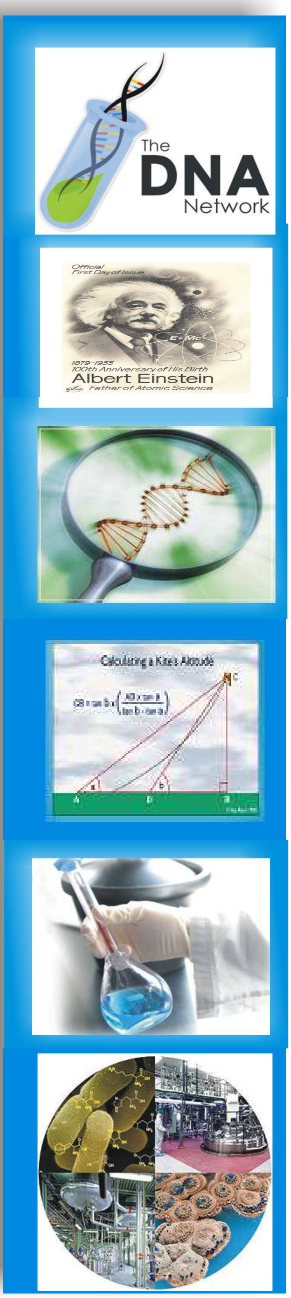


IIT JAM - MATHEMATICS
STATISTICS

SAMPLE THEORY

- EIGEN VALUES
- POISSON DISTRIBUTION
- NEYMAN AND PERSON LEMMA

VPM CLASSES
For IIT-JAM, JNU, GATE, NET, NIMCET and Other Entrance Exams
1-C-8, Sheela Chowdhary Road, Talwandi, Kota (Raj.) Tel No. 0744-2429714
Web Site www.vpmclasses.com E-mail-vpmclasses@yahoo.com

EIGEN VALUES

Let A= $[a_{ij}]_{n\times n}$ be any n rowed square matrix and λ is an indeterminate.

I A - λI is called characteristic matrix.

|A-λ I = 0 is called characteristic equation and roots of this equation is called the characteristic roots or characteristic values or eigen values or latent roots or proper values of the matrix A.

Note. The set of the eigen values of A is called the spectrum of A.

Eigen vectors

If λ is a characteristic root of an n x n matrix A, then a non-zero vector X such that AX = λ X is called a characteristic vector or **eigen vector** of A corresponding to the characteristic root λ.

Remark. 1. λ is a characteristic root of a matrix A if any only if there exists a non- zero vector X such that AX = λX .

2. If X is a characteristic vector of a matrix A corresponding to the characteristic value λ ,then kx is also a characteristic vector of A corresponding to the same characteristic value λ.

Here k is any non-zero scalar.

3. The characteristic vectors corresponding to distinct characteristic roots of a matrix are linearly independent.

Algebraic Multiplicity

If λ_1 be a eigen value of order t of the | A - λI | = 0, then t is called the AM (algebraic multiplicity) of λ_1

Geometric Multiplicity

If s be the number of L.I eigen vector corresponding to the eigen value λ_1, then s is called G.M (geometric multiplity) of λ_1.

Note: G.M.$\leq$ A.M.

Nature of the eigen value of the special types of matrices

1. The eigen value of a Hermitian matrix are all real.
2. The eigen value of a real symmetric matrix are all real.
3. The eigen value of a skew-symmetric matrix are either pure imaginary or zero, for every matrix is Skew-Hermitian.
4. The eigen values of unitary matrices are unit modulus.
5. The eigen roots of an orthogonal matrix are of unit modulus.

Ex . Show that the eigen value of a triangular matrix are just the diagonal elements of the matrix

Let A= $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$ triangular matrix of order 3

| A - λI | = $\begin{bmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ 0 & a_{22}-\lambda & a_{23} \\ 0 & 0 & a_{33}-\lambda \end{bmatrix}$

$= (a_{11}-\lambda)\,(a_{22}-\lambda)\,(a_{33}-\lambda) = 0$

The roots of the equation | A - λ I = 0 are a_{11} , a_{22} , a_{33}

The Cayley-Hamilton theorem

Every square matrix satisfies its characteristic equation, i.e., if for a square matrix A of order n.

$|A-\lambda I| = (-1)^n [\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + ... + a_n]$ then the matrix equation

$x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n I = 0$ is satisfied by $X = A$.

i.e., $A^n + a_1 A^{n-1} + a_2 A^{n-2} + ... + a_n I = 0$

Cor. 1. If A be a non-singular matrix. $|A| \neq 0$.

Premultiplying by A^{-1}

$A^{n-1} + a_1 A^{n-2} + a_2 A^{n-3} + ... + a_{n-1} I + a_n A^{-1} = 0$

or $\quad A^{-1} = -\left(\frac{1}{a_n}\right)(A^{n-1} + a_1 A^{n-2} + ... + a_{n-1} I)$

Cor. 2. If m be a positive integer such that $m \geq n$, then multiplying the results by A^{m-n}

$$A^m + a_1 A^{m-1} + ... + a_n A^{m-n} = 0.$$

Eigen values and eigen vectors

If V is a vector space over the field F and T is a linear operator on V. An **eigen value** of T is a scalar c in F such that there is a non-zero vector $\alpha \in V$ with $T\alpha = c\alpha$. If c is an eigen value of T, then

(a) Any α such that $T\alpha = c\alpha$ is called **eigen vector** of T associated with the eigen value c;

(b) The collection of all c such that $T\alpha = c\alpha$ is called the eigen space associated with c.

Eigen value of matrix A over F

If A is an n × n matrix over the field F, an eigen value of A over F is a scalar c in F such that the matrix (A – cI) is singular (not invertible.)

Eigen polynomial

$f(c) = |A - cI|$.

Diagonalizable

If T is a linear operator on the finite dimensional space V. Then T is diagonalizable if there is a basis for V each vector of which is an eigen vector of T.

Some Important Theorems

1 If T is a linear operator on a finite dimensional space V and c is any scalar. Then following are equivalent

(a) c is an eigen value of T

(b) The operator (T – cI) is singular (not invertible)

(c) det (T – cI) = 0

2 Similar matrices have the same eigen polynomial.

3 If $T\alpha = c\alpha$ and F is any polynomial, then $F(T)\alpha = F(c)\alpha$.

4 Suppose T is a linear operator on the finite, dimensional space V, $c_1 \ldots c_k$ are k-distinct eigen

Values of T and W are the space of eigen vector associated with the eigen value c_i. If

$W = W_1 + W_2 + \ldots + W_k$, then dim W = dim W_1 + dim W_2 + ...+ dim W_k. In fact, if B_i is an ordered basis for W_i then B = (B_1,....,B_k) is an ordered basis for W.

5 If T is a linear operator on finite dimensional values of T and W_i is a null space of (T – c_i I).

The following are equivalent :

(i) T is diagonalizable

(ii) The eigen polynomial for T is $F = (x - c_1)^{d_1} \ldots (x - c_k)^{d_k}$ with dim $W_1 = d_i, i = 1, \ldots k$.

(iii) dim V = dim W_1 + dim W_2 + ...+ dim W_k.

Theorem. If α is a characteristic vector of T corresponding to the characteristic value c, then $k\alpha$ is also a characteristic vector of T corresponding to the same characteristic value c. Here k is any non zero scalar.

Proof. Since α is a characteristic vector of T corresponding to the characteristic value c,

Therefore $\alpha \neq 0$ and

$T(\alpha) = c\alpha$. ... (1)

If k is any non-zero scalar, then $k\alpha \neq \mathbf{0}$.

Also $T(k\alpha) = kT(\alpha) = k(c\alpha) = (kc)\,\alpha$

$= (ck)\,\alpha = c\,(k\alpha)$.

$\therefore$ $k\alpha$ is a characteristic vector of T corresponding to the characteristic value c.

Thus corresponding to a characteristic value c, there may correspond more than one Characteristic vectors.

Theorem. If α is a characteristic vector of T, then α cannot correspond to more than one characteristic values of T.

Proof. Let α be a characteristic vector of T corresponding to two distinct characteristic values c_1

and c_2 of T. Then

$$T\alpha = c_1\alpha$$

and $$T\alpha = c_2\alpha$$

$\therefore$ $$c_1\alpha = c_2\alpha$$

$\Rightarrow$ $$(c_1 - c_2)\,\alpha = 0$$

$\Rightarrow$ $$c_1 - c_2 = 0 \qquad [\because \alpha \neq 0]$$

$\Rightarrow$ $$c_1 = c_2 = 0$$

Theorem . Let T be a linear operator on an n-dimensional vector space V and A be the matrix of T relative to any ordered basis B. Then a vector A in V is an eigenvector of T corresponding to its eigenvalue c if and only if its coordinate vector X relative to the basis B is an eigen-vector of A corresponding to its eigenvalue c.

Proof. We have

$$[T - cI]_B = [T]_B - c[I]_B = A - cI.$$

If $\alpha \neq \mathbf{0}$, then the coordinate vector X of a is also non-zero.

Now $$[(T - cI)(\alpha)]_B = [T - cI]_B\,[\alpha]_B$$

$$= (A - cI)\,X.$$

$\therefore$ $(T - cI)(\alpha) = \mathbf{0}$ iff $(A - cI)\,X = O$

or $T(\alpha) = c\alpha$ iff $AX = cX$

or α is an eigenvector of T iff X is an eigenvector of A.

Thus with the help of this theorem we see that our definition of characteristic vector of a matrix

is sensible Now we shall define the characteristic polynomial of a linear operator. Before doing
so we shall prove the following theorem.

Theorem . Let T be any linear operator on a finite dimensional vector space V, let $c_1, c_2, \ldots, c_k$ be the distinct characteristic values of T, and let W_i be the null space of $(T - c_iI)$. Then the subspace $W_1, \ldots, W_k$ are independent.Further show that if in addition T is diagonalizable, then V is the direct sum of the subspaces of the subspaces $W_1, \ldots, W_k$.

Proof. By definition of W_i, we have

$$W_i = \{\alpha : \alpha \in V \text{ and } (T - c_iI)\,\alpha = 0 \text{ i.e. } T\alpha = c_i\alpha\}.$$

Now let α_i be in W_i, i = 1, ..., k , and suppose that

$$\alpha_1 + \alpha_2 + \ldots + \alpha_k = 0. \qquad \ldots (1)$$

Let j be any integer between 1 and k and let

$$U_j = \prod_{\substack{1 \le i \le k \\ i \ne j}} (T - c_iI).$$

Note that U_j is the product of the operators $(T - c_iI)$ for $i \ne j$. In other words $U_j = (T - c_1I)(T - c_2I)\ldots$.
$(T - c_kI)$ where in the product the factor $T - c_jI$ is missing.

Let us find $U_j\alpha_i$, i = 1, ..., k. By the definition of W_i, we have $(T - c_iI)\,\alpha_i = \mathbf{0}$. Since the operators
$(T - c_iI)$ all commute, being polynomials in T, therefore $U_j\alpha_i = 0$ for $i \ne j$. Not that for each $i \ne j$.
Note that for each $i \ne j$, U_j contains a factor $(T - c_iI)$ and $(T - c_iI)\,\alpha_i = \mathbf{0}$.

Also

$$U_j\alpha_j = [(T - c_1I)\ldots(T - c_kI)]\alpha_j$$

$= [(T - c_1 I)\ldots(T - c_{k-1}I)]\,[T\alpha_j - c_k I\alpha_j)$

$= [(T - c_1 I)\ldots(T - c_{k-1}I)]\,(c_j\,\alpha_j - c_k\alpha_j)$

$[\because T\alpha_j = c_j\alpha_j$ and $I\alpha_j = \alpha_j]$

$= [(T - c_1 I)\ldots(T - c_{k-1}I)]\,(c_j - c_k)\,\alpha_j$

$= (c_j - c_k)\,[(T - c_1 I)\ldots(T - c_{k-1}I)]\alpha_j$

$= (c_j - c_k)\,(c_j - c_{k-1})\ldots(c_j - c_1)\,\alpha_j$, the factor $c_i - c_j$ will be missing. Thus

$$U_j\alpha_j = \left[\prod_{\substack{1\le i\le k\\ i\ne j}}\left(c_j - c_i\right)\right]\alpha_j. \qquad \ldots (2)$$

Now applying U_j to both sides of (1), we get

$U_j\alpha_1 + U_j\alpha_2 + \ldots + U_j\alpha_k = \mathbf{0}$

$\Rightarrow\ U_j\alpha_j = \underline{0} \qquad [\because U_j\alpha_i = \mathbf{0}$ if $i \ne j]$

$\Rightarrow\ \left[\prod_{i\ne j}\left(c_j - c_i\right)\right]\alpha_j = \mathbf{0} \qquad$ [by (2)]

Since the scalars c_i are all distinct, therefore the product

$$\prod_{i\ne j}\left(c_j - c_i\right)$$

is a non-zero scalar. Hence $\left[\prod_{i\ne j}\left(c_j - c_i\right)\right]\alpha_j = 0$

$\Rightarrow\ \alpha_j = 0$. Thus aj = 0 for every integer j between 1 and k.

In this way $\alpha_1 + \ldots + \alpha_k = \mathbf{0}$

$\Rightarrow \alpha_i = \mathbf{0}$ for each i. Hence the subspaces $W_1, \ldots, W_k$ are independent.

Ex. Let V be a n-dimensional vector space over F. What is the characteristic polynomial of

(i) the identity operator on V, (ii) the zero operator on V.

Sol. Let B be any ordered basis for V.

(i) If I is the identity operator on V, then

$$[I]_B = I.$$

The characteristic polynomial of I = det (I – xI)

$$= \begin{vmatrix} 1-x & 0 & \dots & 0 \\ 0 & 1-x & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & 1-x \end{vmatrix} = (1-x)^n.$$

(ii) If $\hat{\mathbf{0}}$ is the zero operator on V, then $\left[\hat{\mathbf{0}}\right]_B = O$ i.e. f the null matrix of order n.

The characteristic polynomial of $\hat{\mathbf{0}} = \det(O - xI)$

$$= \begin{vmatrix} -x & 0 & \dots & 0 \\ 0 & -x & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & -x \end{vmatrix} = (-1)^n x^n.$$

Poisson distribution as a limiting case of the negative Binomial distribution

Negative binomial distribution tends to Poisson distribution as $P \to 0$, $r \to \infty$ such that rP = λ (finite). Proceeding to the limits, we get

$$\lim p(x) = \lim \binom{x+r-1}{r-1} p^r q^x = \lim \binom{x+r-1}{x} Q^{-r} \left(\frac{P}{Q}\right)^x \quad \left[\text{let } p = \frac{1}{2}, q = \frac{p}{2}\right]$$

$$= \lim_{r \to \infty} \frac{(x+r-1)(x+r-2)\dots(r+1)r}{x!} (1+P)^{-r} \left(\frac{P}{1+P}\right)^x$$

$$= \lim_{r \to \infty} \left\{ \frac{1}{x!} \left(1 + \frac{x-1}{r}\right)\left(1 + \frac{x-2}{r}\right)\dots\left(1 + \frac{1}{r}\right).1.r^x (1+P)^{-r} \left(\frac{P}{1+P}\right)^x \right\}$$

$$= \frac{1}{x!} \lim_{r \to \infty} \left\{ (1+P)^{-r} \left(\frac{rP}{1+P} \right)^{x} \right\} = \frac{\lambda^{x}}{x!} \lim_{r \to \infty} \left[\left(1 + \frac{\lambda}{r} \right) \right]^{-r} \lim_{r \to \infty} \left(1 + \frac{\lambda}{r} \right)^{-x} [\because rP = \lambda]$$

$$= \frac{\lambda^{x}}{x!} . e^{-\lambda} . 1 = \frac{e^{-\lambda} \lambda^{x}}{x!}$$

which is the probability function of the Poisson distribution with parameter 'λ'.

Deduction of moments of negative Binomial distribution from those of Binomial distribution

If we write $p = 1/Q$, $q = P/Q$ such that $Q - P = 1$, then the m.g.f. of negative binomial variate X is

given by :

$M_X(t) = (Q - Pe^t)^{-k}$ ...(1)

This is analogous to m.g.f. of binomial variate Y with parameters n and p', viz.,

$M_Y(t) = (q' + p'e^t)^n$; $q' = 1 - p'$ ...(2)

Comparing (1) and (2), we get $q' = Q$, $p' = -P$ and $n = -k$ ...(3)

Using the formulae for moments of binomial distribution, the moments of negative binomial distribution are given by:

Mean $= np' = (-k)(-P) = kP$

Variance $= np'q' = (-k)(-P)Q = kPQ$

$\mu_3 = np'q'(q' - p') = (-k)(-P)Q(Q + P) = kPQ(Q + P)$

$\mu_4 = np'q'[1 + 3p'q'(n - 2)] = (-k)(-P)Q[1 + 3(-P)Q(-k - 2)] = kPQ[1 + 3PQ(k + 2)]$.

Ex. Given the hypothetical distribution:

No. of cells (x) :	0	1	2	3	4	5	Total
Frequency (f) :	213	128	37	18	3	1	400

Fit a negative binomial distribution and calculate the expected frequencies.

Sol. Let X be negative binomial variate with parameters r and p.

$$\mu_1' = \text{Mean} = \frac{\Sigma fx}{\Sigma f} = \frac{273}{400} = 0.6825 = \frac{rq}{p} \quad ; \quad (q = 1 - p)$$

$$\mu_2' = \frac{\Sigma fx^2}{\Sigma f} = \frac{511}{400} = 1.2775 \qquad \text{...(1)}$$

$$\mu_2 = \mu_2' - \mu_1'^2 = 1.2775 - (0.6825)^2 = 0.8117$$

$$\therefore \text{ Variance} = 0.8117 = \frac{rq}{p^2} \qquad \text{...(2)}$$

Dividing (1) by (2), we get $p = \frac{0.6825}{0.8117} = 0.8408$, $q = 1 - p = 0.1592$

$$\therefore \quad r = \frac{p \times 0.6825}{q} = \frac{0.5738}{0.1592} = 3.6043 \simeq 4 \quad \text{[From (1)]} \qquad \text{...(3)}$$

Since, r being the number of successes cannot be fractional.

$$f_0 = p^r = (0.8408)^4 = 0.4978 \;\; 0.5$$

$$f_1 = \frac{r+0}{0+1} q f_0 = rqf_0 = 0.5738 \times 0.5 = 0.2869$$

$(\therefore\ rq = p \times 0.6825 = 0.5738)$ [From (***)]

$$f_2 = \frac{r+1}{1+1} \cdot q \cdot f_1 = \frac{5}{2} \times 0.1592 \times 0.2869 = 0.1142$$

$$f_3 = \frac{r+2}{2+1} \cdot q \cdot f_2 = \frac{6}{3} \times 0.1592 \times 0.1142 = 0.0364$$

$$f_4 = \frac{r+3}{3+1} \cdot q \cdot f_3 = \frac{7}{4} \times 0.1592 \times 0.0364 = 0.0101$$

$$f_5 = \frac{r+4}{4+1} \cdot q \cdot f_4 = \frac{8}{5} \times 0.1592 \times 0.0101 = 0.0026$$

$\therefore$ Expected frequencies are : (N = 400)

$Nf_0 \quad Nf_1 \quad Nf_2 \quad Nf_3 \quad Nf_4 \quad Nf_5$

200 114.76 45.68 14.56 4.04 1.04

∴

Observed frequency :	213	128	37	18	3	1
Expected frequency :	200	115	46	14	4	1

Poisson distribution

A random variable X is said to follow a Poisson distribution if it assumes only non-negative values
and its probability mass function is given by :

$$p(x, \lambda) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!}; x = 0, 1, 2, \ldots; \lambda > 0 \\ 0, \text{otherwise} \end{cases}$$

Moments of the Poisson distribution

$$\mu_1' = E(X) = \sum_{x=0}^{\infty} x\, p(x, \lambda) = \sum_{x=0}^{\infty} x \,.\, \frac{e^{-\lambda}\lambda^{x}}{x!} = \lambda e^{-\lambda}\left\{\sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}\right\} = \lambda e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\right) = \lambda e^{-\lambda} \,.\, e^{\lambda} = \lambda$$

Hence the mean of the Poisson distribution is λ.

$$\mu_2' = E(X^2) = \sum_{x=0}^{\infty} x^2 p(x, \lambda) = \sum_{x=0}^{\infty} \{x(x-1) + x\}\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty} x(x-1)\frac{\lambda^{x}}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^{x}}{x!} = \lambda^2 e^{-\lambda}\left[\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}\right] + \lambda = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda$$

$$\mu_3' = E(X^3) = \sum_{x=0}^{\infty} x^3 p(x, \lambda) = \sum_{x=0}^{\infty} \{x(x-1)(x-2) + 3x(x-1) + x\}\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2)\frac{e^{-\lambda}\lambda^{x}}{x!} + 3\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda}\lambda^3\left\{\sum_{x=3}^{\infty}\frac{\lambda^{x-3}}{(x-3)!}\right\} + 3e^{-\lambda}\lambda^2\left\{\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}\right\} + \lambda = e^{-\lambda}\lambda^3 e^{\lambda} + 3e^{-\lambda}\lambda^2 e^{\lambda} + \lambda = \lambda^3 + 3\lambda^2 + \lambda$$

$$\mu_4' = E(X^4) = \sum_{x=0}^{\infty} x^4 \,.\, p(x, \lambda)$$

$$= \sum_{x=0}^{\infty}\{x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x\}\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda^4(e^{-\lambda} e^{\lambda}) + 6\lambda^3(e^{-\lambda} e^{\lambda}) + 7\lambda^2(e^{-\lambda} e^{\lambda}) + \lambda = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda$$

Coefficients of skewness and kurtosis are given by :

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \quad \text{and} \quad \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}}$$

Also $\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1}{\lambda}$ and $\gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}$

Hence the Poisson distribution is always a skewed distribution.

Proceeding to the limit as $\lambda \to \infty$ and $\beta_2 = 3$.

Mode of the Poisson distribution

$$\frac{p(x)}{p(x-1)} = \frac{\frac{e^{-\lambda}\lambda^x}{x!}}{\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}} = \frac{\lambda}{x}$$

We discuss the following cases :

***Case–I* :**

When λ is not an integer. Let us suppose that S is the integral part of λ, so that

$\lambda = S + f,\ 0 < f < 1.$

We get :

$$\frac{p(x)}{p(x-1)} = \frac{S+f}{x} = \begin{cases} > 1, \text{ if } x = 0, 1, \ldots, S \\ < 1, \text{ if } x = S+1, S+2, \ldots \end{cases}$$

$$\frac{p(1)}{p(0)} > 1, \frac{p(2)}{p(1)} > 1, \ldots, \frac{p(S-1)}{p(S-2)} > 1, \frac{p(S)}{p(S-1)} > 1, \quad \text{and} \quad \frac{p(S+1)}{p(S)} < 1, \frac{p(S+2)}{p(S+1)} < 1, \ldots$$

Combining the above expressions into a single expression, we get

$$p(0) < p(1) < p(2) < \ldots < p(S-2) < p(S-1) < p(S) > p(S+1) > p(S+2) > \ldots ,$$

which shows that p(S) is the maximum value. Hence, in this case, the distribution is unimodal and the

integral part of λ is the unique modal value.

***Case–II* :**

When λ = k(say) is an integer. Here, as in case–I, we have

$$\frac{p(1)}{p(0)} > 1, \frac{p(2)}{p(1)} > 1, \ldots, \frac{p(k-1)}{p(k-2)} > 1 \quad \text{and} \quad \frac{p(k)}{p(k-1)} = 1, \frac{p(k+1)}{pk} < 1, \frac{p(k+2)}{p(k+1)} < 1, \ldots$$

$\therefore$ $p(0) < p(1) < p(2) < \ldots < p(k-2) < p(k-1) = p(k) > p(k+1) > p(k+2) \ldots$

In this case we have two maximum values, viz., p(k – 1) and p(k) and thus the distribution is bimodal

and two modes are at (k – 1) and k, i.e., at (λ – 1) and λ, (since k = λ).

Moment generating function of the Poisson distribution

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}(\lambda e^t)^x}{x!} = e^{-\lambda}\left\{1 + \lambda e^t + \frac{(\lambda e^t)^2}{2!} + \ldots\right\} = e^{-\lambda} \cdot e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

Cumulants Characteristic function of the Poisson distribution

$$\phi_X(t) = \sum_{x=0}^{\infty} e^{itx} . p(x, \lambda) = \sum_{x=0}^{\infty} e^{itx} \frac{e^{-\lambda}\lambda x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^{it})^x}{x!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)}.$$

$$K_X(t) = \log M_X(t) = \log[e^{\lambda(e^t-1)}] = \lambda(e^t - 1)$$

$$= \lambda\left[\left(1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots + \frac{t^r}{r!} + \dots\right) - 1\right] = \lambda\left[t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots + \frac{t^r}{r!} + \dots\right]$$

κ_r = r^{th} cumulant = Coefficient of $\frac{t^r}{r!}$ in $K_X(t) = \lambda \Rightarrow \kappa_r = \lambda$; $r = 1, 2, 3, \dots$

Hence, all cumulants of the Poisson distribution are equal, each being equal to λ. In particular, we have

Mean $= \kappa_1 = \lambda$, $\mu_2 = \kappa_2 = \lambda$, $\mu_3 = \kappa_3 = \lambda$ and $\mu_4 = \kappa_4 + 3\kappa_2^2 = \lambda + 3\lambda^2$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\lambda + 3\lambda^2}{\lambda^2} = \frac{1}{\lambda} + 3$$

Ex. In a Poisson frequency distribution, frequency corresponding to 3 successes is 2/3 times frequency corresponding to 4 successes. Find the mean and standard deviation of the distribution.

Sol. Let X be a random variable following Poisson distribution with parameter λ. Then the frequency function is given by

$$f(x) = N \,.\, p(x) = NP(X = x) = N \,.\, \frac{e^{-\lambda} . \lambda^x}{x!} \; ; x = 0, 1, 2, \dots \qquad \dots(1)$$

Putting x = 3 and 4 in (1), $f(3) = N \,.\, \frac{e^{-\lambda}\lambda^3}{3!}$ and $f(4) = N \,.\, \frac{e^{-\lambda}\lambda^4}{4!}$

We are given : $f(3) = \frac{2}{3} f(4) \Rightarrow N . \frac{e^{-\lambda}\lambda^3}{3!} = \frac{2}{3} N . \frac{e^{-\lambda}\lambda^4}{4!}$

$$\Rightarrow \frac{1}{3!} = \frac{2}{3} . \frac{\lambda}{4!} \qquad \Rightarrow \lambda = \frac{1}{3!} \times \frac{3}{2} \times 4! = 6 .$$

$\Rightarrow$ Mean of poisson distribution

$\lambda = 6$ and s.d. of the

Distribution $=\sqrt{\lambda}=\sqrt{6}$

NEYMAN J. AND PEARSON, E.S LEMMA

This Lemma provides the most powerful test of simple hypothesis against a simple alternative

hypothesis. The theorem, known as Neyman Pearson Lemma, will be provided for density function f(x, θ) of a single continuous variate and a single parameter. However by regarding x and θ as vectors the proof can be easily generalized for any number of random variables $x_1, x_2, \ldots, x_n$ and any number of parameters $\theta_1, \theta_2, \ldots, \theta_k$. The variables $x_1, x_2, \ldots, x_n$ occurring in this theorem are understood to represent a random sample of size n from the population ,whose density function is f(x, θ). The lemma is concerned with a simple hypothesis $H_0 : \theta = \theta_0$ and a simple alternative $H_1 : \theta = \theta_1$.

Neyman Pearson Lemma Let k > 0 be a constant and W be a critical region of size such that

$$W = \left\{ x \in S : \frac{f(x,\theta_1)}{f(x,\theta_0)} > k \right\}$$

$$W = \left\{ x \in S : \frac{L_1}{L_0} > k \right\} \qquad \ldots (1)$$

and $$\overline{W} = \left\{ x \in S : \frac{L_1}{L_0} < k \right\} \qquad \ldots (2)$$

where L_0 and L_1 are the likelihood functions of the sample observations x = ($x_1, x_2, ..., x_n$) under H_0 and H_1 respectively. Then W is the most powerful critical region of the test hypothesis $H_0 : \theta = \theta_0$ against the alternative $H_1 : \theta = \theta_1$.

Proof. We are given

$$P(x \in W \mid H_0) = \int_W L_0\, dx = \alpha \qquad \text{... (1a)}$$

The power of the region is

$$P(x \in W \mid H_1) = \int_W L_1\, dx = 1 - \beta, (\text{say}). \qquad \text{... (1b)}$$

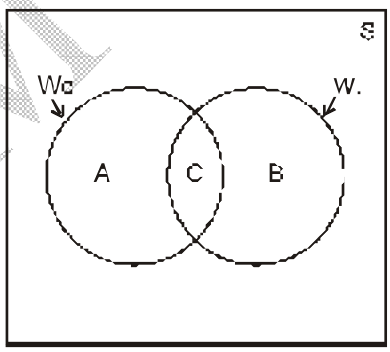
In order to establish the lemma, we have to prove that there exists no other critical region, of size less than or equal to α, which is more powerful than W. Let W_1 be another critical region of size $\alpha_1 \le \alpha$ and power $1 - \beta_1$ so that we have

$$P(\mathbf{x} \in W_1 \mid H_0) = \int_{W_1} L_0\, d\mathbf{x} = \alpha_1 \qquad \text{... (2)}$$

and $$P(x \in W_1 \mid H_0) = \int_{W_1} L_1\, d\mathbf{x} = 1 - \beta_1 \qquad \text{... (3)}$$

Now we have to prove that $1 - \beta \ge 1 - \beta_1$

Let $W = A \cup C$ and $W_1 = B \cup C$



(C may be empty, i.e., W and W1 may be disjoint).

If $\alpha_1 \le \alpha$, we have

$$\int_{W_1} L_0\, dx \le \int_W L_0\, d\mathbf{x}$$

$$\Rightarrow \quad \int_{B\cup C} L_0\, d\mathbf{x} \le \int_{A\cup C} L_0\, d\mathbf{x}$$

$$\Rightarrow \quad \int_B L_0\, d\mathbf{x} \le \int_A L_0\, d\mathbf{x}$$

$$\Rightarrow \quad \int_A L_0\, d\mathbf{x} \ge \int_B L_0\, d\mathbf{x} \qquad \text{... (4)}$$

Since $A \subset W$,

(1) $\Rightarrow \int_A L_1\, d\mathbf{x} > k\int_A L_0\, d\mathbf{x} \ge k\int_B L_0\, d\mathbf{x}$ [Using (4)] ... (4a)

Also [1(a)] implies

$$\frac{L_1}{L_0} \le k \ \forall \ x \in \bar{W}$$

$$\Rightarrow \quad \int \bar{W} L_1\, d\mathbf{x} \le k\int \bar{W} L_0\, d\mathbf{x}$$

This result also holds for any subset of $\bar{W}$, say $\bar{W} \cap W_1 = B$. Hence

$$\int_B L_1\, d\mathbf{x} \le k\int_B L_0\, d\mathbf{x} \le \int_A L_1\, d\mathbf{x} \qquad \text{[From (40)]}$$

Adding $\int_C L_1 d\mathbf{x}$ to both sides, we get

$$\int_{W_1} L_1\, d\mathbf{x} \le \int_{\mathbf{w}} L_1\, d\mathbf{x} \quad \Rightarrow \quad 1-\beta \ge 1-\beta_1$$

Hence the Theorem.

Unbiased Test and Unbiased Critical Region. Let us consider the testing of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$: The critical region W and consequently the test based on it is said to be unbiased if the power of the test exceeds the size of the critical region, i.e, if

Power of the test $\ge$ Size of the C.R.

$$\Rightarrow \quad 1-\beta \ge \alpha$$

$$\Rightarrow \quad P_{\theta_1}(W) \ge P_{\theta_0}(W)$$

$$\Rightarrow \quad P[\mathbf{x} : \mathbf{x} \in W \mid H_1] \ge P[\mathbf{x} : \mathbf{x} \in W \mid H_0]$$

In other words, the critical region W is said to be unbiased if

$\Rightarrow \quad P_\theta(W) \geq P_{\theta_0}(W), \forall \theta (\neq \theta_0) \in \Theta$

Theorem. Every most powerful (MP) or uniformly most powerful (UMP) critical region (CR) is necessarily unbiased.

(a) If W be an MPCR of size α for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ then it is necessarily unbiased.

(b) Similarly if W be UMPCR of size α for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \in \Theta_1$ then it is also unbiased

Proof. (a) Since W is the MPCR of size α for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$, by Neyman Pearson Lemma, we have ; for $\forall k > 0$,

$W = \{\mathbf{x} : L(\mathbf{x}, \theta_1) \geq k\, L(\mathbf{x}, \theta_0)\} = \{x : L_1 k \geq L_0\}$

and $W' = \{\mathbf{x} : L(\mathbf{x}, \theta_1) < k\, L(\mathbf{x}, \theta_0)\} = \{x : L_1 < k\, L_0\}$,

where k is determined so that the size of test is α i.e.,

$$P_{\theta_0}(W) = P(\mathbf{x} \in W \mid H_0) = \int_W L_0\, d\mathbf{x} = \alpha \qquad \text{... (i)}$$

To prove that W is unbiased, we have to show that :

Power of $W \geq \alpha$ i.e., $P_{\theta_1}(W) \geq \alpha$ ... (ii)

We have : $P_{\theta_1}(W) = \int_W L_1\, d\mathbf{x} \geq k \int_W L_0\, d\mathbf{x} = k\alpha$

[$\because$ on W, $L_1 \geq k L_0$ and using (i)] ... (iii)

i.e., $P_{\theta_1}(W) \geq k\alpha, \ \forall k > 0$

Also

$$1 - P_{\theta_1}(w) = 1 - P(\mathbf{x} \in W \mid H_1) = P(\mathbf{x} \in W' \mid H_1) = \int_{W'} L_1\, d\mathbf{x}$$

$$= k\int_{W'} L_0\, d\mathbf{x} = kP(x : x \in W' \mid H_0) \qquad [\because \text{ on } W',.\ L_1 < k\, L_0]$$

$$= k\,[1 - P(\mathbf{x} : \mathbf{x} \in W \mid H_0)] \qquad [\text{Using (i)}]$$

i.e., $1-P_{\theta_1}(W) \le k(1-\alpha), \forall\ k > 0$ ... (iv)

Case (i) $k \ge 1$. if $k \ge 1$, then from (iii), we get :

$$P_{\theta_1}(W) \ge k\alpha \ge \alpha$$

$\Rightarrow$ W is unbiased CR.

Case (ii) 0 < k < 1. If 0 < k < 1, then from (iv), we get :

$1-P_{\theta_1}(W) < 1-\alpha \Rightarrow P_{\theta_1}(W) > \alpha \Rightarrow$ W is unbiased C.R.

Hence MP critical region is unbiased.

(b) if W is UMPCR of size α then also the above proof holds if for θ_1 we write θ such that $\theta \in \Theta_1$. So we have

$P_\theta(W) > \alpha,\ \forall \theta \in \Theta_1 \Rightarrow$ W is unbiased CR.

Optimum Regions and sufficient Statistics. Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from a

population with p.m.f or p.d.f. $f(x, \theta)$ where the parameter θ may be vector. Let T be sufficient for θ.

Let T be a sufficient statistic for θ. Then by Factorization Theorems,

$$L(x,\theta) = \prod_{i=1}^{n} f(x_i,\theta) = g_\theta(t(x)).h(x)$$

where $g_\theta(t(x))$ is the marginal distribution of the statistic T = t(x).

By Neyman Pearson Lemma the MPCR for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ is given by :

$$W = \{\mathbf{x} : L(\mathbf{x},\theta_1) \ge k\,L(\mathbf{x},\theta_0)\}, \forall\ k > 0$$

$$W = \left\{\mathbf{x} : g_{\theta_1}(t(\mathbf{x})).h(\mathbf{x}) \ge k.\ g_{\theta_0}(t(\mathbf{x})).h(\mathbf{x})\right\}, \forall\ k > 0$$

$$= \left\{\mathbf{x} : g_{\theta_1}(t(\mathbf{x})) \ge k.\ g_{\theta_0}(t(\mathbf{x}))\right\}, \forall\ k > 0$$

Hence if T = t (x) is sufficient statistic for θ then the MPCR for the test may be defined in terms of the marginal distribution of $T = t(x)$, rather than the joint distribution of $x_1, x_2, ..., x_n$.

Ex. Use the Neyman-Pearson Lemma to obtain the region for testing $\theta = \theta_0$ against $\theta = \theta_1 > \theta_0$ and

$\theta = \theta_1 < \theta_0$, in the case of the a normal population $N(\theta, \sigma^2)$, where σ^2 is known. Hence find the power of the test.

Sol. $L = \prod_{i=1}^{n} f(x_i, \theta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right\}$

Using Neyman-Pearson Lemma, best critical region (B.C.R) is given by (for k > 0)

$$\frac{L_1}{L_0} = \frac{\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta_1)^2\right\}}{\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta_0)^2\right\}} \geq k$$

$$\Rightarrow \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta_1)^2 - \sum_{i=1}^{n}(x_i - \theta_0)^2\right] \geq k$$

$$\Rightarrow \exp\left[-\frac{n}{2\sigma^2}(\theta_1^2 - \theta_0^2) + \frac{1}{\sigma^2}(\theta_1 - \theta_0)\sum_{i=1}^{n} x_i\right] \geq k$$

$$\Rightarrow -\frac{n}{2\sigma^2}(\theta_1^2 - \theta_0^2) + \frac{1}{\sigma^2}(\theta_1 - \theta_0)\sum_{i=1}^{n} x_i \geq \log k$$

(since log x is an increasing function of x)

$$\Rightarrow \bar{x}(\theta_1 - \theta_0) \geq \frac{\sigma^2}{n}\log k + + \frac{\theta_1^2 - \theta_0^2}{2}$$

Case (i) If $\theta_1 > \theta_0$, the B.C.R. is determined by the relation (right-tailed test) :

$$\bar{x} > \frac{\sigma^2}{n} \cdot \frac{\log k}{\theta_1 - \theta_0} + \frac{\theta_1 + \theta_0}{2}$$

$\Rightarrow \quad \bar{x} > \lambda_1$, (say).

$\therefore \quad$ B.C.R. is $W = \{x : x > \lambda_1\}$ ... (1)

Case (ii) If $\theta_1 < \theta_0$, the B.C.R. is given by the relation (left handed test)

$$\bar{x} < \frac{\sigma^2}{n}\cdot\frac{\log k}{\theta_1+\theta_0} + \frac{\theta_1+\theta_0}{2} = \lambda_2, \text{(say)}.$$

Hence B.C.R. is : $W_1 = \{x : \bar{x} \le \lambda_2\}$ ... (2)

The constants λ_1 and λ_2 are so chosen as to make the probability of each of the relations (18.10) and (18.11) equal to α when the hypothesis H_0 is true. The sampling distribution of $\bar{x}$, when H is true is $N\left(\theta_i, \frac{\sigma^2}{n}\right), (i = 0,1)$. Therefore, the constants λ_1 and λ_2 are determined from the relations :

$$P[\bar{x} > \lambda_1 \mid H_0] = \alpha \text{ and } P[\bar{x} < \lambda_2 \mid H_0] = \alpha$$

$$\therefore \quad P(\bar{x} > \lambda_1 \mid H_0) = P\left[Z > \frac{\lambda_1 - \theta_0}{\sigma/\sqrt{n}}\right] = \alpha; Z \sim N(0,1)$$

$$\Rightarrow \quad \frac{\lambda_1 - \theta_0}{\sigma/\sqrt{n}} = z_\alpha \Rightarrow \lambda_1 = \theta_0 + \frac{\sigma}{\sqrt{n}} z_\alpha \qquad \text{... (3)}$$

where z_α is the upper α-point of the standard normal variate given by :

$$P(Z > z_\alpha) = \alpha \qquad \text{... (i)}$$

Also $P(\bar{x} < \lambda_2 \mid H_0) = \alpha \Rightarrow P(\bar{x} \ge \lambda_2 \mid H_0) = 1 - \alpha$

$$\Rightarrow \quad P\left(Z \ge \frac{\lambda_2 - \theta_0}{\sigma/\sqrt{n}}\right) = 1 - \alpha \Rightarrow \frac{\lambda_2 - \theta_0}{\sigma/\sqrt{n}} = z_{1-\alpha}$$

$$\Rightarrow \quad \lambda_2 = \theta_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha} \qquad \text{... 3(a)}$$

Power of the test. By definition, the power of the test in case (i) is :

$$1 - \beta = P[\mathbf{x} \in W \mid H_1] = P[\bar{x} \ge \lambda_1 \mid H_1]$$

$$= P\left(Z \ge \frac{\lambda_1 - \theta_1}{\sigma/\sqrt{n}}\right) \qquad \left[\because \text{Under } H_1, Z = \frac{\bar{x} - \theta_1}{\sigma/\sqrt{n}} \sim N(0,1)\right]$$

$$= P\left(Z \ge \frac{\theta_0 + \frac{\sigma}{\sqrt{n}} z_\alpha - \theta_1}{\sigma/\sqrt{n}}\right) \qquad \text{[Using (3)]}$$

$$= P\left(Z \ge z_\alpha - \frac{\theta_1 - \theta_0}{\sigma/\sqrt{n}}\right) \qquad (\because\ \theta_1 > \theta_0)$$

$$= 1 - P(Z \le \lambda_3) \qquad \left\{\lambda_3 = z_\alpha - \frac{\theta_1 - \theta_0}{\sigma/\sqrt{n}}, \text{say.}\right\}$$

$$= 1 - \Phi(\lambda_3), \qquad \text{.... (4)}$$

where $\Phi(.)$ is the distribution function of standard normal variate.

Similarly in case (ii), $(\theta_1 < \theta_0)$, the power of the test is

$$1 - \beta = P(\bar{x} < \lambda_2 \mid H_1) = P\left(Z < \frac{\lambda_2 - \theta_1}{\sigma/\sqrt{n}}\right)$$

$$= P\left(Z < \frac{\theta_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha} - \theta_1}{\sigma/\sqrt{n}}\right) \qquad \text{[Using (3a)]}$$

$$= P\left(Z < z_{1-\alpha} + \frac{\theta_0 - \theta_1}{\sigma/\sqrt{n}}\right) = \Phi(\lambda_4), \qquad (\because\ \theta_0 > \theta_1) \qquad \text{... (4a)}$$

where $\lambda_4 = z_{1-\alpha} + \dfrac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} = \dfrac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} - z_\alpha$ ... (4a)

UMP Critical Region. Provided best critical region for testing $H_0 : \theta = \theta_0$ against the hypothesis $\theta = \theta_1$,

provided $\theta_1 > \theta_0$ while defines the best critical region for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$, provided $\theta_1 < \theta_0$. Thus, the best critical region for testing simple hypothesis $H_0 : \theta = \theta_0$ against the simple hypothesis $\theta = \theta_1 + c$, $c > 0$ will not serve as best critical region for testing simple hypothesis $H_0 : \theta = \theta_0$ against simple alternative hypothesis $H_1 : \theta = \theta_0 - c, c > 0$.

Hence in this problem, no uniformly most powerful test exists for testing the simple hypothesis, $H_0 : \theta = \theta_0$ against the composite alternative hypothesis, $H_1 : \theta \neq \theta_0$.

However, for each alternative hypothesis, f$H_1 : \theta = \theta_1 > \theta_0$ or $H_1 : \theta = \theta < \theta_0$,a UMP test exists and is given by and respectively.

Remark : In particular, if we take n = 2, then the B.C.R. for testing $H_0 : \theta = \theta_0$, against $H_1 : \theta_1$ ($> \theta_0$) is given by

$$W = \{x : (x_1 + x_2)/2 \geq \theta_0 + \sigma z_\alpha / \sqrt{2}\}$$

$$= \{x : x_1 + x_2 \geq 2\theta_0 + \sqrt{2}\,\sigma z_\alpha\} \qquad [\because \bar{x} = (x_1 + x_2)/2]$$

$$= \{x : x_1 + x_2 \geq C\}, \text{(say)}$$

where $\quad C = 2\theta_0 + \sqrt{2}\sigma\, z_\alpha = 2\theta_0 + \sqrt{2}\,\sigma \times 1.645, \text{ if } \alpha = 0.05.$ ... (**)

Similarly, the B.C.R. for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ ($< \theta_0$) with n = 2 and α = 0.05 is given by

$$W_1 = \{x : (x_1 + x_2)/2 \leq \theta_0 - \sigma z_\alpha / \sqrt{2}\}$$

$$= \{x : (x_1 + x_2) \leq 2\theta_0 - \sqrt{2}\sigma \times 1.645\}$$

$$= \{x : x_1 + x_2 \leq C_1\}, \text{(say)}, \qquad \text{... (***)}$$

where $C_1 = 2\theta_0 - \sqrt{2}\sigma z_\alpha = 2\theta_0 - \sqrt{2}\sigma \times 1.645$, if $\alpha = 0.05$

The B.C.R. for testing $H_0 : \theta = \theta_0$ against the two tailed alternative

$H_1 : \theta = \theta_1 (\neq \theta_0)$, is given by : $W_2 = \{x : (x_1 + x_2 \geq C) \cup (x_1 + x_2 \leq C_1)\}$ ... (****)

The regions in (**), (***) , and (****) are given by the shaded portions in the following figures (i), (ii) and (iii)

respectively.

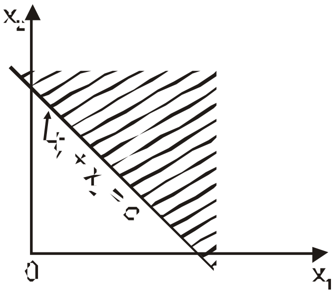


Fig. (i)

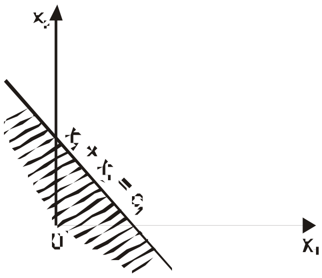


Fig. (ii)

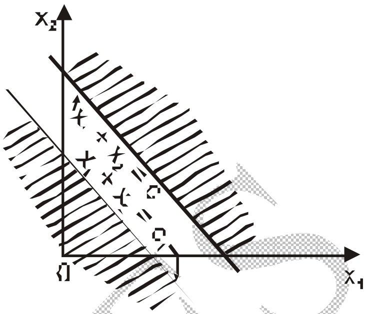


Fig. (iii)

$$\left.\begin{matrix}\text{BCR}\\ \text{for}\end{matrix}\right\}:\begin{matrix}H_0:\theta=\theta_0\\ H_1:\theta=\theta_1(>\theta_0)\end{matrix}\qquad \left.\begin{matrix}\text{BCR}\\ \text{for}\end{matrix}\right\}:\begin{matrix}H_0:\theta=\theta_0\\ H_1:\theta=\theta_1(<\theta_0)\end{matrix}\qquad \left.\begin{matrix}\text{BCR}\\ \text{for}\end{matrix}\right\}:\begin{matrix}H_0:\theta=\theta_0\\ H_1:\theta=\theta_1(\neq\theta_0)\end{matrix}$$

Ex. Examine whether a best critical region exists for testing the null hypothesis $H_0 : \theta = \theta_0$ against the alternative hypothesis $H_1 : \theta > \theta_0$ for the parameter θ of the distribution.

$$f(x,\theta)=\frac{1+\theta}{(x+\theta)^2}, 1\le x<\infty$$

Sol. $\prod_{i=1}^{n} f(x_i,\theta)=(1+\theta)^n\prod_{i=1}^{n}\frac{1}{(x_i+\theta)^2}$

By Neyman-Pearson Lemma, the B.C.R. for $k > 0$, is given by

$$(1+\theta_1)^n\prod_{i=1}^{n}\frac{1}{(x_i+\theta_1)^2}\ge k(1+\theta_0)^n\prod_{i=1}^{n}\frac{1}{(x_i+\theta_0)^2}$$

$$\Rightarrow\quad n\log(1+\theta_1)-2\sum_{i=1}^{n}\log(x_i+\theta_1)\ge\log k+n\log(1+\theta_0)-2\sum_{i=1}^{n}\log(x_i+\theta_0)$$

$$\Rightarrow\quad 2\sum_{i=1}^{n}\log\left(\frac{x_i+\theta_0}{x_i+\theta_1}\right)\ge\log k+n\log\left(\frac{1+\theta_0}{1+\theta_1}\right)$$

Thus the test criterion is $\sum_{i=1}^{n}\log\left(\frac{x_i+\theta_0}{x_i+\theta_1}\right)$, which cannot be put in the form of a function of the sample observations, not depending on the hypothesis. Hence no B.C.R. exists in this case.