Estimation Theory

Department of Statistics & Informative

Fourth Stage

Second Semester (2023-2024)

Le. Zainab A. M.

Contents of the Second Semester

** Same References as First Semester

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

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(θ) 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:

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

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

Note: If the second derivative less than zero that were 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) $\dfrac{\partial \ln\left(L(\underline{x}; \theta)\right)}{\partial \theta} = 0.$

4) $Find \;\; \hat{\theta}.$

Ex1: Let $X_1, X_2, \ldots, X_n$ denote a random sample from Bernoulli dist^n Ber(θ), find the m.l.e for θ.

Sol:

$\because$ X ~ Ber(θ)

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

$\because X's \;\; are \; indep.$

$$L(\theta) = f(x_1, x_1, \ldots, x_1;\; \theta) = \Pi\, f(x_i; \theta)$$

$$= \theta^{\Sigma x_i} (1-\theta)^{n - \Sigma x_i}$$

$$\ln L(\theta) = \Sigma x_i \ln(\theta) + (n - \Sigma x_i) \ln(1-\theta)$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{\Sigma x_i}{\theta} - \frac{n - \Sigma x_i}{1-\theta} \qquad , \qquad \frac{\partial \ln L(\theta)}{\partial \theta} = 0$$

$$\frac{\Sigma x_i}{\theta} - \frac{n - \Sigma x_i}{1-\theta} = 0$$

$$\frac{(1-\theta)\Sigma x_i - \theta(n - \Sigma x_i)}{\theta(1-\theta)} = 0$$

$\Sigma x_i - \hat{\theta}\,\Sigma x_i - n\hat{\theta} + \hat{\theta}\Sigma x_{i_i} = 0$

$\Sigma x_i - n\hat{\theta} = 0$

$\Sigma x_i = n\hat{\theta} \qquad \hat{\theta}_{m.l.e} = \frac{\Sigma X_i}{n} = \overline{X}$

$$\frac{\partial^2 \ln L(\theta)}{\partial\theta^2} = -\frac{\Sigma x_i}{\theta^2} - \frac{n - \Sigma x_i}{(1-\theta)^2} < 0$$

$\therefore \hat{\theta} = \overline{X}$ is m.l.e for θ .

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.

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 $u(\theta)$.

Ex1: In a rss$n$ from exponential dist^n Exp($1/\theta$), find the m.l.e for:

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

Sol:

$X \sim Exp(1/\theta)$

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

$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^n e^{-\theta \sum x_i}$

$\ln L(\theta) = n \ln(\theta) - \theta \sum x_i$

$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{n}{\theta} - \sum x_i \qquad , \qquad \frac{\partial \ln L(\theta)}{\partial \theta} = 0$

$\frac{n}{\theta} - \sum x_i = 0 \quad \Rightarrow \frac{n}{\theta} = \sum x_i \quad \Rightarrow \hat{\theta} = \frac{n}{\sum X_i} = \frac{1}{\overline{X}}$

$\frac{\partial^2 \ln L(\theta)}{\partial \theta^2} = \frac{-n}{\theta^2} < 0 \qquad \Rightarrow \therefore \hat{\theta}$ *is m.l.e for* θ .

1) $u_1(\hat{\theta}) = \frac{1}{\hat{\theta}} = \frac{1}{\frac{1}{\overline{x}}} = \overline{x}$ 2) $u_2(\hat{\theta}) = \frac{\ln(\hat{\theta})}{\hat{\theta}} = \frac{\ln\left(\frac{1}{\overline{x}}\right)}{\frac{1}{\overline{x}}} = \overline{x}\ln\left(\frac{1}{\overline{x}}\right)$

Ex: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n N(θ, σ^2), find**:**
1) 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.
Sol: 1)

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

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

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

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

1) *For* θ ?

$$\ln L(\theta,\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum(x_i-\theta)^2$$

$$\frac{\partial \ln L(\theta,\sigma^2)}{\partial \theta} = zero - zero - \frac{2}{2\sigma^2}\sum(x_i-\theta)\,(-1) = \frac{\sum(x_i-\theta)}{\sigma^2} = \frac{\sum x_i - n\theta}{\sigma^2}$$

$$\frac{\partial \ln L(\theta,\sigma^2)}{\partial \theta} = 0$$

$$\frac{\sum x_i - n\theta}{\sigma^2} = 0 \quad \Rightarrow \sum x_i - n\theta = 0 \quad \Rightarrow \sum x_i = n\theta \quad \Rightarrow \hat{\theta}_{m.l.e} = \frac{\sum x_i}{n} = \overline{x}$$

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial \theta^2} = \frac{-n}{\sigma^2} < 0$$

$\therefore\ \hat{\theta} = \overline{x}$ *is m.l.e for* θ .

2) *For* σ^2 ?

$$\frac{\partial \ln L(\theta,\sigma^2)}{\partial \sigma^2} = zero - \frac{n}{2\sigma^2} + \frac{2\sum(x_i-\theta)^2}{4\sigma^4} = -\frac{n}{2\sigma^2} + \frac{\sum(x_i-\theta)^2}{2\sigma^4}$$

$$= \frac{-n\sigma^2 + \sum(x_i-\theta)^2}{2\sigma^4} = 0$$

$$-n\sigma^2 + \sum(x_i-\theta)^2 = 0 \quad \Rightarrow \quad n\sigma^2 = \sum(x_i-\theta)^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum(x_i-\hat{\theta})^2}{n} = \frac{\sum(X_i-\overline{X})^2}{n} = S^2$$

$$\because \frac{\partial \ln L(\theta,\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum(x_i-\theta)^2}{2\sigma^4} \quad \} \times 2\sigma^2$$

$$= -n + \frac{\sum(x_i-\theta)^2}{\sigma^2} \quad => \quad \frac{-n\,\sigma 2 + (\Sigma x_i - \theta)2}{\sigma 2} = 0 => n\,\sigma 2 = (\Sigma x_i - \theta)2$$

$$=> \sigma 2 = \frac{(\Sigma x_i - \overline{X})2}{n} = S2$$

$$\frac{\partial^2 \ln L(\theta,\sigma^2)}{\partial (\sigma^2)^2} = zero - \frac{\sum(x_i-\theta)^2}{\sigma^4} = -\frac{\sum(x_i-\theta)^2}{\sigma^4} < 0$$

$\therefore\ S^2$ *is m.l.e for* σ^2

2)

$u(\sigma^2) = u(\hat{\sigma}^2)$

$u(\sigma^2) = \sqrt{\sigma^2} = \sigma$

$$u(\hat{\sigma}^2 = S^2) = \sqrt{S^2} = \sqrt{\frac{\sum(X_i-\overline{X})^2}{n}} = S$$

$\therefore\ S$ *is 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:

$$m_k = M_k$$

Ex1: Let $X_1, X_2, \ldots, X_n$ be a rss$n$ from normal dist^n $N(\theta, \sigma^2)$, estimate the parameters θ and σ^2 using moment method.

Sol:

$$m_k = M_k$$

$$m_k = \frac{\sum X_i^k}{n} \quad , \quad M_k = E(X^k)$$

$$m_1 = \frac{\sum X_i}{n} \quad \Rightarrow \quad M_1 = E(X) = \theta$$

$$m_1 = M_1$$

$$\frac{\sum X_i}{n} = \theta \quad \Rightarrow \quad \therefore \hat{\theta} = \overline{X}$$

$$m_2 = \frac{\sum X_i^2}{n} \quad \Rightarrow \quad M_2 = E(X^2)$$

$$M_2 = E(X^2) = V(X) + (E(X))^2 = \sigma^2 + \theta^2$$

$$m_2 = M_2$$

$$\frac{\sum X_i^2}{n} = \sigma^2 + \overline{X}^2$$

$$\therefore \hat{\sigma}^2 = \frac{\sum X_i^2}{n} - \overline{X}^2 = \frac{(\Sigma x_i - \overline{X})2}{n} = S2$$

Ex2: 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.

Sol:

$$m_k = M_k$$

$$m_k = \frac{\sum X_i^k}{n} \quad , \quad M_k = E(X^k)$$

$$m_1 = \frac{\sum X_i}{n}$$

$$M_1 = E(X)$$

$$E(X) = \int_0^1 x\, f(x;\theta)\, dx = \int_0^1 x\, (\theta+1)\, x^{\theta}\, dx = \int_0^1 (\theta+1)\, x^{\theta+1}\, dx = (\theta+1) \left. \frac{x^{\theta+2}}{\theta+2} \right|_0^1 = \frac{\theta+1}{\theta+2}$$

$$m_1 = M_1$$

$$\frac{\sum X_i}{n} = \frac{\theta+1}{\theta+2} \quad \Rightarrow \quad \overline{X} = \frac{\theta+1}{\theta+2} \quad \Rightarrow \quad (\theta+2)\,\overline{X} = \theta+1 \quad \Rightarrow \quad \theta\,\overline{X} + 2\,\overline{X} = \theta+1$$

$$\Rightarrow \theta\,\overline{X} - \theta = 1 - 2\,\overline{X} \quad \Rightarrow \theta(\overline{X} - 1) = 1 - 2\,\overline{X}$$

$$\Rightarrow \therefore \hat{\theta} = \frac{1 - 2\,\overline{X}}{\overline{X} - 1}$$

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 $\hat{\theta}$.

Ex1: In a rssn, find m.v.e. for the parameters of; **1)** Ber(θ). **2)** N(θ, σ^2).

Sol:

1) $X \sim Ber(\theta)$

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

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

$$\ln L(\theta) = \sum x_i \ln(\theta) + (n-\sum x_i)\ln(1-\theta)$$

$$\frac{\partial \ln L(\theta)}{\partial\theta} = \frac{\sum x_i}{\theta} - \frac{n-\sum x_i}{(1-\theta)} = \frac{(1-\theta)\sum x_i - \theta(n-\sum x_i)}{\theta(1-\theta)} = \frac{\sum x_i - \theta\sum x_i - n\theta + \theta\sum x_i}{\theta(1-\theta)}$$

$$= \frac{\sum x_i - n\theta}{\theta(1-\theta)} \quad (\div n)$$

$$= \frac{\bar{x}-\theta}{\frac{\theta(1-\theta)}{n}} = \frac{\hat{\theta}-\theta}{V(\hat{\theta})} \quad , \quad \hat{\theta} = \bar{X} \quad , \quad V(\hat{\theta}) = V(\bar{X}) = \frac{V(X)}{n} = \frac{\theta(1-\theta)}{n}$$

$\therefore\ \bar{X}$ *is m.v.e. for* θ.

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;

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

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

$$p(\theta\,|\,x) = \frac{p(x\,|\,\theta)}{p(x} = ?$$

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

$$p(\theta \mid x) \propto p(x \mid \theta)\ p(\theta)$$
$$\propto L(\theta)\ p(\theta)$$

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;

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

Ex1: Find Bayes estimator for parameter of; **1)** Ber(θ). **2)** Poisson(θ), using non informative prior probability.

Sol:

1) $X \sim Ber(\theta)$

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

$$p(\theta \mid x) \propto L(\theta)\ p(\theta)$$

$$L(\theta) = \theta^{\sum x_i}\ (1-\theta)^{n-\sum x_i}$$

$$p(\theta) \propto (I_s(\theta))^{1/2} \qquad , \ F.I = -E\left(\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right)$$

$$\ln f(x;\theta) = x \ln(\theta) + (1-x)\ln(1-\theta)$$

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

$$\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2} = \frac{-x}{\theta^2} - \frac{1-x}{(1-\theta)^2}$$

$$F.I. - E\left(\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right) = \frac{E(X)}{\theta^2} + \frac{E(1-X)}{(1-\theta)^2}$$
$$= \frac{E(X)}{\theta^2} + \frac{1-E(X)}{(1-\theta)^2} = \frac{1}{\theta} + \frac{1}{(1-\theta)} = \frac{1}{\theta(1-\theta)}$$

$$p(\theta) \propto \left(\frac{1}{\theta(1-\theta)}\right)^{1/2}$$

$$\propto \theta^{-1/2} (1-\theta)^{-1/2}$$

$$p(\theta \mid x) \propto L(\theta)\; p(\theta)$$

$$\propto \theta^{\sum x_i} e^{n-\sum x_i} \theta^{-1/2} (1-\theta)^{-1/2}$$

$$\propto \theta^{\sum x_i - \frac{1}{2}} (1-\theta)^{n-\sum x_i - \frac{1}{2}}$$

$$\alpha - 1 = \sum x_i - \frac{1}{2} \qquad \Rightarrow \alpha = \sum x_i + \frac{1}{2}$$

$$\beta - 1 = n - \sum x_i - \frac{1}{2} \qquad \Rightarrow \beta = n - \sum x_i + \frac{1}{2}$$

$$p(\theta \mid x) \sim Beta(\alpha = \sum x_i + \frac{1}{2}, \beta = n - \sum x_i + \frac{1}{2})$$

When; $X \sim Beta(\alpha, \beta)$,

$$f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad , \quad E(X) = \frac{\alpha}{\alpha+\beta}$$

then the complete p.d.f. of the posterior probability is;

$$p(\theta \mid x) = \frac{\Gamma(n+1)}{\Gamma\left(\sum x_i + \frac{1}{2}\right)\Gamma\left(n - \sum x_i + \frac{1}{2}\right)} \theta^{\sum x_i - \frac{1}{2}} (1-\theta)^{n-\sum x_i - \frac{1}{2}}$$

$$\therefore E(\theta \mid X_1, \ldots, X_n) = \hat{\theta}_{Bayes} = \frac{\alpha}{\alpha+\beta}$$

$$= \frac{\sum X_i + \frac{1}{2}}{n+1} = \frac{\sum X_i}{n+1} + \frac{1}{2n+2}$$

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/θ)	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.

Sol:

1) $X \sim Geo\ (\theta)$

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

$$p(\theta \mid x_1, x_2, \ldots, x_n) \propto L(\theta)\ p(\theta)$$

$$L(\theta) = \theta^{n}(1-\theta)^{\sum x_i}$$

$$p(\theta) \sim Beta\ (\alpha_o\ ,\ \beta_o)$$

$$p(\theta) = \frac{\Gamma(\alpha_o + \beta_o)}{\Gamma(\alpha_o)\ \Gamma(\beta_o)}\ \theta^{\alpha_o - 1}(1-\theta)^{\beta_o - 1}$$

$$\propto\ \theta^{\alpha_o - 1}(1-\theta)^{\beta_o - 1}$$

$$p(\theta \mid x_1, x_2, \ldots, x_n) \propto \theta^{n}(1-\theta)^{\sum x_i}\ \theta^{\alpha_o - 1}(1-\theta)^{\beta_o - 1}$$

$$\propto \theta^{(\alpha_o + n) - 1}(1-\theta)^{(\sum x_i + \beta_o) - 1}$$

$$p(\theta \mid x_1, x_2, \ldots, x_n) \sim Beta(\alpha = \alpha_o + n\ ,\ \beta = \sum x_i + \beta_o)$$

The complete p.d.f. of the posterior probability is;

$$\therefore\ p(\theta \mid x_1, x_2, \ldots, x_n) = \frac{\Gamma(\alpha_o + n + \sum x_i + \beta_o)}{\Gamma(\alpha_o + n)\ \Gamma(\sum x_i + \beta_o)}\ \theta^{(\alpha_o + n) - 1}(1-\theta)^{(\sum x_i + \beta_o) - 1}$$

$$\therefore\ E(\theta \mid X) = \hat{\theta}_{Bayes} = \frac{\alpha}{\alpha + \beta} = \frac{\alpha_o + n}{\alpha_o + n + \sum X_i + \beta_o}$$

Chapter Two

Interval Estimation

Definition:

In a rssn taken 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 – α)% confidence coefficient, where; L_1: is lower confidence limit. L_2: is upper confidence limit.

Confidence Interval (CI)

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 $\overline{X}$ is distributed with mean θ and the variance $\frac{\sigma^2}{n}$ and $Z = \frac{\overline{X} - \theta}{\sigma / \sqrt{n}}$ has standard normal dist^n or: $\overline{X} \sim N(\theta, \frac{\sigma^2}{n})$

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

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

or;

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

Note:

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

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

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

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

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 $n \ge 30$, $Z = \frac{\overline{X} - \theta}{S / \sqrt{n}}$ then:

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

or;

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

b) If a sample size $n < 30$, $T = \dfrac{\overline{X} - \theta}{S/\sqrt{n}}$ 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: 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:

$1 - \alpha = 95\%$

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

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

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

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

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

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 rss$n$ from a normal population with mean μ_Y and unknown variance σ_Y^2, then:

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

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

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

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

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

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

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

Sol:

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

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

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

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

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

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

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;

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

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

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

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

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