(1-\alpha)\%$ c.I for μ is $(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$.

Example (7) A sample of size 36 from a distn that has mean μ , variance $\sigma^2 = 144$, and its m.g.f exist. This sample yields $\bar{x} = 67.53$. Find the approximate 95% c.I for μ .

Solution $n = 36$ (large), $\sigma = 12$, $\bar{x} = 67.53$, so we approximate to normal

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1-\frac{\alpha}{2} = 0.975$$

From $N(0,1)$ distn table, we have $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$

∴ The approximate 95% C.I for μ is $(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$
 $= (67.53 - \frac{12}{6}(1.96), 67.53 + \frac{12}{6}(1.96))$
 $= (63.61, 71.45)$

Case (ii) When σ^2 is unknown and $n \geq 30$ (n large)
 Again, according to C.L.T, the r.v $U = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \stackrel{\text{app.}}{\sim} N(0,1)$

$$\text{Since } S^2 \xrightarrow[\text{sto.}]{\text{Conv.}} \sigma^2 \Rightarrow \frac{S^2}{\sigma^2} \xrightarrow[\text{sto.}]{\text{Conv.}} 1 \Rightarrow V = \frac{S}{\sigma} \xrightarrow[\text{sto.}]{\text{Conv.}} \sqrt{1} = 1$$

Then the r.v $Z = \frac{U}{V}$ has limiting distn as that of U .

That is

$$Z = \frac{\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}}{\frac{S}{\sigma}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \stackrel{\text{app.}}{\sim} N(0,1)$$

So, we can find from $N(0,1)$ distn table two no.s, say $\pm z_{1-\frac{\alpha}{2}}$ such that
 $\Pr(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}) = 1 - \alpha$

∴ As given in the case (i) of normal case, the approximate 100(1- α)% C.I for μ is $(\bar{X} - \frac{S}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{X} + \frac{S}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$

Example (8) A sample of size 100 from a distn that has mean μ , variance σ^2 and its m.g.f exist, this sample observe $\bar{x} = 16$ and $s^2 = 5.76$. Determine a 90% C.I for μ .

Solution We have $n = 100$ (large), σ^2 unknown, $\bar{x} = 16$, $s^2 = 5.76 \Rightarrow s = 2.4$

So, we approximate to normal.

$$1 - \alpha = 0.9 \Rightarrow \alpha = 0.1 \Rightarrow \frac{\alpha}{2} = 0.05 \Rightarrow 1 - \frac{\alpha}{2} = 0.95$$

From $N(0,1)$ distn table, we have $z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.645$

∴ The approximate 90% C.I for μ is $(\bar{x} - \frac{s}{\sqrt{n}} z_{1-\frac{\alpha}{2}}, \bar{x} + \frac{s}{\sqrt{n}} z_{1-\frac{\alpha}{2}})$

$$= (16 - \frac{2.4}{10}(1.645), 16 + \frac{2.4}{10}(1.645))$$

$$= (15.605, 16.395)$$

Non-Normal Case (Small Samples)

For small samples size ($n < 30$) selected from non-normal distns, we cannot expect our degree of confidence to be accurate. However, with assumption that the sample is approximately normal, we can derive an approximate $100(1-\alpha)\%$ C.I. for μ when σ^2 is unknown.

The derivation exactly as given in case (ii) of normal case. That is, the approximate $100(1-\alpha)\%$ C.I. for μ when σ^2 is unknown is

$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}, \bar{X} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}} \right)$$

Example (9) The length of 10 basket ball players in cm are:- 182, 199, 201, 189, 192, 190, 203, 178, 197, 187. Find a 95% C.I. for μ . Assume approximate normal distn for the length.

Solution we have $n=10$ (n small), σ^2 unknown, the assumption for the length is approximately normal.

$$\bar{x} = \frac{1}{n} \sum x_i = \frac{1}{10} (182 + 199 + \dots + 187) = 191.4$$

$$s^2 = \frac{1}{n-1} [\sum x_i^2 - n\bar{x}^2] = \frac{1}{9} [(182)^2 + (199)^2 + \dots + (187)^2 - 10(191.4)^2]$$

$$= 67.733 \Rightarrow s = 8.23$$

$$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025 \Rightarrow 1-\frac{\alpha}{2} = 0.975$$

From t-distn table with $n-1 = 10-1 = 9$ dof, we have $t_{1-\frac{\alpha}{2}} = t_{0.975} = 2.262$

So, the approximate 95% C.I. for μ is $\left(\bar{x} - \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}, \bar{x} + \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}} \right)$

$$= \left(191.4 - \frac{8.23}{\sqrt{10}} (2.262), 191.4 + \frac{8.23}{\sqrt{10}} (2.262) \right)$$

$$= (185.9, 203.6)$$

(14)

Summary For C.I of μ

Normal	Non-Normal approximated to Normal
$\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$, if σ^2 is known	$\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$, if σ^2 is known and $n \geq 30$
$\bar{X} \pm \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}$, if σ^2 is unknown	$\bar{X} \pm \frac{S}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$, if σ^2 is unknown and $n \geq 30$
$\bar{X} \pm \frac{S}{\sqrt{n}} t_{1-\frac{\alpha}{2}}$, if σ^2 is unknown and $n < 30$ with the assumption that the sample is approximately normal	

Confidence Intervals For Difference of Means

Normal Case

Let $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ be two independent random samples of sizes n and m from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively.

Suppose we require c.I for $\mu_1 - \mu_2$ with prob. $1 - \alpha$.

There are two cases :-

Case (i) when σ_1^2 and σ_2^2 are known.

For normal case we know that $\bar{X} \sim N(\mu_1, \frac{\sigma_1^2}{n})$ and $\bar{Y} \sim N(\mu_2, \frac{\sigma_2^2}{m})$
 $\Rightarrow \bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m})$. Then the r.v

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0, 1)$$

Hence, we can find from $N(0, 1)$ distn table two no's, say $\pm z_{1-\frac{\alpha}{2}}$ such that $Pr(-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}}) = 1 - \alpha$

Now

$$-z_{1-\frac{\alpha}{2}} < Z < z_{1-\frac{\alpha}{2}} \equiv -z_{1-\frac{\alpha}{2}} < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} < z_{1-\frac{\alpha}{2}}$$

$$\equiv (\bar{X} - \bar{Y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$$

$\therefore$ The $100(1-\alpha)\%$ c.I for $\mu_1 - \mu_2$ is

$$\left((\bar{X} - \bar{Y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, (\bar{X} - \bar{Y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right)$$

Example (10) Two independent samples of sizes 100 and 72 from $N(\mu_1, 25)$ and $N(\mu_2, 36)$ respectively. These sample yields $\bar{x} = 80$ and $\bar{y} = 75$. Find a 90% c.I for $\mu_1 - \mu_2$

Solution we have $n = 100$, $m = 72$, $\sigma_1^2 = 25$, $\sigma_2^2 = 36$, $\bar{x} = 80$, $\bar{y} = 75$

$1 - \alpha = 0.9 \Rightarrow 1 - \frac{\alpha}{2} = 0.95$. From $N(0, 1)$ distn table, we have

$z_{1-\frac{\alpha}{2}} = z_{0.95} = 1.645$. The 90% c.I for $\mu_1 - \mu_2$ is

$$\left((\bar{x} - \bar{y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, (\bar{x} - \bar{y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right)$$

(17)

$$= ((80-75) - (1.645) \sqrt{\frac{25}{100} + \frac{36}{72}}, (80-75) + (1.645) \sqrt{\frac{25}{100} + \frac{36}{72}})$$

$$= (5 - (1.645) \frac{\sqrt{3}}{2}, 5 + (1.645) \frac{\sqrt{3}}{2}) = (3.575, 6.425)$$

Case (ii) When σ_1^2 and σ_2^2 are unknown but $\sigma_1^2 = \sigma_2^2 = \sigma^2$

For normal case, we know the following

- (1) $\bar{X} \sim N(\mu_1, \frac{\sigma^2}{n})$ and $\bar{Y} \sim N(\mu_2, \frac{\sigma^2}{m})$
- (2) $\frac{(n-1)S_1^2}{\sigma^2} \sim \chi^2(n-1)$ and $\frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(m-1)$
- (3) $\bar{X}$ and S_1^2 are st. indep and $\bar{Y}$ and S_2^2 are st. indep

Now

$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n} + \frac{\sigma^2}{m}) \Rightarrow \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0, 1)$$

Also $\frac{(n-1)S_1^2}{\sigma^2} + \frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(n+m-2)$. Then the t-v

$$T = \frac{\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}}}{\sqrt{\frac{\frac{(n-1)S_1^2 + (m-1)S_2^2}{\sigma^2}}{n+m-2}}} = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\frac{1}{n} + \frac{1}{m}) \cdot \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}} \sim t(n+m-2)$$

Hence from t-distn table with $(n+m-2)$ dof, we can find two nos $\pm t_{1-\frac{\alpha}{2}}$ such that $P_r(-t_{1-\frac{\alpha}{2}} < T < t_{1-\frac{\alpha}{2}}) = 1 - \alpha$

Now

$$-t_{1-\frac{\alpha}{2}} < T < t_{1-\frac{\alpha}{2}} \Rightarrow -t_{1-\frac{\alpha}{2}} < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\frac{1}{n} + \frac{1}{m}) \cdot \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}} < t_{1-\frac{\alpha}{2}}$$

$$\Rightarrow (\bar{X} - \bar{Y}) - t_{1-\frac{\alpha}{2}} \sqrt{(\frac{1}{n} + \frac{1}{m}) \cdot \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}} < \mu_1 - \mu_2 < (\bar{X} - \bar{Y}) + t_{1-\frac{\alpha}{2}} \sqrt{(\frac{1}{n} + \frac{1}{m}) \cdot \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}$$

$\therefore$ The $100(1-\alpha)\%$ C.I for $\mu_1 - \mu_2$ is

$$\left((\bar{X} - \bar{Y}) - t_{1-\frac{\alpha}{2}} \sqrt{(\frac{1}{n} + \frac{1}{m}) \cdot \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}, (\bar{X} - \bar{Y}) + t_{1-\frac{\alpha}{2}} \sqrt{(\frac{1}{n} + \frac{1}{m}) \cdot \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}} \right)$$

As given in case (i) of normal case, the approximate $100(1-\alpha)\%$ C.I for $\mu_1 - \mu_2$ is

$$\left((\bar{X} - \bar{Y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, (\bar{X} - \bar{Y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right)$$

Example (12) A r.s of size $n=100$ taken from a distn whose standard deviation $\sigma_1=5$ yield $\bar{x}=80$. A second r.s of size $m=75$ taken from a different distn whose standard deviation $\sigma_2=3$ yield $\bar{y}=72$. Find an approximate 90% C.I for $\mu_1 - \mu_2$

Solution

$n=100, m=75$ (n and m are large), $\sigma_1^2=25, \sigma_2^2=9, \bar{x}=80, \bar{y}=72$
So, we approximate to normal

$1-\alpha=0.9 \Rightarrow 1-\frac{\alpha}{2}=0.95$. From $N(0,1)$ table, we have $z_{1-\frac{\alpha}{2}}=z_{0.95}=1.645$

$\therefore$ The approximate 90% C.I for $\mu_1 - \mu_2$ is

$$\left((\bar{x} - \bar{y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, (\bar{x} - \bar{y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right)$$

$$\equiv \left((80 - 72) - (1.645) \sqrt{\frac{25}{100} + \frac{9}{75}}, (80 - 72) + (1.645) \sqrt{\frac{25}{100} + \frac{9}{75}} \right)$$

$$\equiv \left(8 - (1.645) \sqrt{0.37}, 8 + (1.645) \sqrt{0.37} \right) \equiv (6.999, 9.001)$$

Case (ii) When σ_1^2 and σ_2^2 are unknown and $n \geq 30, m \geq 30$

According to C.L.T, the r.v

$$U = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \xrightarrow{\text{app}} N(0,1)$$

$$\text{since } \sum_1^2 \xrightarrow[\text{sto}]{\text{conv}} \sigma_1^2 \Rightarrow \frac{\sum_1^2}{n} \xrightarrow{\text{conv}} \frac{\sigma_1^2}{n} \text{ and } \sum_2^2 \xrightarrow[\text{sto}]{\text{conv}} \sigma_2^2 \Rightarrow \frac{\sum_2^2}{m} \xrightarrow{\text{conv}} \frac{\sigma_2^2}{m}$$

$$\Rightarrow \frac{\sum_1^2}{n} + \frac{\sum_2^2}{m} \xrightarrow[\text{sto}]{\text{conv}} \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m} \Rightarrow \frac{\frac{\sum_1^2}{n} + \frac{\sum_2^2}{m}}{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \xrightarrow[\text{sto}]{\text{conv}} 1$$

and the r.v

$$V = \sqrt{\frac{\frac{\sum_1^2}{n} + \frac{\sum_2^2}{m}}{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \xrightarrow[\text{sto}]{\text{conv}} \sqrt{1} = 1$$

$\therefore$ The r.v $Z = \frac{U}{V}$ has limiting as that of U . That is

Example (11) A r.s of size 10 from $N(\mu_1, \sigma^2)$ yield $\bar{x} = 4.2$ and $s_1^2 = 49$. Another r.s (indep. of the 1st sample) of size 7 from $N(\mu_2, \sigma^2)$ yield $\bar{y} = 3.4$ and $s_2^2 = 32$. Find a 90% C.I for $\mu_1 - \mu_2$.

Solution

$$n=10, m=7, \sigma_1^2 = \sigma_2^2 = \sigma^2, \bar{x} = 4.2, s_1^2 = 49, \bar{y} = 3.4, s_2^2 = 32.$$

$1 - \alpha = 0.9 \Rightarrow 1 - \frac{\alpha}{2} = 0.95$. From t-distn table with $n+m-2 = 15$ df, we have $t_{\frac{\alpha}{2}} \equiv t_{0.95} = 1.753$

$\therefore$ The 90% C.I for $\mu_1 - \mu_2$ is

$$\begin{aligned} & \left((\bar{x} - \bar{y}) - t_{\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \cdot \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}}, (\bar{x} - \bar{y}) + t_{\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \cdot \frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}} \right) \\ & \equiv \left((4.2 - 3.4) - (1.753) \sqrt{\left(\frac{1}{10} + \frac{1}{7}\right) \cdot \frac{(9)(49) + (6)(32)}{15}}, (4.2 - 3.4) + (1.753) \sqrt{\left(\frac{1}{10} + \frac{1}{7}\right) \cdot \frac{(9)(49) + (6)(32)}{15}} \right) \\ & \equiv \left(0.8 - (1.753) \sqrt{\frac{17}{70}(42.2)}, 0.8 + (1.753) \sqrt{\frac{17}{70}(42.2)} \right) = (-4.812, 6.412) \end{aligned}$$

Non-Normal Case (Large samples), $n \geq 30, m \geq 30$

If two independent large samples $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ of sizes n and m are selected from non-normal distns with means μ_1 and μ_2 , variances σ_1^2 and σ_2^2 and the m.g.f.s are exist, we can find with help of C.L.T an approximate C.I for $\mu_1 - \mu_2$ with prob. $1 - \alpha$.

There are two cases :-

Case (i) When σ_1^2 and σ_2^2 are known and $n \geq 30, m \geq 30$. According to C.L.T, the r.v.s

$$\frac{\sqrt{n}(\bar{X} - \mu_1)}{\sigma_1} \stackrel{\text{app}}{\sim} N(0,1) \Rightarrow \bar{X} \stackrel{\text{app}}{\sim} N\left(\mu_1, \frac{\sigma_1^2}{n}\right)$$

and

$$\frac{\sqrt{m}(\bar{Y} - \mu_2)}{\sigma_2} \stackrel{\text{app}}{\sim} N(0,1) \Rightarrow \bar{Y} \stackrel{\text{app}}{\sim} N\left(\mu_2, \frac{\sigma_2^2}{m}\right)$$

$$\Rightarrow \bar{X} - \bar{Y} \stackrel{\text{app}}{\sim} N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right). \text{ Then the r.v.}$$

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \stackrel{\text{app}}{\sim} N(0,1)$$

(20)

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\left(\frac{\frac{s_1^2}{n} + \frac{s_2^2}{m}}{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right)^{1/2}} = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}}} \stackrel{\text{app}}{\sim} N(0,1)$$

As given in case (i) of normal case, the approximate $100(1-\alpha)\%$ C.I for $\mu_1 - \mu_2$ is

$$\left((\bar{X} - \bar{Y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}}, (\bar{X} - \bar{Y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}} \right)$$

Example (13) A maths test was given to 75 boys and 50 girls. The boys made an average grade of 82 with standard deviation of 8, while the girls made an average grade of 76 with standard deviation of 6. Find a 96% C.I for $\mu_1 - \mu_2$ where μ_1 is the mean score of all boys and μ_2 is the mean score of all girls.

Solution

$n=75, m=50$ (n and m are large), σ_1^2 and σ_2^2 are unknown, $\bar{x}=82, \bar{y}=76, s_1=8, s_2=6$, we approximate to normal $1-\alpha=0.96 \Rightarrow 1-\frac{\alpha}{2}=0.98$. From $N(0,1)$ table, we have $z_{1-\frac{\alpha}{2}} = z_{0.98} = 2.05$

$\therefore$ The approximate 96% C.I for $\mu_1 - \mu_2$ is

$$\begin{aligned} & \left((\bar{x} - \bar{y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}}, (\bar{x} - \bar{y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}} \right) \\ &= \left((82-76) - (2.05) \sqrt{\frac{64}{75} + \frac{36}{50}}, (82-76) + (2.05) \sqrt{\frac{64}{75} + \frac{36}{50}} \right) \\ &= \left(6 - (2.05)(2.57)^{1/2}, 6 + (2.05)(2.57)^{1/2} \right) = (2.714, 9.286) \end{aligned}$$

Non-Normal (Small Samples)

For small samples sizes ($n < 30, m < 30$) selected from non-normal distns with assumption that the samples are approximately normal, The derivation of an approximate $100(1-\alpha)\%$ C.I for $\mu_1 - \mu_2$ is exactly

(21)

as given in case (ii) for normal case when the unknowns $\sigma_1^2 = \sigma_2^2 = \sigma^2$.
That is, the approximate $100(1-\alpha)\%$ C.I for $\mu_1 - \mu_2$ is

$$\left((\bar{X} - \bar{Y}) - t_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \left(\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}\right)}, (\bar{X} - \bar{Y}) + t_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \left(\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}\right)} \right)$$

Example (14) Records of the past 15 years have shown that the average rainfall in a certain region for the month May is 4.93 cm with standard deviation of 1.14 cm. A second region has had an average rainfall in May 2.64 cm with standard deviation of 0.66 cm during the past 10 years. Find a 95% C.I for the different of the true average rainfalls in these two regions. Assume that the observations come from an approximate two independent normal distn with equal variances.

Solution

$n=15, m=10$ ($n < 30, m < 30$), $\sigma_1^2 = \sigma_2^2 = \sigma^2$ unknown, $\bar{x} = 4.93, \bar{y} = 2.64$
 $S_1 = 1.14, S_2 = 0.66$, the assumption is approximately normal.

$1-\alpha = 0.95 \Rightarrow 1-\frac{\alpha}{2} = 0.975$. From t -distn table with $n+m-2 = 15+10-2 = 23$ dof, we have $t_{1-\frac{\alpha}{2}} = t_{0.975} = 2.069$

$\therefore$ The approximate 95% C.I for $\mu_1 - \mu_2$ is

$$\left((\bar{x} - \bar{y}) - t_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \left(\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}\right)}, (\bar{x} - \bar{y}) + t_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \left(\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}\right)} \right)$$

$$\equiv (4.93 - 2.64) \pm (2.069) \sqrt{\left(\frac{1}{15} + \frac{1}{10}\right) \left(\frac{(14)(1.14)^2 + 9(0.66)^2}{23}\right)}$$

$$\equiv 2.29 \pm (2.069)(0.4) \equiv (1.89, 2.69)$$

Summary of C.I For $\mu_1 - \mu_2$

Normal

$$(\bar{X} - \bar{Y}) \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \text{ For known } \sigma_1^2 \text{ and } \sigma_2^2$$

$$(\bar{X} - \bar{Y}) \pm t_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \left(\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}\right)}, \text{ For unknown } \sigma_1^2 \text{ and } \sigma_2^2$$

with $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Non-Normal Approximated To Normal (Large Samples)

$$(\bar{X} - \bar{Y}) \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \text{ For known } \sigma_1^2, \sigma_2^2 \text{ and } n \geq 30, m \geq 30$$

$$(\bar{X} - \bar{Y}) \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n} + \frac{s_2^2}{m}}, \text{ For unknown } \sigma_1^2, \sigma_2^2 \text{ and } n \geq 30, m \geq 30$$

Non-Normal (Small Samples)

$$(\bar{X} - \bar{Y}) \pm t_{1-\frac{\alpha}{2}} \sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) \left(\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}\right)}$$

For unknown σ_1^2, σ_2^2 and $n < 30, m < 30$
 with $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and the assumption
 that the samples are approximately normal

Confidence Interval For the Variance σ^2

Normal Case

Let $X_1, X_2, \dots, X_n$ be a r.s of size n from $N(\mu, \sigma^2)$. Suppose, we wish to find C.I for the variance σ^2 with prob $1-\alpha$.

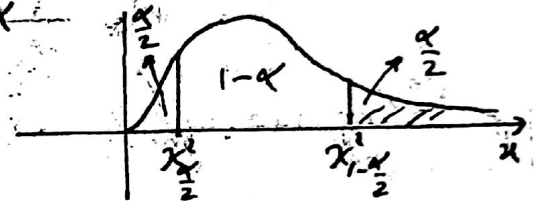
There are two cases :-

Case (i). When μ is known

For $i=1, 2, \dots, n$, we have $X_i \sim N(\mu, \sigma^2) \Rightarrow \frac{X_i - \mu}{\sigma} \sim N(0, 1)$
 $\Rightarrow \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2_{(1)} \Rightarrow Y = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2_{(n)}$

Hence, we can find from χ^2 -distn table with n dof two numbers, say, $\chi^2_{\frac{\alpha}{2}}$ and $\chi^2_{1-\frac{\alpha}{2}}$ such that $P_r(\chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}}) = 1-\alpha$

$$\text{Now } \chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}} = \chi^2_{\frac{\alpha}{2}} < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}$$



$$\begin{aligned} &= \frac{1}{\chi^2_{1-\frac{\alpha}{2}}} < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} < \frac{1}{\chi^2_{\frac{\alpha}{2}}} \\ &= \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{\frac{\alpha}{2}}} \end{aligned}$$

$\therefore$ The $100(1-\alpha)\%$ C.I for σ^2 is $\left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi^2_{\frac{\alpha}{2}}} \right)$

EX(15)

A r.s of size 10 from $N(0, \sigma^2)$ yield $\sum_{i=1}^{10} x_i^2 = 106.6$. Find a 95% C.I for σ^2

Solution $n=10$, $\mu=0$ (known), $\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^{10} (x_i - 0)^2 = \sum_{i=1}^{10} x_i^2 = 106.6$

$1-\alpha = 0.95 \Rightarrow 1 - \frac{\alpha}{2} = 0.975 \Rightarrow \frac{\alpha}{2} = 0.025$. From χ^2 -distn table with $n=10$ dof, we have

$$\chi^2_{\frac{\alpha}{2}} = \chi^2_{0.025} = 3.25 \text{ and } \chi^2_{1-\frac{\alpha}{2}} = \chi^2_{0.975} = 20.5$$

(25)

$$\therefore \text{The } 95\% \text{ C-I for } \sigma^2 \text{ is } \left(\frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^n (x_i - \mu)^2}{\chi^2_{\frac{\alpha}{2}}} \right)$$

$$\equiv \left(\frac{106.6}{20.5}, \frac{106}{3.25} \right) \equiv (5.2, 32.5)$$

Case (ii) when μ is unknown

For normal case, we know that the r.v. $Y = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$
 So, we can find from χ^2 -distrn with $n-1$ dof two numbers, say, $\chi^2_{\frac{\alpha}{2}}$ and $\chi^2_{1-\frac{\alpha}{2}}$ such that $P(\chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}}) = 1-\alpha$

$$\text{Now } \chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}} \equiv \chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}} \equiv \chi^2_{\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}$$

$$\equiv \frac{1}{\chi^2_{1-\frac{\alpha}{2}}} < \frac{\sigma^2}{(n-1)S^2} < \frac{1}{\chi^2_{\frac{\alpha}{2}}} \equiv \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}$$

$$\therefore \text{The } 100(1-\alpha)\% \text{ C-I for } \sigma^2 \text{ is } \left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} \right)$$

Ex(16) The volumes of 10 cans distributed by a certain company are:
 46.4, 46.1, 45.8, 47.0, 46.1, 45.9, 45.8, 46.9, 45.2, 46.0.

Find a 95% C-I for the variance of all such cans if the volume distrn is $N(\mu, \sigma^2)$.

Solution : $n=10$, μ is unknown, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{10}(46.4+46.1+\dots+46.0) = 46.1$

$$\sum_{i=1}^n x_i^2 = (46.4)^2 + (46.1)^2 + \dots + (46.0)^2 = 21273.13$$

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right] = \frac{1}{9} [21273.13 - (10)(46.1)^2] = 0.286$$

$$(n-1)S^2 = 2.576$$

$1-\alpha = 0.95 \Rightarrow 1-\frac{\alpha}{2} = 0.975 \Rightarrow \frac{\alpha}{2} = 0.025$. From χ^2 -distrn table with
 $n-1 = 10-1 = 9$ dof, we have $\chi^2_{\frac{\alpha}{2}} = \chi^2_{0.025} = 2.7$ and $\chi^2_{1-\frac{\alpha}{2}} = \chi^2_{0.975} = 19.023$

$$\therefore \text{The } 95\% \text{ C-I for } \sigma^2 \text{ is } \left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} \right) \equiv \left(\frac{2.576}{19.023}, \frac{2.576}{2.7} \right)$$

$$\equiv (0.135, 0.954)$$

Confidence Interval For the Ratio of two variances

Normal Case

Let $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ be two independent random samples from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. Suppose we wish to find a C.I for the ratio $\frac{\sigma_2^2}{\sigma_1^2}$ with prob. $1-\alpha$.

For Normal case, we have $\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi_{(n-1)}^2$ and $\frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi_{(m-1)}^2$ where S_1^2 and S_2^2 are the samples variance.

Then the r.v $F = \frac{\frac{(n-1)S_1^2}{\sigma_1^2}}{\frac{(m-1)S_2^2}{\sigma_2^2}} = \frac{\sigma_2^2}{\sigma_1^2} \frac{S_1^2}{S_2^2} \sim F(n-1, m-1)$

So, we can find from F-distrib table with $(n-1, m-1)$ dof two number, say, $f_{\alpha/2}, f_{1-\alpha/2}$ such that $P(f_{\alpha/2} < F < f_{1-\alpha/2}) = 1-\alpha$

$$\begin{aligned} \text{Now } f_{\alpha/2} < F < f_{1-\alpha/2} &\equiv f_{\alpha/2} < \frac{\sigma_2^2}{\sigma_1^2} \frac{S_1^2}{S_2^2} < f_{1-\alpha/2} \\ &\equiv \frac{S_2^2}{S_1^2} f_{\alpha/2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2} f_{1-\alpha/2} \end{aligned}$$



$\therefore$ The $100(1-\alpha)\%$ C.I for $\frac{\sigma_2^2}{\sigma_1^2}$ is $\left(\frac{S_2^2}{S_1^2} f_{\alpha/2}, \frac{S_2^2}{S_1^2} f_{1-\alpha/2} \right)$

Remark $f_{\alpha/2}^{(n,m)} = \frac{1}{f_{1-\alpha/2}^{(m,n)}}$

EX(17) Two independent samples of sizes $n=16, m=10$ from two independent $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ distrib yields $S_1^2=4.14, S_2^2=7.25$. Find a 90% C.I for $\frac{\sigma_2^2}{\sigma_1^2}$

Solution

$$n=16, m=10, S_1^2=4.14, S_2^2=7.25$$

$$1-\alpha=0.90 \Rightarrow 1-\frac{\alpha}{2}=0.95 \Rightarrow \frac{\alpha}{2}=0.05$$

From F-distrib table with $(n-1, m-1)=(15, 9)$ dof, we have

(27)

$$f_{1-\frac{\alpha}{2}}^{(15,9)} = f_{0.95}^{(15,9)} = 3.01, \quad f_{\frac{\alpha}{2}}^{(15,9)} = f_{0.05}^{(15,9)} = \frac{1}{f_{0.95}^{(9,15)}} = \frac{1}{2.59} = 0.386$$

$$\therefore \text{The } 90\% \text{ C.I. for } \frac{\sigma^2}{\sigma_0^2} \text{ is } \left(\frac{s^2}{s_0^2} f_{\frac{\alpha}{2}}, \frac{s^2}{s_0^2} f_{1-\frac{\alpha}{2}} \right) \\ = \left(\frac{7.26}{4.14} (0.386), \frac{7.26}{4.14} (3.01) \right) = (0.677, 5.28)$$

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339483207>

الاحصائية الجداول -

Chapter · February 2020

CITATIONS

0

READS

3,069

1 author:



Ibrahim Alali

Tishreen University

20 PUBLICATIONS 0 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



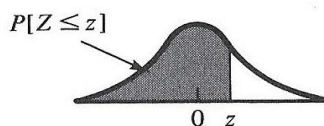
multivariate statistical analysis [View project](#)



أسس التحليل الإحصائي متعدد المتغيرات [View project](#)

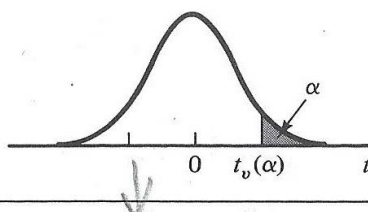
الملاحق: الجداول الاحصائية

TABLE 1 STANDARD NORMAL PROBABILITIES

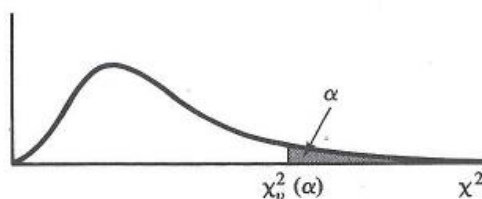


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

TABLE 2 STUDENT'S *t*-DISTRIBUTION CRITICAL POINTS

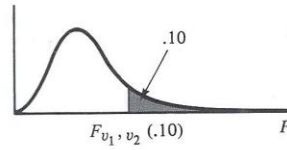


d.f. <i>v</i>	.250	.100	.050	α .025	.010	.00833	.00625	.005
1	1.000	3.078	6.314	12.706	31.821	38.190	50.923	63.657
2	.816	1.886	2.920	4.303	6.965	7.649	8.860	9.925
3	.765	1.638	2.353	3.182	4.541	4.857	5.392	5.841
4	.741	1.533	2.132	2.776	3.747	3.961	4.315	4.604
5	.727	1.476	2.015	2.571	3.365	3.534	3.810	4.032
6	.718	1.440	1.943	2.447	3.143	3.287	3.521	3.707
7	.711	1.415	1.895	2.365	2.998	3.128	3.335	3.499
8	.706	1.397	1.860	2.306	2.896	3.016	3.206	3.355
9	.703	1.383	1.833	2.262	2.821	2.933	3.111	3.250
10	.700	1.372	1.812	2.228	2.764	2.870	3.038	3.169
11	.697	1.363	1.796	2.201	2.718	2.820	2.981	3.106
12	.695	1.356	1.782	2.179	2.681	2.779	2.934	3.055
13	.694	1.350	1.771	2.160	2.650	2.746	2.896	3.012
14	.692	1.345	1.761	2.145	2.624	2.718	2.864	2.977
15	.691	1.341	1.753	2.131	2.602	2.694	2.837	2.947
16	.690	1.337	1.746	2.120	2.583	2.673	2.813	2.921
17	.689	1.333	1.740	2.110	2.567	2.655	2.793	2.898
18	.688	1.330	1.734	2.101	2.552	2.639	2.775	2.878
19	.688	1.328	1.729	2.093	2.539	2.625	2.759	2.861
20	.687	1.325	1.725	2.086	2.528	2.613	2.744	2.845
21	.686	1.323	1.721	2.080	2.518	2.601	2.732	2.831
22	.686	1.321	1.717	2.074	2.508	2.591	2.720	2.819
23	.685	1.319	1.714	2.069	2.500	2.582	2.710	2.807
24	.685	1.318	1.711	2.064	2.492	2.574	2.700	2.797
25	.684	1.316	1.708	2.060	2.485	2.566	2.692	2.787
26	.684	1.315	1.706	2.056	2.479	2.559	2.684	2.779
27	.684	1.314	1.703	2.052	2.473	2.552	2.676	2.771
28	.683	1.313	1.701	2.048	2.467	2.546	2.669	2.763
29	.683	1.311	1.699	2.045	2.462	2.541	2.663	2.756
30	.683	1.310	1.697	2.042	2.457	2.536	2.657	2.750
40	.681	1.303	1.684	2.021	2.423	2.499	2.616	2.704
60	.679	1.296	1.671	2.000	2.390	2.463	2.575	2.660
120	.677	1.289	1.658	1.980	2.358	2.428	2.536	2.617
∞	.674	1.282	1.645	1.960	2.326	2.394	2.498	2.576

TABLE 3 χ^2 CRITICAL POINTS

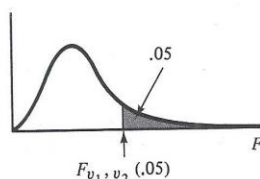
d.f. ν	.990	.950	.900	α .500	.100	.050	.025	.010	.005
1	.0002	.004	.02	.45	2.71	3.84	5.02	6.63	7.88
2	.02	.10	.21	1.39	4.61	5.99	7.38	9.21	10.60
3	.11	.35	.58	2.37	6.25	7.81	9.35	11.34	12.84
4	.30	.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	.55	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75
6	.87	1.64	2.20	5.35	10.64	12.59	14.45	16.81	18.55
7	1.24	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.65	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.95
9	2.09	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.56	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	3.05	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.57	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	4.11	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.66	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	5.23	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.81	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	6.41	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	7.01	9.39	10.86	17.34	25.99	28.87	31.53	34.81	37.16
19	7.63	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	8.26	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.90	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	9.54	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	10.20	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	10.86	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	11.52	14.61	16.47	24.34	34.38	37.65	40.65	44.31	46.93
26	12.20	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	12.88	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.64
28	13.56	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	14.26	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	14.95	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	22.16	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	29.71	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	37.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	45.44	51.74	55.33	69.33	85.53	90.53	95.02	100.43	104.21
80	53.54	60.39	64.28	79.33	96.58	101.88	106.63	112.33	116.32
90	61.75	69.13	73.29	89.33	107.57	113.15	118.14	124.12	128.30
100	70.06	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

TABLE 4 F-DISTRIBUTION CRITICAL POINTS ($\alpha = .10$)



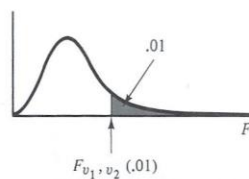
$\begin{matrix} \nearrow \\ P_1 \\ \searrow \end{matrix}$ P_2	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.05	62.26	62.53	62.79
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.17	5.17	5.16	5.15
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.81	2.80	2.78	2.76
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.57	2.56	2.54	2.51
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.27	2.25	2.23	2.21
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.17	2.16	2.13	2.11
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.03	2.01	1.99	1.96
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.93	1.91	1.89	1.86
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.89	1.87	1.85	1.82
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.86	1.84	1.81	1.78
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.83	1.81	1.78	1.75
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.80	1.78	1.75	1.72
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.78	1.76	1.73	1.70
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.76	1.74	1.71	1.68
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.74	1.72	1.69	1.66
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.71	1.69	1.66	1.62
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.68	1.66	1.63	1.59
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.67	1.65	1.61	1.58
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.66	1.64	1.60	1.57
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.65	1.63	1.59	1.56
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.64	1.62	1.58	1.55
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.63	1.61	1.57	1.54
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.50	1.48	1.44	1.40
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24

TABLE 5 F-DISTRIBUTION CRITICAL POINTS ($\alpha = .05$)



α	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60
1	161.5	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	246.0	248.0	249.3	250.1	251.1	252.2
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.46	19.47	19.48
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.63	8.62	8.59	8.57
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.52	4.50	4.46	4.43
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.83	3.81	3.77	3.74
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.40	3.38	3.34	3.30
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.11	3.08	3.04	3.01
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.89	2.86	2.83	2.79
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.73	2.70	2.66	2.62
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.60	2.57	2.53	2.49

12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.50	2.47	2.43	2.38
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.41	2.38	2.34	2.30
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.34	2.31	2.27	2.22
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.28	2.25	2.20	2.16
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.23	2.19	2.15	2.11
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.18	2.15	2.10	2.06
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.14	2.11	2.06	2.02
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.07	2.04	1.99	1.95
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.02	1.98	1.94	1.89
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.00	1.96	1.91	1.86
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.97	1.94	1.89	1.84
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.94	1.90	1.85	1.80
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.92	1.88	1.84	1.79
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.89	1.85	1.81	1.75
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.88	1.84	1.79	1.74
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.78	1.74	1.69	1.64
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.69	1.65	1.59	1.53
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.60	1.55	1.50	1.43
∞	3.84	3.00	2.61	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.51	1.46	1.39	1.32

TABLE 6 F-DISTRIBUTION CRITICAL POINTS ($\alpha = .01$)

α	1	2	3	4	5	6	7	8	9	10	12	15	20	25	30	40	60
1	4052	5000	5403	5625	5764	5859	5928	5981	6023	6056	6106	6157	6209	6240	6261	6287	6313
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.58	26.50	26.41	26.32
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.91	13.84	13.75	13.65
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.45	9.38	9.29	9.20
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.30	7.23	7.14	7.06
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.06	5.99	5.91	5.82
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.26	5.20	5.12	5.03
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.71	4.65	4.57	4.48
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.31	4.25	4.17	4.08
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.01	3.94	3.86	3.78

12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.76	3.70	3.62	3.54
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.57	3.51	3.43	3.34
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.41	3.35	3.27	3.18
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.28	3.21	3.13	3.05
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.16	3.10	3.02	2.93
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.07	3.00	2.92	2.83
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	2.98	2.92	2.84	2.75
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.91	2.84	2.76	2.67
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.84	2.78	2.69	2.61
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.79	2.72	2.64	2.55
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.73	2.67	2.58	2.50
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.69	2.62	2.54	2.45
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.64	2.58	2.49	2.40
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.60	2.54	2.45	2.36
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.57	2.50	2.42	2.33
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.54	2.47	2.38	2.29
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.51	2.44	2.35	2.26
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.48	2.41	2.33	2.23
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.45	2.39	2.30	2.21
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.27	2.20	2.11	2.02
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.10	2.03	1.94	1.84
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.93	1.86	1.76	1.66
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.78	1.70	1.59	1.47