Statistical Inference

Department of Statistics

Fourth Stage

2022-2023

Contents

References

1. Introduction to Mathematical Statistics, 5th edition; By Robert V. Hogg and Craig,1995.

2. Introduction to Probability Theory and Statistical Inference, 3rd edition; By Harold J. Larson, 1982.

3. Statistical inference / George Casella, Roger L. Berger.-2nd edition 2002.

4. Principles of Statistical Inference, D.R. Cox, 2006.

5. An introduction to Probability and Mathematical Statistics, Rohatgi, V.K. , 1976.

6. Theory of Point Estimation, E.L. Lehmann George Casella 2nd edition 1998.

7. Statistical Distributions. Merran Evans, Nicholas Hastings, Brian Peacock, 3rd Edition, 2000.

7. Mathematical Statistics. Ferguson, T.S. 1968.

8. Statistical inference. Silvey 1973.

9. Bayesian Inference in Statistical Analysis. Box and Tiro 1973.

10. The Theory of Statistical Inference. Zacks, S.

11. Introduction to Probability and Statistical Inference. George Roussas 2003.

12. Probability and Mathematical Statistics. Prasanna Sahoo 2013.

Chapter one

First review all subjects and laws of Mathematical Statistics

Statistical Distributions

First: The Discrete Distributions

1) Discrete Uniform Distribution X ~ D.U (*n*)

$$f(x;n) = \begin{cases} \dfrac{1}{n} & x = 1,2,\ldots,n \\ 0 & o.w \end{cases}$$

$$mean = \frac{n+1}{2} \quad , \quad \operatorname{var}(X) = \frac{n^2-1}{12}$$

2) Bernoulli Distribution X ~ Ber (*θ*)

$$f(x;\theta) = p(x) = \begin{cases} \theta^x (1-\theta)^{1-x} & , \; x = 0,1 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \theta \qquad \operatorname{var}(X) = \theta(1-\theta)$$

3) Binomial Distribution X ~ Bin (*n*, *θ*)

$$f(x;n,\theta) = \begin{cases} C_x^n \, \theta^x (1-\theta)^{n-x} & , \; x = 0,1,2,\ldots.,n \\ 0 & o.w \end{cases}$$

$$mean = E(X) = n\theta \qquad \operatorname{var}(X) = n\theta(1-\theta)$$

4) Negative Binomial Distribution X ~ N.Bin (*r*, *θ*)

$$f(x;r,\theta) = \begin{cases} C_x^{x+r-1} \theta^r (1-\theta)^x & , \; x = 0,1,2,\ldots.. \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \frac{r(1-\theta)}{\theta} \quad , \quad v(X) = \frac{r(1-\theta)}{\theta^2}$$

5) Geometric Distribution X ~ Geo (*θ*)

$$f(x;\theta) = \begin{cases} \theta(1-\theta)^x & , \; x = 0,1,2,\ldots\ldots.. \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \frac{(1-\theta)}{\theta} \quad , \quad v(X) = \frac{(1-\theta)}{\theta^2}$$

6) The Poisson Distribution X ~ Poi (*θ*)

$$f(x;\theta) = \begin{cases} \dfrac{e^{-\theta}\theta^x}{x!} & , \; x = 0,1,2,\ldots. \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \theta \quad , \quad v(X) = \theta$$

Second: The Continuous Distributions

1. Continuous Uniform Distribution X ~ C.U(a , b)

$$f(x;\, a\,,\, b) = \begin{cases} \dfrac{1}{b-a} & ,\ a \le x \le b \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \frac{a+b}{2} \qquad , \quad v(X) = \frac{(b-a)^2}{12}$$

2. Beta Distribution X ~ *Beta* (α, *β*)

$$f(x;\, \alpha\,,\, \beta) = \begin{cases} \dfrac{1}{B(\alpha\,,\, \beta)}\, x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \frac{\alpha}{\alpha+\beta} \qquad , \quad v(X) = \frac{\alpha\,\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

3. Gamma Distribution

a) Gamma, Distribution

1. X ~ Γ(α, *θ*)

$$f(x;\, \alpha\,, \theta) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\,\theta^{\alpha}}\, x^{\alpha-1} e^{-x/\theta} & ,\ x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \alpha\theta \qquad , \qquad v(X) = \alpha\theta^2$$

2. X ~ Γ(α, 1/*θ*)

$$f(x;\, \alpha\,, \theta) = \begin{cases} \dfrac{\theta^{\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1} e^{-\theta x} & ,\ x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \alpha/\theta \qquad , \qquad v(X) = \alpha/\theta^2$$

b) Exponential Distribution

1. X ~ Exp(*θ*)

$$f(x;\, \theta) = \begin{cases} \dfrac{1}{\theta}\, e^{-x/\theta} & ,\ x > 0 \\ 0 & o.w \end{cases}$$

$$mean = E(X) = \theta \qquad , \qquad v(X) = \theta^2$$

2. X ~ Exp(1/θ)

$$f(x;\,\theta) = \begin{cases} \theta\, e^{-\theta x} & , \quad x > 0 \\ 0 & \quad o.w \end{cases}$$

$$mean = E(X) = 1/\theta \qquad , \qquad v(X) = 1/\theta^2$$

c) Chi-Square Distribution X ~ $\chi^2_{(r)}$

$$f(x;r) = \begin{cases} \dfrac{1}{\Gamma(r/2)\, 2^{r/2}}\, x^{\frac{r}{2}-1} e^{-x/2} & , \quad x > 0 \qquad , \; r > 0 \\ 0 & \quad o.w \end{cases}$$

$$mean = E(X) = r \qquad , \qquad v(X) = 2r$$

4. Normal (Gaussian) Distribution X ~ N(θ, σ^2)

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

$$mean = E(X) = \theta \qquad , \qquad v(X) = \sigma^2 \qquad , \; -\infty < \theta < \infty \quad , \quad \sigma > 0$$

5. Log- Normal Distribution X ~ LogN(θ, σ^2)

If a r.v. X has a p.d.f.;

$$f(\log x;\theta,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma\, x}\, e^{-\frac{1}{2\sigma^2}(\log\, x\, -\, \theta)^2} \qquad , \quad x > 0$$

With parameters $(-\infty < \theta < \infty)$ and $(\sigma > 0)$.

$$E(X) = \exp\left(\theta + \frac{\sigma^2}{2}\right) \qquad , \; Var(X) = \left(\exp(\sigma^2) - 1\right)\, \exp\left(2\,\theta + \sigma^2\right)$$

Chapter Two

Distributions of Functions of Random Variables

Transformations of the Discrete Random Variables

If X is a discrete r.v., having p.d.f. $f(x)$, taking values in sample space $\mathcal{S}$, $A = \{ x;\ x = x_1, x_2, \ldots, x_n\}$, at each of which $f(x) > 0$, and let a r.v. $y = g(x)$ define a **one-to-one** transformation that maps A onto B, $B = \{y;\ y = y_1, y_2, \ldots, y_n\}$.
If we solve $y = g(x)$ for x in terms of y, say, $x = w(y)$, then for each $y \in B$, we have $x = w(y) \in A$.
Consider the r.v. $Y = g(X)$. If $y \in B$, then $x = w(y) \in A$. Then to find the p.d.f. of Y, we simple substitute into (1-1) case as follows;

$$f(y) = p(Y = y) = p(X = w(y)) = \begin{cases} f[w(y)] & y \in B \\ 0 & o.w \end{cases}$$

Ex: Let $X \sim poi(\theta)$ and $Y = 4X$ by using transformation technique, find the p.d.f. of Y.
Sol:

Ex: Let X have the binomial p.d.f. . $X \sim b(3\,,\,2/3)$, where $Y = X^2$, by using one-to-one transformation, find the p.d.f. of Y.

Ex: Let X have a p.d.f.

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

Find; the p.d.f. of Y = 2X + 1.

Ex: Let X have a p.d.f.

$$f(x) = \begin{cases} \left(\frac{1}{2}\right)^x & , \quad x = 1,2,3,\ldots\ldots \\ 0 & \quad o.w \end{cases}$$

Find; the p.d.f. of $Y = X^3$.

Definition for the J.P.D.F.

Let $f(x_1, x_2)$ be the j.p.d.f. of two discrete r.v.'s X_1 and X_2 with A the (two dimensional) set of points. Which $f(x_1, x_2) > 0$, let $y_1 = g_1(x_1, x_2)$ and $y_2 = g_2(x_1, x_2)$ define a one-to-one transformation that maps A onto B (two dimensional), then the j.p.d.f. of *the two new r.v.'s* $Y_1 = g_1(X_1, X_2)$, $Y_2 = g_2(X_1, X_2)$ is given;

$$f(y_1, y_2) = p(Y_1 = y_1\,,\, Y_2 = y_2) = \begin{cases} f\left[w_1(y_1, y_2)\,,\, w_2(y_1, y_2)\right] & , \quad (y_1, y_2) \in B \\ 0 & \quad o.w \end{cases}$$

Ex: Let X_1 and X_2 be two stochastically independent r.v.'s that have Poisson distribution with means θ_1 , θ_2 respectively, the j.p.d.f. of X_1 and X_2 is;

$$f(x_1, x_2) = \begin{cases} \dfrac{\theta_1^{x_1}\, \theta_2^{x_2}\, e^{-\theta_1 - \theta_2}}{x_1!\, x_2!} & , \quad x_1 = 0,1,2,3,\ldots.. \quad , \quad x_2 = 0,1,2,3,\ldots.. \\ 0 & \quad o.w \end{cases}$$

Where $Y_1 = X_1 + X_2$, $Y_2 = X_2$. **Find:** the j.p.d.f. of Y_1 and Y_2. and $f_1(y_1)$.
HW: $f_2(y_2)$

Ex: Let X_1 and X_2 have a joint p.d.f as follows;

$$f(x_1, x_2) = \begin{cases} \left(\frac{2}{3}\right)^{x_1 + x_2} \left(\frac{1}{3}\right)^{2 - x_1 - x_2} & , \quad (x_1, x_2) = (0,0), (0,1), (1,0), (1,1) \\ 0 & o.w \end{cases}$$

Find the joint p.d.f of $Y_1 = X_1 - X_2$ and $Y_2 = X_1 + X_2$ and the marginal p.d.f of Y_1 and Y_2.

Transformations of the Continuous Random Variables

Definition: Let X be continuous r.v. having a p.d.f of $f(x)$. Let A be the one-dimension space $A = \{x : x \in R(x)\}$, where $f(x) > 0$. *Consider the r.v.* $Y = g(X)$, where $y = g(x)$ define a one-to-one transformation that maps the set A onto the set B. Let the inverse of $y = g(x)$ be denoted by $x = w(y)$, then;

$$y = g(x) \quad \Rightarrow \quad x = w(y)$$

$$|w'(y)| = \left|\frac{dx}{dy}\right| = |J| \qquad \textit{is called the Jacobian}$$

Then the p.d.f. of the r.v. $Y = g(X)$ is given by;

$$f(y) = \begin{cases} f(w(y))|J| & , \quad y \in B \\ 0 & o.w \end{cases}$$

When we have two r.v.'s X_1 and X_2. Let $y_1 = g_1(x_1 , x_2)$ and $y_2 = g_2(x_1 , x_2)$ define a one-to-one transformation that maps a (two dimensional) set A *in the x_1x_2-plane* onto a (two dimensional) set B *in the y_1y_2-plane*. Then;

$$y_1 = g_1(x_1, x_2) \quad \Rightarrow \quad x_1 = w_1(y_1, y_2)$$
$$y_2 = g_2(x_1, x_2) \quad \Rightarrow \quad x_2 = w_2(y_1, y_2)$$

Then the j.p.d.f. of the r.v. Y_1 and Y_2 is given by;

$$f(y_1, y_2) = \begin{cases} f(w_1(y_1, y_2), w_2(y_1, y_2))|J| & , \quad (y_1, y_2) \in B \\ 0 & o.w \end{cases}$$

Then the Jacobian |J| is the determinant of order (2); $|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$

Ex: Let X be the continuous r.v., having p.d.f.;

$$f(x) = \begin{cases} 2x & , \quad 0 < x < 1 \\ 0 & \quad o.w \end{cases}$$

Let $Y = 8X^3$, **Find** the p.d.f. of Y.

Ex: Let X have the p.d.f.;

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

where $Y = -2\ln X$, **find** the p.d.f. of Y.

Ex: Let X is a uniform random variable on the interval (-2, 2), find the p.d.f. of Y; **1.** $Y = 4X + 3$. **2.** $Y = |X|$.

Ex: Let $X \sim \Gamma(r/2, \theta)$, $Y = \frac{2X}{\theta}$, find; the p.d.f. of Y.

Ex: Let X have the p.d.f.;

$$f(x) = \begin{cases} 2xe^{-x^2} & , \quad 0 < x < \infty \\ 0 & \quad o.w \end{cases}$$

, where $Y = X^2$, **find** the p.d.f. of Y.

Ex: Let X_1 and X_2 be two stochastically independent r.v.'s, which have gamma distribution, with parameters (α, θ) and (β, θ) respectively, and $Y_1 = X_1 + X_2$, $Y_2 = \frac{X_1}{X_1 + X_2}$, find the j.p.d.f. of Y_1 and Y_2, $f(y_1, y_2)$, $f(y_1)$ and $f(y_2)$.

Distribution of Order Statistics

Let $X_1, X_2, \ldots, X_n$ denote a random sample and be independent identically distributed r.v's with a p.d.f. $f(x)$, and let $Y_1 < Y_2 < \ldots < Y_n$ be their ascending ordered values, i.e.;

Y_1: is a smallest value of $(X_1, X_2, \ldots, X_n)$ (min).
Y_2: is the second smallest value of $(X_1, X_2, \ldots, X_n)$.

.

.

Y_n: the largest value of $(X_1, X_2, \ldots, X_n)$ (max).

Then Y_i $(i = 1, 2, \ldots, n)$ is called the i-th order statistic of the random sample $X_1, X_2, \ldots, X_n$. and $Y_1 < Y_2 < \ldots < Y_n$ are called the order statistics corresponding of the random sample $X_1, X_2, \ldots, X_n$.

Then the j.p.d.f. of $X_1, X_2, \ldots, X_n$ is given by;

$$f(x_1, x_2, \ldots, x_n) = f(x_1).\, f(x_2) \ldots f(x_n)$$

The j.p.d.f. of the order statistics $Y_1, Y_2, \ldots, Y_n$ is given by;

$$g(y_1, y_2, \ldots, y_n) = (n!)\; g(y_1)\; g(y_2) \ldots g(y_n)$$

$$= \begin{cases} (n!) \prod_{i=1}^{n} g(y_i) & , \quad a < y_1 < y_2 < \ldots < y_n < b \\ 0 & o.w \end{cases}$$

Explain:

Let $(n = 2)$, then we have two probabilities;

$X_1 > X_2$	or	$X_1 < X_2$
$Y_1 = X_2$		$Y_1 = X_1$
$Y_2 = X_1$		$Y_2 = X_2$

Discrete

$$g(y_1, y_2) = g(y_1 = x_2)\; g(y_2 = x_1) + g(y_1 = x_1)\; g(y_2 = x_2)$$

$$= (2!)\; g(y_1)\; g(y_2)$$

$$= \begin{cases} (2!) \prod_{i=1}^{2} g(y_i) & , \quad a < y_1 < y_2 < b \\ 0 & o.w \end{cases}$$

When $(n = 3)$

$$g(y_1, y_2, y_3) = (3!)\; g(y_1)\; g(y_2)\; g(y_3)$$

$$= \begin{cases} (3!) \prod_{i=1}^{3} g(y_i) & , \quad a < y_1 < y_2 < y_3 < b \\ 0 & o.w \end{cases}$$

Continuous

When $(n = 2)$

$$J_1 = \begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1 \quad , \quad J_2 = \begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

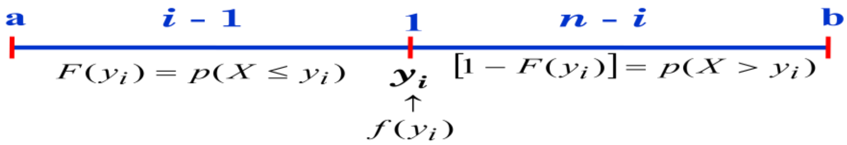
Note: Always J equal to one because be continuous

$$g(y_1, y_2) = g(y_1 = x_2)\ g(y_2 = x_1)\ J_1 + g(y_1 = x_1)\ g(y_2 = x_2)\ J_2$$

The Marginal P.D.F. of an Individual Order Statistics

The marginal p.d.f. of the i-th order statistics is given by:

$$g(y_i) = \frac{n!}{(i-1)!(n-i)!} f(y_i)\left[F(y_i)\right]^{i-1}\left[1 - F(y_i)\right]^{n-i} \quad , \quad a < y_i < b$$



The P.D.F. of the Smallest Order Statistics

If (i = 1) then the distribution of y_1 is given by:

$$g(y_1) = \frac{n!}{(1-1)!(n-1)!} f(y_1)\left[F(y_1)\right]^{1-1}\left[1 - F(y_1)\right]^{n-1} \quad , \quad a < y_1 < b$$

$$g(y_1) = \frac{n(n-1)!}{(n-1)!} f(y_1)\left[1 - F(y_1)\right]^{n-1}$$

$$g(y_1) = n\, f(y_1)\left[1 - F(y_1)\right]^{n-1} \quad , \quad a < y_1 < b$$

The P.D.F. of the Largest Order Statistics

If (i = n) then the distribution of y_n is given by:

$$g(y_n) = \frac{n!}{(n-1)!(n-n)!} f(y_n)\left[F(y_n)\right]^{n-1}\left[1 - F(y_n)\right]^{n-n} \quad , \quad a < y_n < b$$

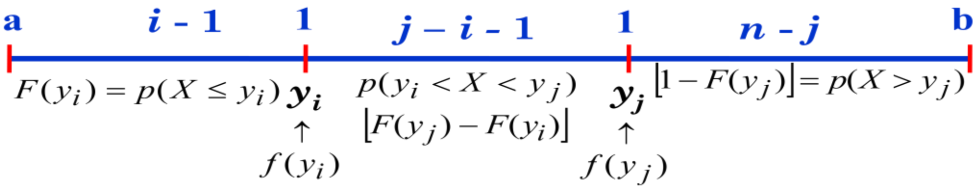
$$g(y_n) = \frac{n(n-1)!}{(n-1)!} f(y_n)\left[F(y_n)\right]^{n-1}$$

$$g(y_n) = n\, f(y_n)\left[F(y_n)\right]^{n-1} \quad , \quad a < y_n < b$$

The Joint Probability Density Fun. of Two Order Statistics

The joint p.d.f. of any two order statistics Y_i and Y_j ($i < j$) is given by:

$$g(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \times$$
$$\times \left[F(y_i)\right]^{i-1}\left[F(y_j) - F(y_i)\right]^{j-i-1}\left[1 - F(y_j)\right]^{n-j} f(y_i)\, f(y_j) \quad , \quad a < y_i < y_j < b$$
$$0 \qquad\qquad o.w$$



Ex: let $X_1, X_2, \dots, X_n$ be a random sample of size (n) rssn taken from C.U(0,1). let $Y_1 < Y_2 < \dots < Y_n$ be the order statistics of this sample. **Find** the p.d.f. of Y_1 and Y_n, the j.p.d.f. of Y_1 and Y_n

Ex: let $X_1, X_2, \dots, X_n$ be a rssn taken from Exp($1/\theta$), let $Y_1 < Y_2 < \dots < Y_n$ be the order statistics of this sample. **Find** $g(y_2)$ and $g(y_{n-1})$.

Ex: let X_1, X_2, X_3, X_4 be a random sample of size (4) from a distribution having p.d.f.;

$$f(x) = \begin{cases} 2x & , \ 0 < x < 1 \\ 0 & o.w \end{cases}$$

$Y_1 < Y_2 < Y_3 < Y_4$ denote the order statistics of random sample.
Find $g(y_3)$, and $p(Y_3 > ½)$.

Ex: let X_1, X_2, X_3, X_4 be a rss(4) from a distribution having p.d.f.;

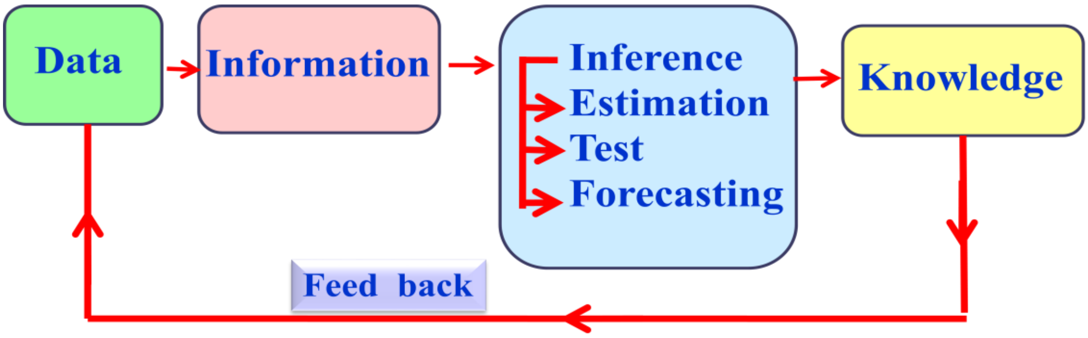
$$f(x) = \begin{cases} 1 & , \ 0 < x < 1 \\ 0 & o.w \end{cases}$$

$Y_1 < Y_2 < Y_3 < Y_4$ denote the order statistics of random sample, let $Z_1 = Y_3 - Y_1$, $Z_2 = Y_3$. **Find** $g(y_1)$, $E(Y_1)$ and $h(z_1)$.

Ex: let X_1, X_2, X_3 be a rss(3) from Beta distribution Beta(3,1), and $Y_1 < Y_2 < Y_3$ be the order statistics of this sample. **Find** $g(y_1)$, $g(y_2)$, $g(y_3)$ and $g(y_1, y_2)$.

Statistical Inference

Statistical Inference: making conclusions about the whole population on the basis of a sample, i.e., use a random sample to learn something about a large population.
Precondition for statistical inference: A sample is randomly selected from the population.



Concepts and Important Definitions about Stat. Inference

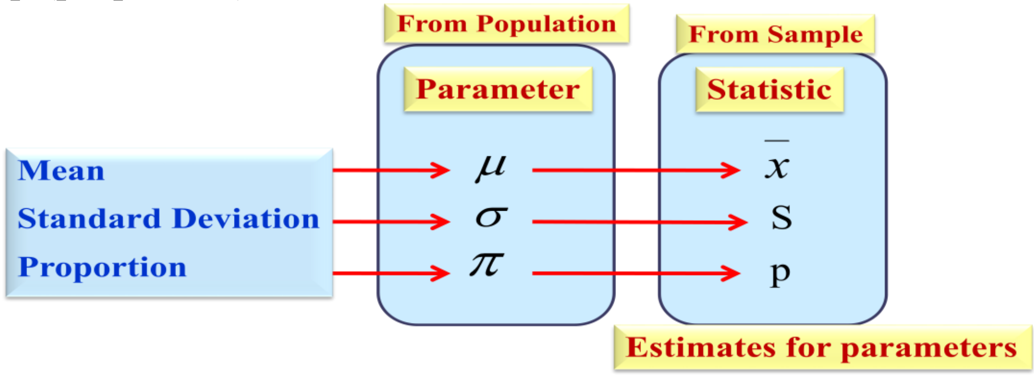
1. $\underline{X} = (X_1, X_2, \ldots, X_n) \equiv$ rssn $\equiv$ Data

2. Statistic: is a function of the random variable (r.v.) only in the sample data.

3. Parameter: It is a characteristic or a measure that is calculated from the population under study. **Ex:** The unemployment rate in Erbil. The average of assumption life for a particular device.

[Parameter = Statistic ± It's Error].

4. Population parameters are denoted using Greek letters μ (mean), σ (standard deviation), π (proportion). Sample values are denoted $\bar{x}$ (mean), S (standard deviation), p (proportion).



5. Estimator: is a function.

6. Estimate: is a value of the estimator.

$$\bar{X} = \underset{Estimator}{\frac{\sum X_i}{n}} = \underset{Estimate}{15}$$

7.

$$Quantitative\ Variable \Rightarrow Standard\ \ Error = SE(Mean) = S\ /\ \sqrt{n}$$

$$Qualitative\ Variable \Rightarrow Standard\ \ Error = SE(p) = \sqrt{p(1-p)\ /\ n}$$

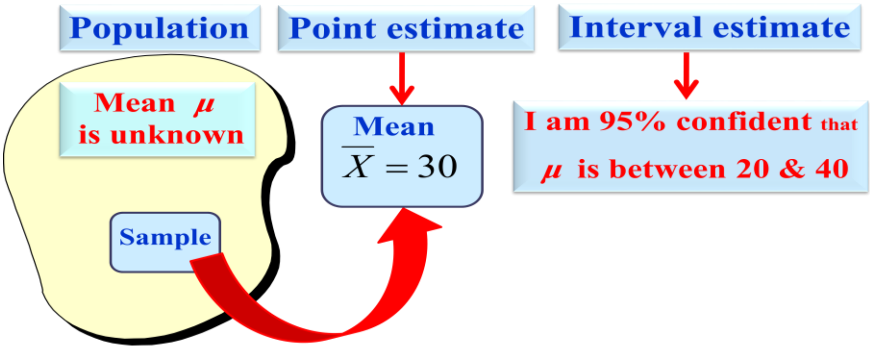
There are two steps to make inference:

1. Estimation of the population parameters

- **a)** Point Estimation.
- **b)** Intervals Estimation.

2. Testing of Hypotheses about the right of the values of population parameters.

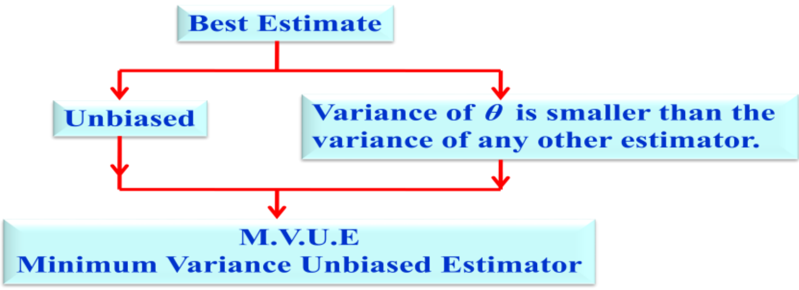
Estimation of Parameters



First: Point Estimation

Let $X_1, X_2, \ldots, X_n$ be a rssn from the p.d.f. $f(x;\theta)$, θ is unknown. We want to estimate θ from the information in the data.

$$\hat{\theta} = estimator\ \ of\ \ \theta$$



Properties of Estimator

1. Unbiased Estimator

An estimator $\left(\hat{\theta} = t(x_1, \ldots, x_n)\right)$ from a sample of size (n) with p.d.f. $f(x;\theta)$ is said to be an unbiased estimator for a population parameter θ if:

$$E(\hat{\theta}) = \theta$$

The quantity $\left(E(\hat{\theta}) - \theta\right)$ is called bias of an estimator $\hat{\theta} = t(X)$ of θ.

$$bias(\hat{\theta}) = E(\hat{\theta}) - \theta$$

Ex: In a random sample of size (n) taken from exponential dist^n Exp(θ). Show that; **1.** $T_1 = \overline{X}$ is unbiased estimator for the parameter (θ).

2. $T_2 = \frac{n}{n+1}\overline{X}^2$ is unbiased estimator for the parameter (θ^2).

Ex: In a random sample of size (n) from normal dist^n $N(\theta, \sigma^2)$. Show that;

1) $S^2 = \frac{1}{n-1}\sum(X_i - \overline{X})^2$ is unbiased estimator for the parameter (σ^2).

2) Is $T = \overline{X}^2$ unbiased estimator for θ^2.

Ex: In a random sample of size (n). Is $T = \overline{X}$ unbiased estimator for $\phi(\theta) = \theta$ of;

1. Ber(θ). **2.** Poisson(θ).

Unbiased In Limit

An estimator $\hat{\theta}$ for known parameter θ of p.d.f. $f(x;\theta)$ is unbiased in limit if:

$$\lim_{n\to\infty} E(\hat{\theta}) = \theta$$

Ex: In a rss(n) from uniform dist^n C.U$(0, \theta)$.

1) Is Y_n unbiased in limit estimator for θ; (Note: Y_n estimator θ).

2) Is $\overline{X}$ unbiased in limit estimator for θ.

3) Is $\overline{X}$ unbiased in limit estimator for $\theta/2$.

Ex: In a random sample of size (n) from normal dist^n $N(\theta, \sigma^2)$. Is $s^2 = \frac{1}{n}\sum(X_i - \overline{X})^2$ unbiased estimator for the parameter (σ^2).

2. Consistency Estimator

Def: An estimator $\hat{\theta}$ of the parameter θ of $f(x;\theta)$ is called consistent estimator for θ if;

$$\lim_{n\to\infty} p\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) = 1 \quad , \quad \forall\ \varepsilon > 0$$

$$or;\ \lim_{n\to\infty} p\left(\left|\hat{\theta} - \theta\right| \geq \varepsilon\right) = 0$$

Note: Consistency means the estimator equal to the parameter or converges stochastically to the parameter θ.

A consistent estimator: That the estimator gets closer to the parameter value as n increases without limit.

$|\hat{\theta} - \theta| \Rightarrow$ *called estimated error*

$$\left.\begin{array}{l} p\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) \geq 1 - \dfrac{v(\hat{\theta})}{\varepsilon^2} \\ p\left(\left|\hat{\theta} - \theta\right| \geq \varepsilon\right) < \dfrac{v(\hat{\theta})}{\varepsilon^2} \end{array}\right\} \rightarrow (\textit{Chebycheve inequality})$$

Th: Let $\hat{\theta}$ be an estimator for the population parameter θ of $f(x;\theta)$, then $\hat{\theta}$ is said to be consistent estimator for θ if:

$$1)\ \lim_{n\to\infty} E(\hat{\theta}) = \theta \qquad 2)\ \lim_{n\to\infty} V(\hat{\theta}) = 0$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist^n, show that $\hat{\theta} = \overline{X}$ is consistent estimator for θ.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n $N(\theta, \sigma^2)$, show that $S^2 = \frac{1}{n}\sum(X_i - \overline{X})^2$ is consistent estimator for σ^2.

Ex: Show that $\hat{\theta} = Y_n$ is consistent estimator for θ from C.U$(0, \theta)$, (by theorem).

Ex: In a rss$n$, show that $\overline{X}$ is consistent estimator for the parameter θ, from;
1) $N(\theta, \sigma^2)$. **2)** Geo(θ).

The Score Function

The score function is the partial derivative of Log the function $f(x;\theta)$ with respect to the parameter θ, is defined as;

$$S(x;\theta) = \frac{\partial}{\partial\theta}\ln f(x;\theta) = \frac{1}{f(x;\theta)}\frac{\partial}{\partial\theta} f(x;\theta)$$

Properties

1) The mean of the score is zero, $E(S(X;\theta))$ = zero

Proof:

$$E(S(X;\theta)) = \int_{R_X} s(x;\theta)\, f(x;\theta)\, dx = \int_{R_X} \frac{1}{f(x;\theta)}\frac{\partial}{\partial\theta} f(x;\theta)\, f(x;\theta)\, dx$$

$$= \int_{R_X} \frac{\partial}{\partial\theta} f(x;\theta)\, dx = \frac{\partial}{\partial\theta}\int_{R_X} f(x;\theta)\, dx = \frac{\partial}{\partial\theta}(1) = zero$$

2) The variance of the score is known as the Fisher Information (F.I), which is measure the information in the sample $\mathcal{S}$ about the parameter θ, and can be written as;

$$F.I = I(\theta) = E\left(\frac{\partial}{\partial\theta}\ln f(x;\theta)\right)^2 \qquad , because\ mean = zero$$

Or;

$$F.I = I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta)\right)$$

If Fisher Information multiply by (n), we get;

$$n\,I(\theta) = F.I \text{ in a } rss(n)$$

Ex: Let $X_1, \ldots, X_n$ be a rss$n$ from exponential dist^n Exp$(1/\theta)$. Find the F.I. of X.

3. Sufficiency Estimator

Sufficiency estimator is containing all the information in the data about the parameter θ.

First Method (Fisher Information)

Definition 1: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from the dist^n with p,d.f. $f(x\,;\,\theta)$, an estimator $\hat{\theta}$ is sufficient estimator for the parameter θ if the Fisher information in $\hat{\theta}$ is equal to the Fisher information in a rss(n).

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Bernoulli dist^n Ber(θ). Show that $\hat{\theta} = \sum X_i$ is sufficient estimator for the parameter θ.

Ex: Show that $\overline{X}$ is sufficient estimator for the mean of $N(\theta, \sigma^2)$.

Ex: In a rss$n$ from Poisson dist^n Poi(θ), is $\sum x_i$ sufficient estimator for θ?

Second Method (Conditional)

Definition 2: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from the dist^n with p,d.f. $f(x\,;\,\theta)$, and $\hat{\theta}$ be an estimator for θ, an estimator $\hat{\theta}$ is sufficient estimator for the parameter θ if the conditional p.d.f. of $(X_1, X_2, \ldots, X_n)$ given $\hat{\theta}$ does not contain the parameter θ:

$$f(x_1, x_2, \ldots, x_n \mid \hat{\theta}) = \frac{f(x_1, x_2, \ldots, x_n; \theta)}{g(\hat{\theta})}$$

Note: If the range depends on the parameter, in this case we can't find F.I; therefore, we use the second method (Conditional).

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a dist^n with p.d.f.:

$$f(x;\theta) = e^{2\theta - x} \quad , \quad x \geq 2\,\theta$$

Show that Y_1 is sufficient estimator for the parameter θ.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist^n Poi(θ), show that $\hat{\theta} = \sum X_i$ is sufficient estimator for θ?

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$. Is $\overline{X}$ sufficient estimator for θ? of;

1) Exp($1/\theta$). **2)** $N(\theta, \sigma^2)$.

Third Method: Factorization Theorem

Definition 3: Let $\hat{\theta}$ be an estimator for the parameter of $f(x;\theta)$ such that the range does not depend on θ. Then the necessary and sufficient condition for an estimator $\hat{\theta}$ to be sufficient estimator, if there are two non-negative functions, such that:

$$f(x_1, x_2, \ldots, x_n;\, \theta) = g(\hat{\theta};\, \theta) \cdot H(x)$$

Theorem:
Let $\hat{\theta}$ be sufficient estimator for the parameter θ, and $u(\hat{\theta})$ be a one-to-one transformation, then $u(\hat{\theta})$ is sufficient estimator for θ.

Note: 1) $\bar{x}$ is one to one transformation to $\sum x_i$. $\Rightarrow$ $\sum X_i = n\bar{X}$.
2) If we have more than one parameter, we use factorization theorem (third method) for sufficiency.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Bernoulli dist^n Ber(θ). Show that $\hat{\theta} = \sum X_i$ is sufficient estimator for the parameter θ.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist^n Poi(θ), show that $\hat{\theta} = \sum X_i$ is sufficient estimator for θ?

Ex: from Exp($1/\theta$). Is $\sum_{i=1}^{n} X_i$ sufficient estimator for θ? (by factorization theorem).

Ex: from Beta($\theta, \beta = 1$), $f(x;\theta) = \theta x^{\theta - 1}$, $0 < x < 1$
Is $\prod_{i=1}^{n} X_i$ sufficient estimator for θ? (by factorization theorem).

Note: $x_1^{\alpha-1} \cdot x_2^{\alpha-1} \cdot \ldots \cdot x_n^{\alpha-1} = \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}$.

H.W: Let $X_1, X_2, \ldots, X_n$ be a rss$n$. Is $\sum X_i^2$ sufficient estimator for θ? From N(0, θ).

Multi-Parameters Case (Joint Sufficient Estimator)

Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a (k) parameters dist^n $f(x;\theta_1,\theta_2,\ldots,\theta_k)$, then $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ are jointly sufficient estimators for parameters $(\theta_1,\theta_2,\ldots,\theta_k)$ respectively if the j.p.d.f. of $(X_1, X_2, \ldots, X_n)$ can be expressed as:

$$f(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k) = g(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k; \theta_1, \theta_2, \ldots, \theta_k) \cdot H(x)$$

Where; H(x) independent of the parameters $(\theta_1,\theta_2,\ldots,\theta_k)$.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Gamma dist^n $\Gamma(\alpha, 1/\theta)$, find the jointly sufficient estimators for the parameters (α, θ).

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n $N(\theta, \sigma^2)$, show that $\sum X_i$, $\sum X_i^2$ are the jointly sufficient estimators for the parameters (θ, σ^2) respectively.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from C.U$(\theta_1 - \theta_2, \theta_1 + \theta_2)$, and $Y_1 < Y_2 < \ldots < Y_n$ be the order statistics, show that Y_1 and Y_n are the jointly sufficient estimators for the parameters (θ_1, θ_2) respectively.

The Exponential Class of Probability Density Functions

Let X has a p.d.f. $f(x;\theta)$, then the family of $f(x;\theta)$ is belong to exponential class of distribution if it can be expressed as:

$$\begin{aligned} f(x;\theta) &= Exp(\ln f(x;\theta)) \\ &= Exp\big(p(\theta)\ K(x) + S(x) + q(\theta)\big) \end{aligned}$$

Such that: $p(\theta)$ $K(x)$ must have to be for exponential class.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Bernoulli dist^n Ber(θ), show that if the dist^n of X can be written in exponential form?

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist^n Poi(θ), show that if the dist^n of X can be written in exponential form?

H.W: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from exponential dist^n Exp(θ), show that if the exponential dist^n belongs to the exponential family?

H.W: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n $N(0, \theta)$, show that if the normal dist^n belongs to the exponential family?

Theorem

Let $f(x;\theta)$ belongs to exponential class of distributions, then the j.p.d.f. of $(X_1, X_2, \ldots, X_n)$ is:

$$f(x_1, x_2, \ldots, x_n; \theta) = Exp\big(p(\theta) \sum K(x_i) + \sum S(x_i) + n\, q(\theta)\big)$$

Using factorization theorem then the j.p.d.f. can be written as;

$$f(x_1, x_2, \ldots, x_n; \theta) = Exp\big(p(\theta) \sum K(x_i) + n\, q(\theta)\big) \cdot Exp\big(\sum S(x_i)\big)$$

Then we say that $\sum K(X_i)$ is minimal sufficient estimator for θ.

Ex: In a rss*n*. Find minimal sufficient estimators for parameters of:
1) Poisson(θ). **2)** Beta(α, β).

Ex: In a rss*n*. Find minimal sufficient estimators for θ from $\Gamma(2, \theta)$.

Ex: In a rss*n*. Find minimal sufficient estimators for θ , σ^2 from $N(\theta, \sigma^2)$.

The Rao-Blackwell Theorem

Let X has a p.d.f. $f(x;\theta)$, and u be an unbiased estimator for parameter θ, and T be sufficient estimator, then;

1) $E(U) = E(E(U|T))$.

2) $Var(U) \geq Var(E(U|T))$

Proof(1):

$Let;\quad E(U|T) = W$

$$E(W) \overset{?}{=} E(U)$$

$$E(W) = \int w\, f(t)\, dt = \int E(U|T)\, f(t)\, dt = \int\left(\int u\, f(u|t)\, du\right) f(t)\, dt$$

$$= \int\int u\, \frac{f(u,t)}{f(t)}\, f(t)\, dt\, du = \int u \int f(u,t)\, dt\, du = \int u\, f(u)\, du = E(U)$$

$$\therefore E(W) = E(U) = \theta$$

Proof(2):

$$Var(U) = E(U-\theta)^2 \qquad \{\mp W$$

$$= E\left((U-W)+(W-\theta)\right)^2$$

$$= E(U-W)^2 + 2E(U-W)(W-\theta) + E(W-\theta)^2$$

$$\therefore Var(U) = E(U-W)^2 + Var(W)$$

$$\Rightarrow \therefore Var(U) \geq Var(W)$$

H.W: $E(U-W)(W-\theta) \overset{?}{=} 0$

$$E(U-W)(W-\theta) = \left(E(U)-E(W)\right)(W-\theta)$$

$$= (\theta-\theta)(W-\theta)$$

$$= zero$$

Ex: Let X and Y be two random variables with j.p.d.f.;

$$f(x,y) = \frac{2}{\theta^2}\, e^{-(x+y)/\theta} \qquad , \quad 0 < x < y < \infty$$

Show that; 1) $E(Y) = E(E(Y|X))$.

2) $Var(Y) \geq Var(E(Y|X))$.

Ex: In a rss3 from C.U(0, θ). Show that [E($2Y_2$) = E{E($2Y_2$ | Y_3)}], and compare the variances of ($2Y_2$) and [E($2Y_2$ |Y_3)].

4. Completeness

Let $f(x\ ;\ \theta)$ denote a family of probability density function, let u(x) be a continuous function of (X), then if [E{u(X)}= 0] implies ($u(x)$ = 0) at each point of (X), we say that the family of p.d.f. is complete.

Note: If the range does depend on θ, then we use the general rule to derivative of integral

$$Let;\quad G(\theta) = \int_{a(\theta)}^{b(\theta)} f(x;\theta)\ dx \qquad , where \quad f : is\ \ any\ \ function$$

$$\frac{\partial G(\theta)}{\partial \theta} = \int_{a(\theta)}^{b(\theta)} \frac{\partial\ f(x;\theta)}{\partial \theta}\ dx + f\big(b(\theta),\ \theta\big) \times b'(\theta) - f\big(a(\theta)\ ,\ \theta\big) \times a'(\theta)$$

Ex: Let X be a random variable from; **1)** Bernoulli $dist^n$. **2)** Poisson $dist^n$ **3)** Normal $dist^n$. Show that the family of X is complete.

Ex: Let X be a r.v. with p.d.f.;

$$f(x;\theta) = \frac{1}{\theta} \qquad , \qquad 0 < x < \theta \qquad , \qquad \theta > 0$$

Show that $f(x;\ \theta)$ is complete?

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ a $dist^n$ with p.d.f.;

$$f(x;\theta) = e^{-(x-\theta)} \qquad , \quad \theta < x < \infty$$

Show that Y_1 is complete.

5) Uniqueness Estimator (M.V.U.E)

Th: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a $dist^n$ with p.d.f. $f(x\ ;\ \theta)$, let Y_1 be a sufficient estimator for θ, and let g(y_1; θ) be complete if there is a continuous function of Y_1 which is an unbiased estimator for θ, $\phi(\theta)$ such that $E\big(\phi(\theta)\big) = \theta$, then $\phi(\theta)$ is the unique best estimator for θ (M.V.U.E).

Proof: (In case order statistics)

$\because Y_1$ *has complete p.d.f.*

Let; $\phi(y_1)$ *be an estimator for* θ.

and $\psi(y_1)$ *be another an estimator for* θ.

$$E(\phi(y_1)) = \theta$$

$$\underline{E(\psi(y_1)) = \theta} \qquad \textit{by subtraction two functions, we get;}$$

$$E(\phi(y_1)) - E(\psi(y_1)) = \theta - \theta$$

$$E(\phi(y_1) - \psi(y_1)) = 0$$

Let $u(y_1) = \phi(y_1) - \psi(y_1)$

$\because Y_1$ *has complete p.d.f.*

$\therefore\ u(y_1) = 0$

$\Rightarrow \phi(y_1) - \psi(y_1) = 0$

$\therefore \phi(y_1) = \psi(y_1)$

Note: If an estimator does not complete then we does not find the unique and if have complete then we find a unique estimator.

Ex: Let X be a r.v. with p.d.f.;

$$f(x;\theta) = \frac{1}{2\theta} \quad , \quad -\theta < x < \theta \quad , \quad \theta > 0$$

Show that $f(x;\theta)$ is not complete? If it is then find the unique estimator for θ.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poisson dist^n poi(θ). Show that $Y = \sum X_i$ is a complete sufficient estimator for θ. Find the unique continuous function of Y which is the best estimator for θ (M.V.U.E).

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Bernoulli dist^n Ber(θ). Show that $Y = \sum X_i$ is a complete sufficient estimator for θ. Find the unique continuous function of Y which is the best estimator for θ (M.V.U.E).

Ex: Let $X_1, X_2, \ldots, X_n$ is a rss$n$ from Gamma dist^n $\Gamma(4, \theta)$, $0 < \theta < \infty$. **1)** Show that $Y = \sum X_i$ is a complete sufficient estimator for θ. **2)** Find the unique continuous function of Y which is the best estimator for θ (M.V.U.E).

Ex: Let $X_1, X_2, \ldots, X_n$ denote a random sample of size $n > 2$ from a dist^n with p.d.f. $f(x;\theta) = \theta e^{-\theta x} \quad 0 < x < \infty$ *, and* $\theta > 0$. **1)** Show that $Y = \sum X_i$ is a complete sufficient estimator for θ. **2)** Prove that (n – 1)/Y is the unique continuous function of Y which is the best estimator for θ (M.V.U.E).

Sol:

Ex: (Functions of Parameter): Let $X_1, X_2, \dots, X_n$ denote a random sample from a dist^n which is Ber(1, θ), find the best estimator for the variance $n\theta(1-\theta)$ of $Y = \sum X_i$ (M.V.U.E).

Ex: (Functions of Parameter): Let $X_1, X_2, \dots, X_n$ denote a random sample from a dist^n which is N(0, θ). Then $Y = \sum X_i^2$ is a sufficient estimator for θ. Find the best estimator for θ^2 (M.V.U.E).

The Rao- Cramer Inequality

Let $X_1, X_2, \dots, X_n$ be a rss$n$ from a dist^n with p.d.f. $f(x;\theta)$, and let $T = u(X_1, X_2, \dots, X_n)$ be an unbiased estimator for $\phi(\theta)$, then the variance of T satisfies the inequality;

$$V(T) \geq \frac{(\phi'(\theta))^2}{n\,E\left(\frac{\partial \ln f(x;\theta)}{\partial\theta}\right)^2}$$

Proof:

$$E(T) = \phi(\theta)$$

$$E(T) = \int t\ f(x;\theta)\,dx = \phi(\theta)$$

$$\frac{\partial E(T)}{\partial\theta} = \int t\,\frac{\partial f(x;\theta)}{\partial\theta}\,dx = \phi'(\theta)$$

$$= \int\left(t\,\frac{1}{f(x;\theta)}\frac{\partial f(x;\theta)}{\partial\theta}\right) f(x;\theta)\,dx = \phi'(\theta)$$

$$= \int t\,\frac{\partial \ln f(x;\theta)}{\partial\theta}\ f(x;\theta)\,dx = \phi'(\theta)$$

$$= \int t\,s\ f(x;\theta)\,dx = \phi'(\theta)$$

$$= E(T\,S) = \phi'(\theta)$$

$$Cov(T\,S) = E(T\,S) - E(T)\,E(S) \qquad , \ E(S) = 0$$

$$Cov(T\,S) = E(T\,S) = \phi'(\theta)$$

$$\rho_{T\,S}^2 = \frac{Cov^2(T\,S)}{V(T)\,V(S)}$$

$$Cov^2(T\,S) = \rho_{T\,S}^2\,V(T)\,V(S)$$

$$\therefore Cov^2(T\,S) \leq V(T)\,V(S)$$

$$Cov(T\,S) = \phi'(\theta) \quad \Rightarrow \quad Cov^2(T\,S) = (\phi'(\theta))^2$$

$$(\phi'(\theta))^2 \leq V(T)\,V(S)$$

$$V(T) \geq \frac{(\phi'(\theta))^2}{V(S)} \qquad , \ V(S) = F.I \qquad , \ \textit{its proved}$$

Notes:

1)

$$\frac{(\phi'(\theta))^2}{V(S)} \text{ is called Rao–Cramer Lower Bound (RCLB) (Minimum variance bound (MVB))}$$

2) If T unbiased estimator for θ, $E(T) = \theta$;

$\phi(\theta) = \theta \quad \to \quad \phi'(\theta) = 1$

$$\therefore \left(RCLB = \frac{1}{V(S)} \right)$$

3)

$$V(T) \geq \frac{(\phi'(\theta))^2}{nE\left(\frac{\partial \ln f(x;\theta)}{\partial\theta}\right)^2} = \frac{(\phi'(\theta))^2}{-nE\left(\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right)}$$

4) In normal distribution case, we apply the second law is easier.

5) We do not use (n) in case using the likelihood function in law.

6. Efficient Estimator

Def^n: The ratio of the RCLB to the actual variance of any unbiased estimator for θ is called the efficiency;

$$eff = \frac{RCLB}{V(T)} \quad , \quad 0 \leq eff \leq 1$$

$if \ eff = 1 \quad \Rightarrow \ T$ *is called efficient estimator for* θ.

Def^n: Let T be an unbiased estimator for $\phi(\theta)$, then we say that T is an efficient estimator for θ iff;

$$V(T) = RCLB$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Poison dist^n Poi(θ), if $T = \overline{X}$ is an efficient estimator for $\phi(\theta) = \theta$.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from exponential dist^n Exp(θ);

1) If $T = \overline{X}$ is an efficient estimator for $\phi(\theta) = \theta$.

2) Find RCLB for each of $[\phi(\theta) = \ln\theta, \phi(\theta) = 2\theta]$.

Ex: In a rss$n$ from N(θ, σ^2). Show that;

1) If $T = \overline{X}$ is an efficient estimator for $\phi(\theta) = \theta$.

2) $S^2 = \dfrac{\Sigma(x_i - \overline{x})^2}{n}$ or $S^2 = \dfrac{\Sigma(x_i - \overline{x})^2}{n-1}$ is an efficient estimator for $\phi(\sigma^2) = \sigma^2$.

Mean Square Error (MSE)

One way of measuring the accuracy of an estimator is via its mean square error. The mean square error of an estimator $\hat{\theta}$ is defined as:

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = Var(\hat{\theta}) + b^2(\hat{\theta})$$

Proof:

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

$$= E\left(\hat{\theta} - \theta \mp E(\hat{\theta})\right)^2$$

$$= E\left(\{\hat{\theta} - E(\hat{\theta})\} + \{E(\hat{\theta}) - \theta\}\right)^2$$

$$= E\left(\hat{\theta} - E(\hat{\theta})\right)^2 + E\left(E(\hat{\theta}) - \theta\right)^2 + 2\,E\left(\hat{\theta} - E(\hat{\theta})\right)\left(E(\hat{\theta}) - \theta\right)$$

$$= Var(\hat{\theta}) + b^2(\hat{\theta}) + zero$$

$$\therefore MSE(\hat{\theta}) = Var(\hat{\theta}) + b^2(\hat{\theta})$$

Note: If $\hat{\theta}$ is unbiased estimator for θ then; $MSE(\hat{\theta}) = Var(\hat{\theta})$

Relative Efficient Estimator

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two estimators for parameter θ of $f(x\,;\,\theta)$, the relative efficient of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is given by:

$$R.Eff.(\hat{\theta}_1 | \hat{\theta}_2) = \frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} < 1$$

$$i.e.,\quad MSE(\hat{\theta}_1) < MSE(\hat{\theta}_2)$$

$$\therefore \hat{\theta}_1 \text{ is more efficient than } \hat{\theta}_2.$$

Ex: In a rss2 from Bernoulli dist^n Ber(θ), let $T_1 = X_1$ and $T_2 = \frac{\sum X_i}{n+1}$ be two estimators for parameter θ, show that which of them more efficient.

Ex: In a rssn from normal dist^n N(θ, σ^2), let $S_1^2 = \frac{\sum (X_i - \overline{X})^2}{n}$ and $S_2^2 = \frac{\sum (X_i - \overline{X})^2}{n-1}$ be two estimators for parameter σ^2, show that which of them more efficient.

Ex: Given $f(x;\theta)=1/\theta$, $0 < x < \theta$, *with* $\theta > 0$, formally compute the reciprocal of; $n\,E\left(\left\{\frac{\partial \ln f(X;\theta)}{\partial \theta}\right\}^2\right)$

Compare this with the variance of $(n+1)\,Y_n / n$, where Y_n is the largest item of a random sample of size (n) from this distribution (*nth* order statistic)

Methods of Estimation

First: Maximum Likelihood Estimation (MLE)

Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a dist^n with a p.d.f. $f(x;\theta)$, the joint p.d.f. of $X_1, X_2, \ldots, X_n$ denote $L(\theta)$ is called the likelihood function, and the value of $\hat{\theta}$ which maximizes the likelihood function is called Maximum Likelihood Estimator (MLE) for θ, or the m.l.e is solution of:

$$j.p.d.f.(x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n;\theta) = L(x_1, x_2, \ldots, x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$\left(\frac{\partial \ln L(\theta)}{\partial \theta} = 0 \quad , \quad with \quad \frac{\partial^2 \ln L(\theta)}{\partial \theta^2} < 0\right)$$

Note: If the second derivative less than zero that be the maximum.

The Steps of Maximum Likelihood Estimation

1) $Find\;\; L(x_1, x_2, \ldots, x_n;\theta) = L(\underline{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta).$
2) $Find\;\; \ln\left(L(\underline{x};\theta)\right).$
3) $\frac{\partial \ln\left(L(\underline{x};\theta)\right)}{\partial \theta} = 0.$
4) $Find\;\; \hat{\theta}.$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n $N(\theta, 1)$, find the m.l.e for θ.
Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from Binomial dist^n $Bin(m, \theta)$, find m.l.e for θ.

Invariance Property of the (m.l.e)

In a rss$n$ from a dist^n with p.d.f. $f(x;\theta)$, let $\hat{\theta}$ be a m.l.e. for the parameter θ, and $u(\theta)$ be a (one-to-one) function of θ, then $u(\hat{\theta})$ is a m.l.e. for $u(\theta)$.

Ex: In a rss$n$ from exponential dist^n $Exp(1/\theta)$, find m.l.e for:

1) $u_1(\theta) = \frac{1}{\theta}$ 2) $u_2(\theta) = \frac{\ln(\theta)}{\theta}$

Remarks:

1) The m.l.e. $\hat{\theta}$ is a function of the sufficient estimator.
2) The m.l.e. $\hat{\theta}$ is not always unbiased estimator for θ.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n N(θ, σ^2), **1)** find m.l.e for parameters θ and σ^2. **2)** If S^2 is m.l.e. for σ^2, then find m.l.e. for σ.

Second: Moments Estimation Method (MEM)

Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from a dist^n with a p.d.f. $f(x;\theta)$, the average value of the kth powers of ($X_1, X_2, \ldots, X_n$); $m_k = \frac{\sum X_i^k}{n}$ is the kth sample moment , M_k = E(X^k) is the kth population moment about origin. The moment's method estimator is the value of the unknown parameter $\hat{\theta}$ that makes:

$$m_k = M_k$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n N(θ, σ^2), estimate the parameters θ and σ^2 using moment method.

Ex: In a rss$n$ from a dist^n with p.d.f.; $f(x;\theta) = (\theta+1)\, x^{\theta}\ ,\ 0 < x < 1$, estimate the parameter θ using moment method.

Ex: Estimate the parameters of $\Gamma(\alpha, 1/\theta)$, using moment method.

Ex: Estimate the parameter by using moment method for:
1) Ber(θ). **2)** Exp($1/\theta$). **3)** Geo(θ).

Third: Minimum Variance Method (MVM)

Let L(θ) be the likelihood function of a rssn with p.d.f. $f(x;\theta)$, then the parameter θ has minimum variance unbiased estimator (m.v.u.e.) if it is possible to express $\left(\frac{\partial}{\partial\theta}\ln L(\theta)\right)$ in the following form;

$$\frac{\partial}{\partial\theta}\ln L(\theta) = \frac{\hat{\theta}-\theta}{V(\hat{\theta})}$$

Where; $\hat{\theta}$: is (m.v.e.) , V($\hat{\theta}$): is variance of $\hat{\theta}$.

Ex: In a rssn, find m.v.e. for the parameters of; **1)** Ber(θ). **2)** N(θ, σ^2).

Fourth: Bayesian Estimation Method (BEM)

Philosophy: Observed data X is fixed, and the unknown generating parameter θ is random. (Certainty about θ depends on both empirical information X and prior knowledge about θ).

In Bayesian estimation method the parameters treats as a random variable with prior probability p(θ), or we have prior informative about the parameter θ.

Let A and B be two events, then the conditional probability of A given B is;

$$p(A \mid B) = \frac{p(A\, B)}{p(B)} = \frac{p(B \mid A)\ p(A)}{p(B)}$$

Let; A = θ and B = x, then in a rssn with p.d.f. $f(x;\theta)$ and prior probability p(θ);

$$p(\theta \mid x) \ = \frac{p(x \mid \theta)\ p(\theta)}{p(x)}$$

p(x) does not contain θ, we can write it as;

$$\begin{aligned} p(\theta \mid x) &\propto p(x \mid \theta)\ p(\theta) \\ &\propto L(\theta)\ p(\theta) \end{aligned}$$

Where;

p(θ | x): is called posterior probability and Bayes estimator denote $\hat{\theta}_{Bayes}$ is the mean of posterior probability E(θ | X).

L(θ): is likelihood function.

p(θ): is prior probability.

We have two types of prior probability:

1) Non Informative prior probability (Jeffery's rule).

2) Informative prior probability.

First: Non Informative prior probability (Jeffery's rule)

It is proportional to the square root of Fisher information;

$$p(\theta) \propto \left(I_S(\theta)\right)^{1/2} \qquad , \ I_S = F.I.$$

Ex: Find Bayes estimator for parameter of; **1)** Exp(1/θ). **2)** Ber(θ)., using non informative prior probability.

Ex: Find Bayes estimator for parameters of; **1)** N(θ, σ^2). **2)** Poisson(θ)., using non informative prior probability.

Second: Informative prior probability

The form of prior probability for parameters of some dist^n as follows:

ID	Probability Distribution	Informative Prior Probability

1	Bernoulli ~ Ber(θ)	Beta ~ Beta(α_o, β_o)
2	Binomial ~ Bin(n ,θ)	Beta ~ Beta(α_o, β_o)
3	Geometric ~ Geo(θ)	Beta ~ Beta(α_o, β_o)
4	Poisson ~ Poi(θ)	Gamma ~ $\Gamma(\alpha_o, \beta_o)$
5	Exponential ~ Exp($1/\theta$)	Gamma ~ $\Gamma(\alpha_o, \beta_o)$
6	Exponential ~ Exp(θ)	Inverse Gamma ~ $\Gamma^{-1}(\alpha_o, \beta_o)$
7	Normal ~ N(θ, σ^2) (θ known)	Inverse Gamma ~ $\Gamma^{-1}(\alpha_o/2, \beta_o/2)$
8	Normal ~ N(θ, σ^2) (σ^2 known)	Normal ~ N(θ_o, σ_o^2)

Ex: Estimate the parameters of; **1)** Geo(θ). **2)** Poisson (θ). **3)** Exp(θ). **4)** N(θ, σ^2) (θ known) and (σ^2 known)., using Bayesian informative prior probability.

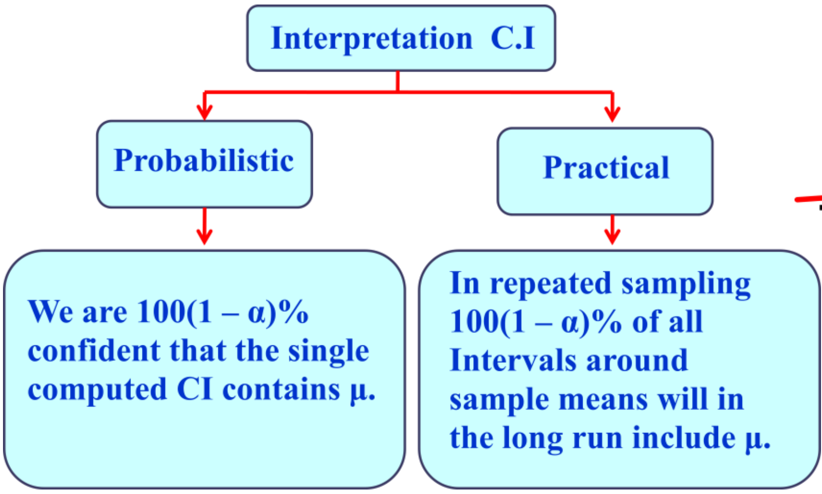
Interval Estimation

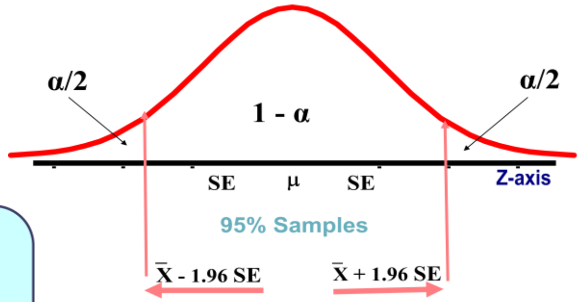
Definition:

In a rssn from a dist^n with p.d.f. $f(x, \theta)$, let L_1 and L_2 be two statistics, then the confidence interval (CI) of parameter θ is;

$$p(L_1 \le \theta \le L_2) = 1 - \alpha$$

With $100(1 - \alpha)\%$ confidence coefficient, where; L_1: is lower confidence limit. L_2: is upper confidence limit.





1) Confidence Interval for Mean when the Variance is known

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with unknown θ, and known variance σ^2, then the sample mean $\bar{x}$ is distributed with mean θ and the variance $\frac{\sigma^2}{n}$ and $Z = \frac{\bar{X} - \theta}{\sigma / \sqrt{n}}$ has standard normal dist^n or:

$$p\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \theta \leq \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

or;

$$\left(L_1 = \bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\ , L_2 = \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \text{ is } 100(1-\alpha)\ \%\ CI \ \text{ for } \theta.$$

$If \ \ CI = 90\% \ \ \Rightarrow z_{\alpha/2} = 1.65$

$If \ \ CI = 95\% \ \ \Rightarrow z_{\alpha/2} = 1.96$

$If \ \ CI = 99\% \ \ \Rightarrow z_{\alpha/2} = 2.58$

2) Confidence Interval for Mean when the Variance is unknown

Let $X_1, X_2, \ldots, X_n$ be a random sample from a population with unknown θ, and unknown variance σ^2, we have two cases:

a) If a sample size $n \geq 30$, $Z = \frac{\bar{X} - \theta}{S / \sqrt{n}}$ then:

$$p\left(\bar{X} - z_{\alpha/2}\frac{S}{\sqrt{n}} \leq \theta \leq \bar{X} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

or;

$$\left(L_1 = \bar{X} - z_{\alpha/2}\frac{S}{\sqrt{n}}\ , L_2 = \bar{X} - z_{\alpha/2}\frac{S}{\sqrt{n}}\right) \text{ is } 100(1-\alpha)\ \%\ CI \ \text{ for } \theta.$$

b) If a sample size $n < 30$, $T = \frac{\overline{X} - \theta}{S / \sqrt{n}}$ has t-distribution with (n – 1) df, then;

$$p\left(-t_{(\alpha/2,\, n-1)} \le T \le t_{(\alpha/2,\, n-1)}\right) = 1 - \alpha$$

$$p\left(-t_{(\alpha/2,\, n-1)} \le \frac{\overline{X} - \theta}{S / \sqrt{n}} \le t_{(\alpha/2,\, n-1)}\right) = 1 - \alpha$$

$$p\left(\overline{X} - t_{(\alpha/2,\, n-1)} \frac{S}{\sqrt{n}} \le \theta \le \overline{X} + t_{(\alpha/2,\, n-1)} \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

Ex: In arss100 taken from normal dist^n with mean θ and variance ($\sigma^2 = 225$), and found that $\overline{X}$ of the sample is (125). Find (95%) confidence interval for θ.

Sol:

$1 - \alpha = 95\%$

$1 - \alpha = 0.95$, $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$p\left(\overline{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \le \theta \le \overline{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$p\left(125 - (1.96)\frac{15}{\sqrt{100}} \le \theta \le 125 + (1.96)\frac{15}{\sqrt{100}}\right) = 1 - 0.05$$

$$p(125 - 2.94 \le \theta \le 125 + 2.94) = 0.95$$

$\therefore (122.06 \le \theta \le 127.94)$

Ex: Let $X_1, X_2, \ldots, X_9$ be a rss9 from a distribution with mean θ and variance σ^2, and ($\overline{X} = 19.74$, $S^2 = 0.65$). Find (99%) confidence interval (CI) for θ.

Sol:

$1 - \alpha = 99\%$

$1 - \alpha = 0.99$, $\alpha = 0.01$, $t_{(\alpha/2,\, n-1)} = t_{(0.005,\, 8)} = 3.355$

$$p\left(\overline{X} - t_{(\alpha/2,\, n-1)} \frac{S}{\sqrt{n}} \le \theta \le \overline{X} + t_{(\alpha/2,\, n-1)} \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

$$p\left(19.74 - (3.355)\frac{0.806}{\sqrt{9}} \le \theta \le 19.74 + (3.355)\frac{0.806}{\sqrt{9}}\right) = 1 - 0.01$$

$$p(19.74 - 0.9 \le \theta \le 19.74 + 0.9) = 0.99$$

$\therefore (18.84 \le \mu \le 20.64)$

Ex: A rss(50) taken from normal population with mean (θ) and variance σ^2, and ($\overline{X} = 5.67$, $S = 1.94$). Find (95%) confidence interval (CI) for θ.

Sol:

$1 - \alpha = 95\%$

$1 - \alpha = 0.95$, $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$p\left(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \theta \le \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$p\left(5.67 - (1.96)\frac{1.94}{\sqrt{50}} \le \theta \le 5.67 + (1.96)\frac{1.94}{\sqrt{50}}\right) = 1-0.05$$

$$p(5.67 - 0.538 \le \theta \le 5.67 + 0.538) = 0.95$$

$$\therefore (5.132 \le \mu \le 6.208)$$

Ex: An epidemiologist studied the blood glucose level of a random sample of 100 patients. The mean was 170, with a SD of 10. Find (95%) confidence interval for θ.

Sol:

$1-\alpha = 95\%$

$1-\alpha = 0.95$, $\alpha = 0.05$, $z_{\alpha/2} = z_{0.025} = 1.96$

$$p\left(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \theta \le \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$p\left(170 - (1.96)\frac{10}{\sqrt{100}} \le \theta \le 170 + (1.96)\frac{10}{\sqrt{100}}\right) = 1-0.05$$

$$p(170 - 1.96 \le \theta \le 170 + 1.96) = 0.95$$

$$\therefore (168.04 \le \mu \le 171.96)$$

3) Confidence Interval for Difference Between two Means

Let $\overline{X}$ be a sample mean for a rssn from a normal population with mean μ_X and unknown variance σ_X^2 and $\overline{Y}$ be a sample mean for a rssn from a normal population with mean μ_Y and unknown variance σ_Y^2, then:

$$\bar{X} \sim N(\mu_X , \frac{\sigma_X^2}{n}) \quad , \quad \bar{Y} \sim N(\mu_Y , \frac{\sigma_Y^2}{m})$$

$$(\bar{X} - \bar{Y}) \sim N\left(\mu_X - \mu_Y , \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$$

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

$$p(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}) = 1 - \alpha$$

$$p\left(-Z_{\alpha/2} \le \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \le Z_{\alpha/2}\right) = 1 - \alpha$$

$$p\left((\bar{X} - \bar{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le (\mu_X - \mu_Y) \le (\bar{X} - \bar{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right) = 1 - \alpha$$

Ex: Let $\bar{X}$ be a sample mean for a rss15 from a normal population with mean μ_X and known variance $\sigma_X^2 = 60$ and $\bar{Y}$ be a sample mean for a rss18 from a normal population with mean μ_Y and known variance $\sigma_Y^2 = 40$, we find that $(\bar{X} = 70.1)$, $(\bar{Y} = 75.3)$, find 90% CI for $(\mu_X - \mu_Y)$.

Sol:

$$\alpha = 0.1 \quad , \quad \frac{\alpha}{2} = 0.05 \quad , \quad Z_{\alpha/2} = Z_{0.05} = 1.645$$

$$\bar{X} - \bar{Y} = 70.1 - 75.3 = -5.2$$

$$p\left((\bar{X} - \bar{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le (\mu_X - \mu_Y) \le (\bar{X} - \bar{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right) = 1 - \alpha$$

$$p\left(-5.2 - 1.645\sqrt{\frac{60}{15} + \frac{40}{18}} \le (\mu_X - \mu_Y) \le -5.2 + 1.645\sqrt{\frac{60}{15} + \frac{40}{18}}\right) = 1 - 0.1$$

$$p(-9.303 \le (\mu_X - \mu_Y) \le -1.097) = 0.9$$

$$(-9.303 \le (\mu_X - \mu_Y) \le -1.097)$$

4) Confidence Interval for the Variance

Let $X_1, X_2, \ldots, X_n$ be a random sample from normal population with unknown mean, and unknown variance, then;

$$\chi^2 = \frac{(n-1)\,S^2}{\sigma^2} \quad is\ distributed\ as\ \chi^2\ with\ (n-1)\ d.f.$$

$$p\left(\chi^2_{\frac{\alpha}{2},\,n-1} \le \chi^2 \le \chi^2_{1-\frac{\alpha}{2},\,n-1}\right) = 1-\alpha$$

$$p\left(\chi^2_{\frac{\alpha}{2},\,n-1} \le \frac{(n-1)\,S^2}{\sigma^2} \le \chi^2_{1-\frac{\alpha}{2},\,n-1}\right) = 1-\alpha$$

$$p\left(\frac{1}{\chi^2_{\frac{\alpha}{2},\,n-1}} \ge \frac{\sigma^2}{(n-1)\,S^2} \ge \frac{1}{\chi^2_{1-\frac{\alpha}{2},\,n-1}}\right) = 1-\alpha$$

$$p\left(\frac{(n-1)\,S^2}{\chi^2_{1-\frac{\alpha}{2},\,n-1}} \le \sigma^2 \le \frac{(n-1)\,S^2}{\chi^2_{\frac{\alpha}{2},\,n-1}}\right) = 1-\alpha$$

Ex: Let $X_1, X_2, \ldots, X_{20}$ be a random sample from normal population with unknown mean, and unknown variance, we found that $(\overline{X} = 76.1\,,\ S^2 = 88.36)$, find 99% CI for σ^2.

Sol:

$\alpha = 0.01 \quad , \frac{\alpha}{2} = 0.005 \quad , \chi^2_{\frac{\alpha}{2},\,n-1} = \chi^2_{0.005,19} = 6.84$

$, \chi^2_{1-\frac{\alpha}{2},\,n-1} = \chi^2_{0.995,19} = 38.6$

$$p\left(\frac{(n-1)\,S^2}{\chi^2_{1-\frac{\alpha}{2},\,n-1}} \le \sigma^2 \le \frac{(n-1)\,S^2}{\chi^2_{\frac{\alpha}{2},\,n-1}}\right) = 1-\alpha$$

$$p\left(\frac{19\,(88.36)}{38.6} \le \sigma^2 \le \frac{19\,(88.36)}{6.84}\right) = 1-0.01$$

$$p\left(45.87 \le \sigma^2 \le 245.22\right) = 0.995$$

$$\therefore\ (45.87 \le \sigma^2 \le 245.22)$$

Note:

$\alpha = 5\% \quad \rightarrow 1-\alpha = 95\% \quad \Rightarrow Z_{0.025} = 1.96$

$\alpha = 10\% \quad \rightarrow 1-\alpha = 90\% \quad \Rightarrow Z_{0.05} = 1.645$

$\alpha = 1\% \quad \rightarrow 1-\alpha = 99\% \quad \Rightarrow Z_{0.005} = 2.58$

$\alpha = 2\% \quad \rightarrow 1-\alpha = 98\% \quad \Rightarrow Z_{0.01} = 2.326$

Ex: Let

$$X \sim N(\theta_X, \sigma_X^2) \quad , \quad Y \sim N(\theta_Y, \sigma_Y^2)$$

$$n = 10 \quad , \quad \overline{X} = 4.2 \quad , \quad \sigma_X^2 = 49$$

$$m = 7 \quad , \quad \overline{Y} = 3.4 \quad , \quad \sigma_Y^2 = 32$$

Find 90 % CI for $(\theta_X - \theta_Y)$.

Sol:

$$1 - \alpha = 0.9 \quad \Rightarrow \alpha = 0.1, \; Z_{\alpha/2} = Z_{0.05} = 1.6450.005 \quad , \; \chi^2_{\frac{\alpha}{2}, n-1} = \chi^2_{0.005,19} = 6.84$$

$$p\left((\overline{X} - \overline{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le (\theta_X - \theta_Y) \le (\overline{X} - \overline{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \right) = 1 - \alpha$$

$$p\left((4.2 - 3.4) - 1.645\sqrt{\frac{49}{10} + \frac{32}{7}} \le (\theta_X - \theta_Y) \le (4.2 - 3.4) - 1.645\sqrt{\frac{49}{10} + \frac{32}{7}} \right) = 1 - 0.1$$

$$p\left(0.8 - 5.063 \le (\theta_X - \theta_Y) \le 0.8 + 5.063\right) = 0.9$$

$$p\left(-4.263 \le (\theta_X - \theta_Y) \le 5.863\right) = 0.9$$

$$\therefore \; (-4.263 \le (\theta_X - \theta_Y) \le 5.863)$$

Ex: from $N(\theta, \sigma^2)$, we have ($n = 9$, $S^2 = 7.63$), find 95 % CI for σ^2.

Sol:

$$1 - \alpha = 0.95 \quad \to \alpha = 0.05 \quad , \; \chi^2_{\frac{\alpha}{2}, n-1} = \chi^2_{0.025,8} = 2.18$$

$$, \; \chi^2_{1-\frac{\alpha}{2}, n-1} = \chi^2_{0.975,8} = 17.5$$

$$p\left(\frac{(n-1)\, S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}} \le \sigma^2 \le \frac{(n-1)\, S^2}{\chi^2_{\frac{\alpha}{2}, n-1}} \right) = 1 - \alpha$$

$$p\left(\frac{8\,(7.63)}{17.5} \le \sigma^2 \le \frac{8\,(7.63)}{2.18} \right) = 1 - 0.05$$

$$p\left(3.488 \le \sigma^2 \le 28\right) = 0.95$$

$$\therefore \; (3.488 \le \sigma^2 \le 28)$$

Testing of Statistical Hypotheses

Hypothesis testing: A statistical method that uses sample data to evaluate a hypothesis about a population parameter. It is intended to help researchers differentiate between real and random patterns in the data.

What is a Hypothesis? It is an allegation about the population parameter. (Hypothesis is a belief concerning a parameter).
Ex: I assume the mean BP of participants is 120 mmHg!
Ex: I believe that mean weight of male students is 70 kilograms!

Null & Alternative Hypotheses

Null hypothesis is prevalent opinion, previous knowledge, basic assumption, prevailing theory,...
Null hypothesis is assumed to be true as long as we find evidence against it.
Alternative hypothesis is rival opinion.
H_0: Null hypothesis states the assumption to be tested e.g. BP of participants = 120 (H_0: $\mu = 120$).
H_1: Alternative hypothesis is the opposite of the null hypothesis (BP of participants ≠ 120 (H_1: $\mu \neq 120$). It may or may not be accepted and it is the hypothesis that is believed to be true by the researcher.

Simple Hypothesis: the statistical hypothesis completely specifies the distribution(the value of a parameter is specified).
Ex: H: $\mu = 10$, H: $\theta = 200$, H: $f(x) = e^{-x}$

Composite Hypothesis: the statistical hypothesis does not completely specifies the distribution (the value of the parameter is not specified).
Ex: H: $\theta \geq 10$, H: $\mu < 30$, H: $f(x) \neq N(0, 1)$

Level of Significance (α): Defines unlikely values of sample statistic if null hypothesis is true. Called rejection region of sampling distribution.

Type I Error (α) : pr(reject H_o | H_o is true).
Type II Error (β) : pr(accept H_o | H_o is false).
Power of Test: probability of right decision.
1. pr(accept H_o | H_o is true).
2. pr(reject H_o | H_o is false).

Result Possibilities

Hypothesis Test		
	Actual Situation	
Decision	H_o True	H_o False
Accept H_o	Power $(1 - \alpha)$	Type II Error (β) ← **False Negative**
Reject H_o	Type I Error (α) ← **False Positive**	Power $(1 - \beta)$

Hypothesis Examples

H_o: Mean height of males equals 174.
H_1: Mean height is bigger than 174.

H_o: Half of the population is in favour of nuclear power plant.
H_1: More than half of the population is in favour of nuclear power plant.

H_o: The amount of overtime work is equal for males and females.
H_1: The amount of overtime work is not equal for males and females.

H_o: There is no correlation between interest rate and gold price.
H_1: There is correlation between interest rate and gold price.

Steps of Hypothesis Testing

Test the hypothesis that the true mean BP of participants is 120 mmHg.

State H_o $\quad H_0: \mu = 120$
State H_1 $\quad H_1: \mu \neq 120$
Choose α $\quad \alpha = 0.05$
Choose n $\quad n = 100$
Choose Test: $\quad$ Z, t, X^2 *Test*
Compute Test Statistic *(or compute P value)*
Search for Critical Value
Make Statistical Decision rule
Express Decision

Definition: Let C be that subset of the sample space which, in accordance with a prescribed test, leads to the rejection of the hypothesis under consideration. Then C is called the critical region of the test.

Certain Best Tests (book p273)

Let C denote a subset of the sample space. Then C is called a best critical region of size α for testing the simple hypothesis H_o: $\theta = \theta'$ against the alternative simple hypothesis H_o: $\theta = \theta''$ if, for every subset A of the sample space for which $p((X_1, \ldots, X_n) \in A; H_o) = \alpha$,

$i)\ \ p((X_1, \ldots, X_n) \in C; H_o) = \alpha;$

$ii)\ \ p((X_1, \ldots, X_n) \in C; H_1) \geq p((X_1, \ldots, X_n) \in A; H_1)$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rssn from $f(x; \theta) = x$
H_o: $\theta = \theta_o$ (Simple Hypothesis)
H_1: $\theta = \theta_1$
$\mathbf{S} = C \cup C^c$

$C_o = C$: Critical region for H_o if $\underline{X} \in C_o$

$\Rightarrow$ reject H_o

$C_1 = C^c$: Critical region for H_1 if $\underline{X} \in C_1$

$\Rightarrow$ reject H_1

α: size of C_o = size of test , level of the test.

$$\alpha = pr(\underline{X} \in C_o | H_o) \quad , \quad \beta = pr(\underline{X} \in C_1 | H_1)$$

Ex: Let $X_1, X_2, \dots, X_n$ be a rssn from Poisson dist[n];

$$f(x;\theta) = e^{-\theta}\,\theta^{x} / x!$$

$$H_o : \theta = \theta_o \quad , \quad H_1 : \theta = \theta_1 \qquad , \quad Find: \; C_o ?$$

Sol:

$$L(\underline{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\left(e^{-\theta}\,\theta^{x} / x!\right) = e^{-n\theta}\,\theta^{\sum x_i} / \prod_{i=1}^{n} x_i!$$

$$\frac{L(\underline{x};\theta_o)}{L(\underline{x};\theta_1)} = \frac{e^{-n\theta_o}\,\theta_o^{\sum x_i} / \prod_{i=1}^{n} x_i!}{e^{-n\theta_1}\,\theta_1^{\sum x_i} / \prod_{i=1}^{n} x_i!} \le k_o$$

$$e^{-n\theta_o}\,\theta_o^{\sum x_i} / e^{-n\theta_1}\,\theta_1^{\sum x_i} \le k_o \qquad (\ln \; for \; both \; sides)$$

$$-n\theta_o + \sum x_i \ln(\theta_o) + n\theta_1 - \sum x_i \ln(\theta_1) \le \ln(k_o)$$

$$n(\theta_1 - \theta_o) + \sum x_i (\ln(\theta_o) - \ln(\theta_1)) \le \ln(k_o)$$

$$\sum x_i (\ln(\theta_o) - \ln(\theta_1)) \le \ln(k_o) - n(\theta_1 - \theta_o)$$

$$\sum x_i \le \left.\frac{\ln(k_o) - n(\theta_1 - \theta_o)}{\ln(\theta_o) - \ln(\theta_1)} \quad \right\} C$$

$$\therefore \sum x_i \le C$$

$$C_o = \{\underline{X} : \sum x_i \le C\}$$

$$\alpha = pr\{\sum x_i \le C\}$$

Neyman–Pearson Theorem

Let $X_1, X_2, \dots, X_n$ be a rssn from $f(x;\theta)$,

let H_o: $\theta = \theta_o$

H_1: $\theta = \theta_1$

Then C is the best critical region (CR) to test H_o against H_1, if;

1. $\dfrac{L(\underline{x};\theta_o)}{L(\underline{x};\theta_1)} \le k \quad , \; if \;\; \underline{X} \in C$
2. $\dfrac{L(\underline{x};\theta_o)}{L(\underline{x};\theta_1)} \ge k \quad , \; if \;\; \underline{X} \in C^c$
3. $\alpha = pr\{\underline{X} \in C \,|\, \theta = \theta_o\}$

Ex: Let X_1, X_2, … , X_n be a rssn from N(θ, 1). Find the best critical region (B.C.R) to test H_o: $\theta = 0$, H_1: $\theta = 1$, ($\alpha = 0.05$)

Sol:

$$X \sim N(\theta, 1) \;\Rightarrow\; f(x;\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right)$$

$$L(\underline{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}\sum (x_i-\theta)^2}$$

$$\frac{L(\underline{x};\theta_o)}{L(\underline{x};\theta_1)} \le k$$

$$\frac{\left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}\sum (x_i-\theta_o)^2}}{\left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}\sum (x_i-\theta_1)^2}} \le k \qquad \Rightarrow \qquad \frac{e^{-\frac{1}{2}\sum (x_i-\theta_o)^2}}{e^{-\frac{1}{2}\sum (x_i-\theta_1)^2}} \le k$$

By Hypotheses $(\theta_o = 0 \;,\; \theta_1 = 1)$

$$\frac{e^{-\frac{1}{2}\sum x_i^2}}{e^{-\frac{1}{2}\sum (x_i-1)^2}} \le k$$

$$e^{-\frac{1}{2}\sum x_i^2 + \frac{1}{2}\sum (x_i-1)^2} \le k \qquad \Rightarrow \qquad e^{\frac{-\sum x_i^2 + \sum x_i^2 - 2\sum x_i + 1}{2}} \le k$$

$$e^{\frac{-2\sum x_i + 1}{2}} \le k$$

$$\frac{-2\sum x_i + 1}{2} \le \ln(k)$$

$$-2\sum x_i \le 2\ln(k) - 1 \qquad (\div(-2))$$

$$\therefore \sum x_i \ge \frac{1}{2} - \ln(k) \quad \Big\} \; C$$

$$\therefore \sum x_i \ge C$$

$$\therefore C_o = \{\underline{X} : \sum x_i \ge C\} \qquad \text{is BCR}$$

To find C

$$\alpha = 0.05 = pr\left\{\underline{X} \in C \mid \theta = \overset{\theta_o}{0}\right\} = pr\{\sum x_i \ge C \mid \theta = 0\}$$

n; the sample size is known

$$0.05 = p\left\{\frac{\sum x_i}{n} \ge \frac{C}{n} \mid \theta = 0\right\}$$

$$0.05 = p(\overline{X} \ge C' \mid \theta = 0) \qquad\qquad note: \mu_{\overline{x}} = \mu_x \;,\; \sigma^2_{\overline{x}} = \frac{\sigma_x^2}{n}$$

when H_o *is true* $(\theta = 0)$

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

$$\overline{X} \sim N(0, \tfrac{1}{n})$$

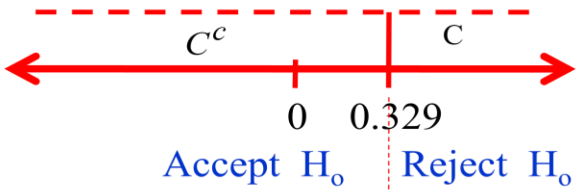
$$Let \;\; n = 25 \qquad \Rightarrow \overline{X} \sim N(0, \tfrac{1}{25})$$

$$0.05 = p\left\{Z \geq \frac{C' - 0}{1/5}\right\}$$

$$0.05 = p\{Z \geq 5C'\}$$

$$\therefore 5C' = 1.645 \qquad \Rightarrow \;\; C' = 0.329$$

$$\therefore \; C = \{\underline{X} : \overline{X} \geq 0.329\}$$



We accept if it is less than this amount and we reject if the opposite.

Ex: Let $X_1, X_2, \ldots, X_n$ be a rssn from $f(x)$. Find the best critical region (B.C.R) to test;

$$H_o \; : \; f(x) = \frac{e^{-x}}{x!} \qquad , \quad x = 0 \,, 1 \,, 2 \,, 3 \,, \ldots\ldots$$

$$H_1 \; : \; f(x) = \left(\frac{1}{2}\right)^{x+1} \qquad , \quad x = 0 \,, 1 \,, 2 \,, 3 \,, \ldots\ldots$$

H_o: $\theta = 0$, H_1: $\theta = 1$, $(\alpha = 0.05)$

Sol:

$$L(\underline{x}) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}\sum(x_i - \theta)^2}$$

$$L(\underline{x} \mid H_o) = \prod_{i=1}^{n} \frac{e^{-x_i}}{x_i!} = \frac{e^{-\sum x_i}}{\prod_{i=1}^{n} x_i!}$$

$$L(\underline{x} \mid H_1) = \prod_{i=1}^{n} \left(\frac{1}{2}\right)^{x_i + 1} = \left(\frac{1}{2}\right)^{\sum x_i + n}$$

$$\frac{L(\underline{x}|H_o)}{L(\underline{x}|H_1)} \leq k$$

$$\frac{e^{-\sum x_i} \Big/ \prod_{i=1}^{n} x_i!}{\left(\frac{1}{2}\right)^{\sum x_i + n}} \leq k \qquad \Rightarrow \qquad \frac{e^{-\sum x_i} \; 2^{\sum x_i + n}}{\prod_{i=1}^{n} x_i!} \leq k$$

$$-\sum x_i + (\sum x_i + n) \ln(2) - \ln\left(\prod_{i=1}^{n} x_i!\right) \leq k$$

$$\therefore \; The \;\; BCR$$

$$C = \left\{\underline{x} : -\sum x_i + (\sum x_i + n) \ln(2) - \ln\left(\prod_{i=1}^{n} x_i!\right) \leq k\right\}$$

$$Let; \quad n=1 \quad , \quad k = 1$$

$$-x + (x+1)\ln(2) - \ln(x!) \le \ln(1)$$

$$-x + x\ln(2) + \ln(2) - \ln(x!) \le 0$$

$$x\,(\ln(2)-1) - \ln(x!) \le -\ln(2)$$

$$\therefore C = \{x: -0.3\,x - \ln(x!) \le -0.7\}$$

$$C = \{0,3,4,5,.......\}$$

Ex9.14: Let $X_1, X_2, \ldots, X_n$ denote a rssn from a dist^n having the p.d.f.

$$f(x;\theta) = \theta^{x}(1-\theta)^{1-x} \quad , \quad x = 0,1$$

Show that $C = \left\{(x_1, x_2, \ldots, x_n) : \sum_{i=1}^{n} x_i \le c\right\}$ is the best critical region for testing

$$H_o : \theta = \frac{1}{2} \; aganst \; H_1 : \theta = \frac{1}{3}$$

Use the central limit theorem to find n and c so that approximately

$$\Pr\left(\sum_{i=1}^{n} X_i \le c : H_o\right) = 0.10 \;\; and \; \Pr\left(\sum_{i=1}^{n} X_i \le c : H_1\right) = 0.80$$

Sol:

من دفتر اخضر

Uniformly Most Powerful Test

Uniformly Most Powerful Test: at least one of hypotheses is composite, for instance;

$$H_o : \theta = \theta_o \quad , \quad H_1 : \theta > \theta_o$$

$$H_o : \theta > \theta_o \quad , \quad H_1 : \theta \le \theta_o$$

$$H_o : \theta = \theta_o \quad , \quad H_1 : \theta > \theta_1$$

$$H_o : \theta = \theta_o \quad , \quad H_1 : \theta = \theta_1 \quad , \quad \theta_1 > \theta_o$$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rssn from $N(0, \theta)$.

Let; $H_o: \theta = \theta_o$, $H_1: \theta > \theta_o$, Find the U.M.P.T.

Sol:

$Let; \quad \theta = \theta_1 > \theta_o$

$H_o : \theta = \theta_o \quad , \quad H_1 : \theta = \theta_1$

$$X \sim N(0, \theta) \Rightarrow f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{x^2}{2\theta} \right)$$

$$L(\underline{x};\theta) = (2\pi\theta)^{-n/2} \exp\left(-\frac{\sum x_i^2}{2\theta} \right)$$

$$L(\underline{x};H_o) = (2\pi\theta_o)^{-n/2} \exp\left(-\frac{\sum x_i^2}{2\theta_o} \right)$$

$$L(\underline{x};H_1) = (2\pi\theta_1)^{-n/2} \exp\left(-\frac{\sum x_i^2}{2\theta_1} \right)$$

$$\frac{L(\underline{x};H_o)}{L(\underline{x};H_1)} \le k \qquad \Rightarrow \quad \frac{(2\pi\theta_o)^{-n/2} \exp\left(-\frac{\sum x_i^2}{2\theta_o} \right)}{(2\pi\theta_1)^{-n/2} \exp\left(-\frac{\sum x_i^2}{2\theta_1} \right)} \le k$$

$$\left(\frac{\theta_1}{\theta_o} \right)^{n/2} \exp\left(-\frac{\sum x_i^2}{2\theta_o} + \frac{\sum x_i^2}{2\theta_1} \right) \le k$$

$$\left(\frac{\theta_1}{\theta_o} \right)^{n/2} \exp\left(\frac{\sum x_i^2}{2} \left\{ \frac{\theta_o - \theta_1}{\theta_o \theta_1} \right\} \right) \le k$$

$$\frac{n}{2} \ln\left(\frac{\theta_1}{\theta_o} \right) + \frac{\sum x_i^2}{2} \left\{ \frac{\theta_o - \theta_1}{\theta_o \theta_1} \right\} \le \ln(k)$$

$$\frac{\sum x_i^2}{2} \left\{ \frac{\theta_o - \theta_1}{\theta_o \theta_1} \right\} \le \ln(k) - \frac{n}{2} \ln\left(\frac{\theta_1}{\theta_o} \right)$$

$$\sum x_i^2 (\theta_o - \theta_1) \le 2\theta_o \theta_1 \left\{ \ln(k) - \frac{n}{2} \ln\left(\frac{\theta_1}{\theta_o} \right) \right\} \quad \Rightarrow = C^* : Constant$$

$\because \theta_1 > \theta_o \quad \Rightarrow (\theta_o - \theta_1)$ *negative then we divide on negative*

$$\therefore \sum x_i^2 \ge \frac{C^*}{\theta_o - \theta_1} = C^\bullet$$

$$\therefore C = \left\{ \underline{x} : \sum x_i^2 \ge C^\bullet \right\}$$

$$Let; \quad \alpha = 0.05 \ , \quad n = 15 \quad , \quad \theta_o \ = 3$$

$$0.05 = pr\{\underline{x} \in C / H_o\}$$

$$0.05 = pr\left\{\sum x_i^2 \geq C^\bullet / \theta_o = 3\right\}$$

$$X_i \sim N(0, \theta)$$

$$\frac{x_i - 0}{\sqrt{\theta}} \sim N(0,1)$$

$$\frac{x_i^2}{\theta} \sim \chi_1^2$$

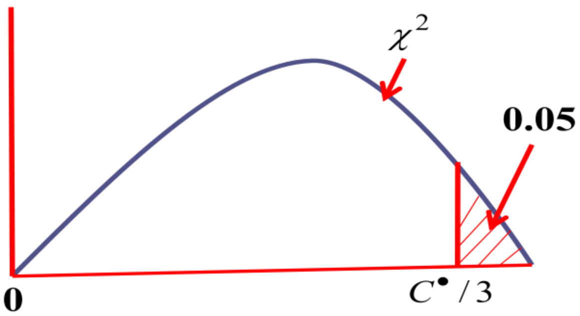
$$\frac{\sum x_i^2}{\theta} \sim \chi_n^2$$

$$0.05 = pr\left\{\frac{\sum_{i=1}^{15} x_i^2}{3} \geq \frac{C^\bullet}{3}\right\}$$

$$0.05 = pr\left\{\chi_{15}^2 \geq \frac{C^\bullet}{3}\right\}$$

$$\frac{C^\bullet}{3} = 25 \qquad \Rightarrow C^\bullet = 75$$

$$\therefore C = \{\underline{x}: \chi_{15}^2 \geq 75\} \rightarrow \ for \ \ reject$$



Ex: Let $X_1, X_2, \dots, X_n$ be a rssn from N(θ, 1).
Let: $H_o : \theta = \theta' = 0 \ , \ H_1 : \theta = \theta'' = -1$,
$\bar{x}$, $n = 25$, $\alpha = p(\theta') = 0.05$, $p(\theta'') = 0.999$ Find the U.M.P.T. Or find the critical region (C)?

Sol:

1. $\frac{L(\underline{x};\theta')}{L(\underline{x};\theta'')} \leq k \quad , \ if \ \ \underline{X} = X_1, X_2, \dots, X_n \in C$

2. $\frac{L(\underline{x};\theta')}{L(\underline{x};\theta'')} > k \quad , \ if \ \ \underline{X} = X_1, X_2, \dots, X_n \in C^c$

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}}$$

$$L(\underline{x};\theta) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum (x_i - \theta)^2}{2}}$$

$$\frac{L(\underline{x};\theta')}{L(\underline{x};\theta'')} = \frac{L(0)}{L(-1)} \le k$$

$$\frac{\frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum x_i^2}{2}}}{\frac{1}{(2\pi)^{n/2}} e^{-\frac{\sum (x_i+1)^2}{2}}} \le k$$

$$e^{-\frac{\sum x_i^2}{2}} \; e^{\frac{\sum (x_i+1)^2}{2}} \le k$$

$$e^{\frac{-\sum x_i^2 + \sum x_i^2 + 2\sum x_i + n}{2}} \le k$$

$$e^{\frac{2\sum x_i + n}{2}} \le k$$

$$\frac{2\sum x_i + n}{2} \le \ln(k)$$

$$\sum x_i + \frac{n}{2} \le \ln(k) \quad \Rightarrow \quad \sum x_i \le \ln(k) - \frac{25}{2} \Big\} = c \quad \Big\} \div n$$

$$\therefore C = r(c)$$

$$\therefore\ C = \{\underline{x}: \sum x_i \le c\}\ ,\ C = \{\underline{x}: \bar{x} \le c\}\ \textit{ is the Best C.R.}$$

3. $\alpha = pr\{\underline{X} \in C \,|H_o\}$

$$= pr\{reject\ H_o\,|H_o \text{ is true}\} \qquad , X \sim N(\theta,1)\ ,\ \bar{X} \sim N(\theta,\tfrac{1}{n})$$

$$= pr\{\bar{x} \le c|\theta = 0\}$$

$$= pr\left\{\frac{\bar{x}-\theta}{1/5} \le \frac{c-0}{1/5}\right\}$$

$$\Rightarrow N\left(\frac{c-0}{1/5}\right) = 0.05 \quad , \ c = \frac{\ln(k) - 25/2}{25}$$

$$p(\theta'') = pr\{\underline{X} \in C\,|H_1\}$$

$$= pr\left\{\frac{\bar{X}-\theta}{1/5} \le \frac{c+1}{1/5}\right\} = 0.999$$

$$\Rightarrow N\left(\frac{c+1}{1/5}\right) = 0.999$$

Ex: Let $X_1, X_2, \dots, X_n$ be a rssn from $N(0, \sigma^2)$.
Let; $H_o: \sigma^2 = 1$, $H_1: \sigma^2 = 2$, $n = 10$, Find the Best C.R.?
Sol:

$$X \sim N(0, \sigma^2) \;\Rightarrow\; f(x;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

$$L(\underline{x};\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\sum x_i^2}{2\,\sigma^2}\right)$$

$$\frac{L(1)}{L(2)} \le k \qquad \Rightarrow \qquad \frac{\frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{\sum x_i^2}{2}\right)}{\frac{1}{(4\pi)^{n/2}} \exp\left(-\frac{\sum x_i^2}{4}\right)} \le k$$

$$\frac{\exp\left(-\frac{\sum x_i^2}{2}\right)}{\left(\frac{1}{2}\right)^{n/2} \exp\left(-\frac{\sum x_i^2}{4}\right)} \le k$$

$$2^{n/2}\, e^{-\frac{\sum x_i^2}{2}}\, e^{\frac{\sum x_i^2}{4}} \le k \qquad \Rightarrow \; 2^5\, e^{\frac{-2\sum x_i^2 + \sum x_i^2}{4}} \le k$$

$$e^{-\frac{\sum x_i^2}{4}} \le 2^{-5} \times k$$

$$-\frac{\sum x_i^2}{4} \le -5\ln(2) + \ln(k)$$

$$-\sum x_i^2 \le -20\ln(2) + 4\ln(k)$$

$$\sum x_i^2 \ge 20\ln(2) - 4\ln(k)$$

$$\therefore \sum x_i^2 \ge C$$

$$\therefore C = \left\{\underline{x} : \sum x_i^2 \ge C\right\} \; \textit{is best C.R.}$$

Likelihood Ratio Test

In this test we test the hypotheses of type: composite with composite and composite with simple. To illustrate this we take the example:

Ex: Let $X_1, X_2, \dots, X_n$ be a rssn from $N(\theta_1, \theta_2)$, where $(-\infty < \theta_1 < \infty \;,\; \theta_2 > 0)$
Let; $H_o: \theta_1 = 0 \;,\; \theta_2 > 0$, $H_1: \theta_1 \ne 0 \;,\; \theta_2 > 0$.

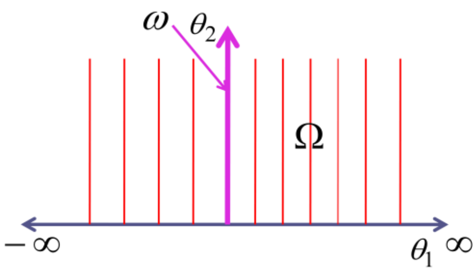
Sol:

$$\Omega = parameters\ space = \{(\theta_1,\theta_2): -\infty < \theta_1 < \infty\ ,\ \theta_2 > 0\}$$

$$\omega = \{(0,\theta_2)\ ;\ \theta_2 > 0\}$$

$$\Omega = H_o \cup H_1$$

$$H_o \sim \omega \quad , \quad H_1 \sim \Omega - \omega$$



$$L(x_1,x_2,\ldots,x_n \mid \underset{\substack{\theta_1=0 \\ \theta_2>0}}{} H_0) = \prod_{i=1}^{n} N(0,\theta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}}\, e^{-\frac{x_i^2}{2\theta_2}}$$

$$L(\omega) = \left(\frac{1}{2\pi\theta_2}\right)^{n/2} e^{-\frac{\sum x_i^2}{2\theta_2}}$$

$$\omega \subset \Omega$$

$$L(\Omega) = \prod_{i=1}^{n} N(\theta_1,\theta_2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}}\, e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}} = \left(\frac{1}{2\pi\theta_2}\right)^{n/2} e^{-\frac{\sum (x_i-\theta_1)^2}{2\theta_2}}$$

$$\lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = Likelihood\ Ratio\ Test = L.R.test \qquad , \ \lambda\ or\ \lambda(\underline{x})$$

$$L(\hat{\omega}) = L(x_1,x_2,\ldots,x_n \mid \theta_1 = 0\ ,\ \theta_2 = \hat{\theta}_2) = m.l.e.\ for\ \underline{\theta}\ under\ H_o\,(restricted)\,m.l.e$$

$$L(\hat{\Omega}) = L(x_1,x_2,\ldots,x_n \mid \theta_1 = 0\ ,\ \theta_2 = \hat{\theta}_2) = m.l.e.\ for\ \underline{\theta}\ under\ \Omega\,(unrestricted)\,m.l.e$$

$$\hat{\theta} = m.l.e.\ for\ \ \theta_2$$

$$\ln L(\omega) = -\frac{n}{2}\ln(2\pi\theta_2) - \frac{\sum x_i^2}{2\theta_2}$$

$$\frac{\partial \ln L(\omega)}{\partial \theta_2} = -\frac{n}{2} \times \frac{1}{2\pi\theta_2}(2\pi) + \frac{\sum x_i^2}{2\theta_2^2} \qquad \} = 0$$

$$-\frac{n}{2} \times \frac{1}{2\pi\hat{\theta}_2}(2\pi) + \frac{\sum x_i^2}{2\hat{\theta}_2^2} = 0$$

$$\Rightarrow \frac{\sum x_i^2}{2\hat{\theta}_2^2} = \frac{n}{2\hat{\theta}_2} \qquad \Rightarrow \frac{\sum x_i^2}{2\hat{\theta}_2} = \frac{n}{2}$$

$$2\sum x_i^2 = 2n\hat{\theta}_2 \qquad \Rightarrow \hat{\theta}_2 = \frac{\sum x_i^2}{n} \quad is\ m.l.e.\ \ for\ \ \theta_2\ in\ \omega$$

$$L(\hat{\omega}) = \left(\frac{n}{2\pi\sum x_i^2}\right)^{n/2} e^{-\frac{n}{2}}$$

$$L(\Omega) = \left(\frac{1}{2\pi\theta_2}\right)^{n/2} e^{-\frac{\sum (x_i-\theta_1)^2}{2\theta_2}}$$

$$\ln L(\Omega) = -\frac{n}{2} \times \ln\,(2\pi\theta_2) - \frac{\sum (x_i-\theta_1)^2}{2\theta_2}$$

$$\frac{\partial \ln L(\Omega)}{\partial \theta_1} = -\frac{2\sum(x_i-\theta_1)(-1)}{2\theta_2} = \frac{\sum(x_i-\theta_1)}{\theta_2} \qquad \text{....... (1)}$$

$$\frac{\partial \ln L(\Omega)}{\partial \theta_2} = -\frac{n}{2}\times\frac{1}{2\pi\theta_2}(2\pi) + \frac{\sum(x_i-\theta_1)^2}{2\theta_2^2} \qquad \text{....... (2)}$$

$$\frac{\sum(x_i-\hat{\theta}_1)}{\hat{\theta}_2} = 0 \qquad \text{....... (1)}$$

$$-\frac{n}{2}\times\frac{1}{\hat{\theta}_2} + \frac{\sum(x_i-\hat{\theta}_1)^2}{2\hat{\theta}_2^2} = 0 \qquad \text{....... (2)}$$

$$\therefore\ \hat{\theta}_1 = \frac{\sum x_i}{n} = \bar{x} \qquad , \quad \hat{\theta}_2 = \frac{\sum(x_i-\bar{x})^2}{n}$$

$$L(\hat{\Omega}) = \left(\frac{1}{2\pi\frac{\sum(x_i-\bar{x})^2}{n}}\right)^{n/2} e^{-\frac{\sum(x_i-\bar{x})^2}{\frac{2\sum(x_i-\bar{x})^2}{n}}}$$

$$= \left(\frac{n}{2\pi\sum(x_i-\bar{x})^2}\right)^{n/2} e^{-\frac{n}{2}}$$

$$\therefore\ \lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \frac{\left(\frac{n}{2\pi\sum x_i^2}\right)^{n/2} e^{-\frac{n}{2}}}{\left(\frac{n}{2\pi\sum(x_i-\bar{x})^2}\right)^{n/2} e^{-\frac{n}{2}}} = \left(\frac{\sum(x_i-\bar{x})^2}{\sum x_i^2}\right)^{n/2}$$

$$= \left(\frac{\sum(x_i-\bar{x})^2}{\sum(x_i-\bar{x})^2 + n\,\bar{x}^2}\right)^{n/2} = \left(\frac{1}{1+\frac{n\,\bar{x}^2}{\sum(x_i-\bar{x})^2}}\right)^{n/2} \le \lambda_o$$

$\omega = \{\theta \in H_o\}$, $\Omega = space$

$Let;\quad C = \{\underline{x} : \lambda(\underline{x}) < \lambda_o\}$

$$Type\ one\ Error = \alpha = pr\{\underline{x} \in c\} = pr\{\lambda(\underline{x}) < \lambda_o\} \qquad , \ \alpha : known$$

$$= pr\left\{\left(\frac{1}{1+\frac{n\bar{x}^2}{\sum(x_i-\bar{x})^2}}\right)^{n/2} < \lambda_o\right\} = pr\left\{\frac{1}{1+\frac{n\bar{x}^2}{\sum(x_i-\bar{x})^2}} < \lambda_o^{2/n}\right\}$$

$$= pr\left\{1+\frac{n\bar{x}^2}{\sum(x_i-\bar{x})^2} > \lambda_o^{-2/n}\right\}$$

$$= pr\left\{\frac{n\bar{x}^2}{\sum(x_i-\bar{x})^2} > \lambda_o^{-2/n} - 1\right\} \qquad , Z \sim N(0,1)$$

$$t-dist \rightarrow \quad T = \frac{Z}{\sqrt{\chi_r^2}} = \frac{Z}{\sqrt{\frac{W}{r}}}$$

$$W \sim \chi_r^2 \quad , \therefore Z, W \ \ are\ indep.$$

$$\therefore \bar{x}\ and\ \sum(x_i-\bar{x})\ \ are\ \ indep.$$

$$\frac{\sum(x_i-\bar{x})^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$\frac{nS^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$S^2 = \frac{\sum(x_i-\bar{x})^2}{n} \qquad , \ X \sim N(0,\theta_2) \quad , \ \bar{X} \sim N(0,\frac{\theta_2}{n})$$

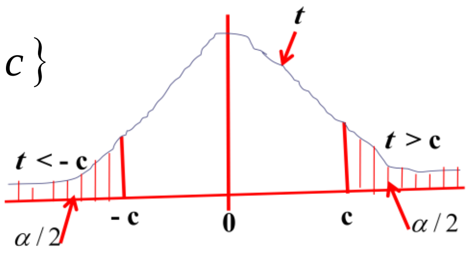
$$Z = \frac{\bar{x}-0}{\sqrt{\frac{\theta_2}{n}}} \sim N(0,1)$$

$$= \frac{\sqrt{n}\,\bar{x}}{\sqrt{\theta_2}}$$

$$\therefore pr\left\{\frac{n\bar{x}^2}{\sum(x_i-\bar{x})^2} > \lambda_o^{-2/n} - 1\right\} = pr\left\{\frac{n\bar{x}^2}{\frac{\sum(x_i-\bar{x})^2}{n-1}} > (n-1)(\lambda_o^{-2/n} - 1)\right\}$$

$$= pr\left\{\frac{n\bar{x}^2}{S^2} > (n-1)(\lambda_o^{-2/n} - 1)\right\} = pr\left\{\frac{\sqrt{n}\,|\bar{x}|}{\sqrt{S^2}} > \sqrt{(n-1)(\lambda_o^{-2/n} - 1)}\right\}$$

$$= pr\left\{ \frac{\sqrt{n}\,|\bar{x}|}{\sqrt{S^2}} > c \right\} = pr\left\{ \frac{|\bar{x}|}{\sqrt{\frac{S^2}{n}}} > c \right\} = pr\{|t| > c\}$$



$$|t| > c \Rightarrow t > c \quad or \quad -t > c \quad , \quad t < -c$$

H.W.: P(306): (10.1) and (10.6)

Ex9.17:p256: Let $Y_1 \le Y_2 \le Y_3 \le Y_4$ is order statistic, and $y_4 \le 1/2 \quad or \quad y_{4 \ge 1}$

$f(x;\theta) = \frac{1}{\theta}$, and let; $H_o : \theta = 1 \quad , \quad H_1 : \theta \ne 1$. Find $K(\theta)$?

Sol:

$$Let; \quad \theta = \hat{\theta} \ne 1$$

$$k(\theta) = pr\{reject\ H_o \mid H_1\ is\ true\}$$

$$1 - \beta = pr\{x \in c \mid \theta \ne 1\}$$

$$g(y_4;\theta) = 4\, f(y_4)\left(F(y_4)\right)^3 \quad , \; F(y_4) = pr(X \le y_4) = \int_0^{y_4} \frac{1}{\theta}\, dx = \frac{y_4}{\theta}$$

$$= 4\,\frac{1}{\theta}\left(\frac{y_4}{\theta}\right)^3 = \frac{4y_4^3}{\theta^4} \quad , \; 0 < y_4 < \theta$$

$$\therefore 1 - \beta = pr\left\{ y_4 \le \frac{1}{2} \,|\theta \ne 1 \right\} + pr\{y_4 \ge 1 |\theta = 1\}$$

$$= \int_0^{1/2} \frac{4y_4^3}{\theta^4}\, dy_4 + \int_1^{\theta} \frac{4y_4^3}{\theta^4}\, dy_4 \;\; | \; \theta \ne 0 \qquad , \; 0 < x < \theta$$

$$To\ find\ g(y_4;\theta)\ we\ put\ the\ first\ time\ \theta = \frac{1}{2}, and\ the\ second\ time\ \theta = 2;$$

$$\therefore 1 - \beta = \int_0^{1/2} \frac{4y_4^3}{\theta^4}\, dy_4 + \left(1 - \int_0^1 \frac{4y_4^3}{\theta^4}\, dy_4\right) = \left.\frac{y_4^4}{\theta^4}\right|_0^{1/2} + 1 - \left.\frac{y_4^4}{\theta^4}\right|_0^1$$

$$= \frac{1/16}{\theta^4} + 1 - \frac{1}{\theta^4}$$

$$when\ \ \theta = 1/2 \quad \Rightarrow \therefore 1 - \beta = \frac{1/16}{1/16} + 1 - \frac{1}{1/16} = 2 - 16 = ignor$$

$$when\ \ \theta = 2 \quad \Rightarrow \therefore 1 - \beta = \frac{1/16}{16} + 1 - \frac{1}{16} = \frac{1}{16 \times 16} + \frac{15}{16} = \frac{241}{256}$$

Ex9.19: Let $N_1(\mu_1, 400)$, $N_2(\mu_2, 225)$, and let $\theta = \mu_1 - \mu_2$,

$$H_o : \theta = 0 \qquad \bar{x} - \bar{y} \geq c$$

$$H_1 : \theta > 0 \qquad Find \quad (n) \quad and \quad c?$$

$\alpha = K(0) = 0.05$, $K(10) = 0.90$

Sol:

$$Let; \quad z = x - y \qquad \Rightarrow \bar{z} = \bar{x} - \bar{y}$$

$$Z \sim N(\mu_1 - \mu_2, 625) \Rightarrow \sim N(\theta, 625)$$

$$\bar{Z} \sim N(\theta, \frac{625}{n})$$

$$f(z;\theta) = \frac{1}{25\sqrt{2\pi}} e^{-\frac{(z-\theta)^2}{2(625)}} \qquad , -\infty < z < \infty$$

$$Let \quad \theta = 10 > 0$$

$$0.05 = pr\{\bar{Z} \geq c \mid \theta = 0\} \quad ..... (1)$$

$$\underline{0.90 = pr\{\bar{Z} \geq c \mid \theta = 10\} \quad ..... (2)}$$

$$0.05 = pr\left\{\frac{\bar{Z} - \theta}{\sigma / n} \geq \frac{c - \theta}{\sigma / n} \mid \theta = 0\right\}$$

$$0.05 = 1 - N\left(\frac{c - 0}{25 / n}\right) \quad ..... (1)$$

$$0.90 = pr\left\{\frac{\bar{Z} - \theta}{\sigma / n} \geq \frac{c - \theta}{\sigma / n} \mid \theta = 10\right\}$$

$$0.90 = 1 - N\left(\frac{c - 10}{25 / n}\right) \quad ..... (2)$$

$$H.W(from\ table)$$

$$c = c$$