BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024

Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Statistical Inference-II

Course Code: 23ST11CC

Unit-IV

Maximum Likelihood Estimator

Dr. T. Jai Sankar

Associate Professor and Head

Department of Statistics

Ms. I. Angel Agnes Mary

Guest Faculty

Department of Statistics

Likelihood Ratio (LR) Test

The likelihood ratio (LR) test is a test of hypothesis in which two different maximum likelihood estimates of a parameter are compared in order to decide whether to reject or not to reject a restriction on the parameter.

Likelihood ratio test is useful for testing simple or composite hypothesis. If $f(x, \theta)$ is the density function of a population and $L(\theta)$ is a likelihood function of sample observations $x_1, x_2, x_3, \ldots, x_n$ then the likelihood ratio λ is defined as

$$\lambda = \frac{\text{Maximum of Likelihood function } L(\theta) | H_0}{\text{Maximum of } L(\theta)}$$

If the parameter θ is replaced by its maximum likelihood estimator $\hat{\theta}$, then we get $L(\hat{\theta})$. ie., $H_0 : \theta = \theta_0$, then we get $L(\hat{\theta})$. (ie) Max $L(\theta) = L(\hat{\theta})$

$$\lambda = \frac{L(\theta_0)}{L(\hat{\theta})}$$

Any test for testing H_0 against H_1 is called likelihood ratio test. If it is based on likelihood ratio λ and the critical region $0 \le \lambda \le \lambda_0$ such that $\int_0^{\lambda} g(\lambda | H_0) d\lambda = \alpha$

Likelihood Ratio Test for the Mean of a Normal Distribution (μ and σ^2 are Unknown)

Statement

Let $x_1, x_2, \ldots, x_n$ form a random sample of size n from the normal distribution with mean μ and variance σ^2, where μ and σ^2 are unknown. Consider the problem of testing the null hypothesis $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$.

Proof

Let $x_1, x_2, \ldots, x_n$ be a random sample from $N(\mu, \sigma^2)$ where $x, \mu \in \Re, \sigma > 0$.

The joint pdf of $x_1, x_2, \ldots, x_n$ is

$$\prod_{i=1}^{n} f(x_i : \mu, \sigma^2) = \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2} \right], x_i, \mu \in \Re, \sigma > 0$$

The likelihood function of the sample observation $x_1, x_2, \ldots, x_n$ is given by,

$$L(\mu, \sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n e^{-\frac{1}{2}\Sigma\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

$$L(\mu_0, \sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum\left(\frac{x_i - \mu_0}{\sigma}\right)^2}$$

MLE of μ and σ^2 are $\hat{\mu} = \bar{x}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2 = s^2$

The maximum of likelihood function is given by

$$L(\hat{\mu}, \hat{\sigma}^2) = \left(\frac{1}{s\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum\left(\frac{x_i - \bar{x}}{s}\right)^2} = (2\pi s^2)^{-n/2} e^{-n/2} \text{----(1)}$$

Maximum Likelihood estimator of for σ^2 when H_0: $\mu = \mu_0$ is true given by

$$\sigma^2 = \frac{1}{n}\sum(x_i - \mu_0)^2 = \frac{1}{n}\sum(x_i - \bar{x} + \bar{x} - \mu_0)^2 = s^2 + (\bar{x} - \mu_o)^2$$

Therefore, $\hat{\sigma}^2 = s_0^2$

Maximum likelihood function under H_0 is

$$L(\hat{\sigma}^2 | H_0) = \left(\frac{1}{s_0\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum\left(\frac{x_i - \bar{x}}{s_o}\right)^2} = (2\pi s_o^2)^{-n/2} e^{-n/2} \text{----(2)}$$

The UMP critical region of size α is given by $0 \le \lambda \le \lambda_0$

$$\Rightarrow \frac{\text{Maximum Likelihood function} | H_o}{\text{Maximum Likelihood function}} \le \lambda_0$$

Using (1) and (2),

$$\Rightarrow \frac{(2\pi s_o^2)^{-n/2} e^{-n/2}}{(2\pi s^2)^{-n/2} e^{-n/2}} \le \lambda_0$$

$$\Rightarrow \left(\frac{s_0^2}{s^2}\right)^{-n/2} \le \lambda_0 \Rightarrow \left(\frac{s^2 + (\bar{x} - \mu_o)^2}{s^2}\right)^{-n/2} \le \lambda_0 \Rightarrow \left(\frac{1}{1 + \frac{(\bar{x} - \mu_o)^2}{s^2}}\right)^{n/2} \le \lambda_0$$

Where λ_0 is fixed such that size of CR is

$$t = \frac{(\bar{x} - \mu_0)}{s/\sqrt{n-1}} \sim t_{n-1}$$

$$\lambda = \left[\frac{1}{1 + \frac{t^2}{n-1}}\right]^{n/2} \le \lambda_0$$

Therefore, t-distribution can be used to find the value for given α and degrees of freedom (n-1).

Therefore, UMPT of size α for testing mean of the normal distribution when σ^2 is unknown is based on t-distribution.

The UMP CR of Size α is given by

$$\left[\frac{1}{1+\frac{t^2}{n-1}}\right]^{n/2} \le \lambda_0 \Rightarrow \left[\frac{1}{1+\frac{t^2}{n-1}}\right] \le \lambda_0^{2/n}$$

$$\frac{1}{(\lambda_0)^{2/n}} \le 1 + \frac{t^2}{(n-1)}$$

$$\frac{1}{(\lambda_0)^{2/n}} - 1 \le \frac{t^2}{(n-1)}$$

$$t^2 \ge (n-1)\left[\frac{1}{(\lambda_0)^{2/n}} - 1\right]$$

Hence proved.

Properties of Likelihood Ratio Test

- Likelihood ratio test leads to uniformly most powerful test if it exists.
- When the sample size n is large $-2\log_e \lambda \sim \chi^2$ distribution with respective degrees of freedom
- Under certain conditions likelihood ratio tests are consistent.
- If the distribution *f(x, θ)* has a monotone likelihood ratio in D(x) then there exists UMP test for testing $H_0: \theta \le \theta_0$ or $H_0: \theta \ge \theta_0$ against $H_1: \theta < \theta_0$.

Asymptotic Distribution of LR Test Statistic

Statement:

For testing $H_0: \theta = \theta_0$ versus $H_1: \theta \neq \theta_0$ suppose $X_1, \dots, X_n$ are i.i.d. $f(x|\theta)$, $\hat{\theta}$ is the MLE of θ and satisfies the regularity conditions. Then under H_o as $n \to \infty$, $-2\log\lambda(X) \to \chi_1^2$ in distribution, where χ_1^2 is a χ^2 random variable with 1 degree of freedom

Proof:

From the test statistic of likelihood ratio test,

$$\lambda(X) = \frac{L(\widetilde{\theta_0}|x)}{L(\hat{\theta}|x)}$$

Taking log on both sides,

$$log(\lambda(X)) = log(\frac{L(\widetilde{\theta_0}|x)}{L(\hat{\theta}|x)})$$

$$-2log\ \lambda(X) = \ -2(log(L(\widetilde{\theta_0}|x) - log(L(\hat{\theta}|x))$$

$$= -2log(L(\widetilde{\theta_0}|x)) + 2log(L(\hat{\theta}|x))$$

First expand $log\ L(\theta|x) = \ l(\theta|x)$ in a Taylor series around $\hat{\theta}$, giving

$$l(\theta|x)\ =\ l(\hat{\theta}\ |x) + l'(\hat{\theta}\ |x)(\theta - \hat{\theta}\ ) + l''(\hat{\theta}\ |x)\frac{(\theta - \hat{\theta}\ )^2}{2} + \ldots$$

Similarly expand $log\ L(\theta_0\ |x) = \ l(\theta_0\ |x)$

Substitute the expansion for $l(\hat{\theta}\ |x)$ and $l(\widetilde{\theta_0}\ |x)$ into (1)

And get

$$-2\ log\lambda(X)\ \approx \frac{(\theta - \hat{\theta}\ )^2}{l''(\hat{\theta}\ |x)}$$

Where use the fact that $l'(\hat{\theta}\ |x)$=0. Since the denominator is the observed information $\hat{X}_i = \frac{1}{n}\sum_{i=1}^{n} X_i$ and it follows from theorem (consistency of $\hat{x}$) form theorem and Slutsky's theorem that -2 logλ(X) $\rightarrow \chi_1^2$

Consistency of LR Test

The likelihood ratio (LR) test is a statistical test used in hypothesis testing, especially in the context of comparing nested models in regression analysis. The test assesses the goodness of fit of two models: a null model and an alternative model.

The consistency of the LR test refers to its ability to correctly reject the null hypothesis when the null model is false as the sample size grows. In other words, as you collect more data, the LR test should become more reliable in detecting a true difference between the null and alternative models.

The LR test is consistent, which means that as the sample size increases, the test's power (the probability of correctly rejecting the null hypothesis when it's false) approaches 1, and its type I error rate (the probability of incorrectly rejecting the null hypothesis when it's true) approaches 0. This property is important in statistical inference as it ensures that the test will correctly identify significant relationships or differences when they exist in the population, given a sufficiently large sample size.

Construction of LR Tests for Standard Statistical Distributions

LR Test for the Variance of a Normal Distribution (Mean and Variance are Unknown)

Let $X_1, X_2, \ldots, X_n$ form a random sample of size n from the normal distribution with mean μ and variance σ^2, where μ and σ^2 are unknown. Consider the problem of testing the null hypothesis

$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{vs.} \quad H_a : \sigma^2 \neq \sigma_0^2$$

Proof

Let $x_1, x_2, \ldots, x_n$ be a random sample from $N(\mu, \sigma^2)$ where $x, \mu \in \Re, \sigma > 0$.

The parameter space is

$$\Theta = \Theta_0 \cup \Theta_a = \{(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 > 0\}.$$

$$\Theta_a = \{(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 \neq \sigma_0^2\},$$

The joint pdf of $x_1, x_2, \ldots, x_n$ is

$$\prod_{i=1}^{n} f(x_i : \mu, \sigma^2) = \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2} \right], x_i, \mu \in \Re, \sigma > 0$$

The likelihood function of the sample observation $x_1, x_2, \ldots, x_n$ is given by,

$$L = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} . \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right\}$$

$$L(\hat{\Theta}) = \left(\frac{1}{2\pi s^2}\right)^{n/2} \exp\left(-\frac{n}{2}\right)$$

In θ_0, we have only one variable parameter, viz., μ and

$$L(\Theta_0) = \left(\frac{1}{2\pi\sigma_0^2}\right)^{n/2} \exp\left[-\frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i - \mu)^2\right]$$

The MLE for μ is given by

$$\frac{\partial}{\partial \mu} \log L = 0 \quad \Rightarrow \quad \hat{\mu} = \bar{x}$$

$$\therefore \qquad L(\hat{\Theta}_0) = \left(\frac{1}{2\pi\sigma_0^2}\right)^{n/2} \exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]$$

$$= \left(\frac{1}{2\pi\sigma_0^2}\right)^{n/2} \exp\left[-\frac{ns^2}{2\sigma_0^2}\right]$$

The likelihood ratio criterion is given by

$$\lambda = \frac{L(\hat{\Theta}_0)}{L(\hat{\Theta})} = \left[\frac{s^2}{\sigma_0^2}\right]^{n/2} \exp\left[-\frac{1}{2}\left(\frac{ns^2}{\sigma_0^2} - n\right)\right]$$

We know that under H_0 the statistic

$$\chi^2 = \frac{ns^2}{\sigma_0^2}$$

Follows chi-square distribution with (n -1) d.f. In terms of χ^2, we have

$$\lambda = \left[\frac{\chi^2}{n}\right]^{n/2} . \exp\left[-\frac{1}{2}(\chi^2 - n)\right]$$

Since λ is a monotonic function of χ^2, the test may be done using χ^2 as a criterion. The critical region $0 < \lambda < \lambda_0$ is now equivalent to

$$(\chi^2/n)^{n/2} \exp\left[-\frac{1}{2}(\chi^2 - n)\right] < \lambda_0$$

Since χ^2, has chi-square distribution with (n - 1) d.f.

Monotone Likelihood Ratio Property

The monotone likelihood ratio (MLR) property for a family of probability mass function or probability density function denoted by $\{p(x, \theta) : \theta \in \Theta \subset R\}$. we exploit this property to derive the UMP level α tests for one-sided null against one-sided alternative hypotheses in some situations.

Monotone Likelihood Ration (MLR) family of distribution

A real parametric family $\{p(x, \theta) : \theta \in \Theta \subset R\}$ is said to have MLR property in a

real valued statistic $T(x)$ if, for any $\theta_1 < \theta_2 \in \Theta$, the following are satisfied.

i) $p(x, \theta_1) \neq p(x, \theta_2)$

[Distributions are distinct corresponding to distinct parameter points]

ii) The ratio

$$R(x) = \frac{p(x, \theta_2)}{p(x, \theta_1)}$$

is non-decreasing in $T(x)$ on the set $\{x : max(p(x, \theta_2), p(x, \theta_1)) > 0\}$.

Note: If $p(x, \theta_2) = 0$ and $p(x, \theta_1) > 0$, $R(x) = 0$.

$p(x, \theta_2) > 0$ and $p(x, \theta_1) = 0$, $R(x) = \infty$.

Example on MLR families

Let $\{p(x, \theta), \theta \in \Theta \subset R\}$: One parameter Exponential family. Then we can express $p(x, \theta)$ in the form,

$$p(x, \theta) = u(\theta)\ exp(q(\theta)\ T(x))\ v(x)$$

such that $u(\theta)$ and $q(\theta)$ depends only on θ, $v(x)$ is independent of θ and $T(x)$ depends only on x. We set $T(x)$ such that $Q(\theta)$ is a strictly increasing function of θ. Then we have for $\theta_1 < \theta_2$,

$$\frac{p(x, \theta_2)}{p(x, \theta_1)} = \frac{u(\theta_2)}{u(\theta_1)} exp\{(Q(\theta_2) - Q(\theta_1))T(x)\},$$

increasing in $T(x)$ because $Q(\theta)$ is a strictly increasing function of θ. Hence, $\{p(x, \theta), \theta \in \Theta\}$ has MLR in $T(x)$.

Note: If $(X_1, X_2, \ldots, X_n)$ is a random sample of size n from the population with p.m.f or p.d.f. $p(x, \theta)$ then $p(x, \theta)$ has MLR in $\sum_{i=1}^n T(x_i)$.

1. Monotone likelihood ratio based on Bernoulli Population

Let $X = (X_1, X_2, \ldots, X_n)$, be a random sample of size n from $Bernoulli(\theta)$ population.

$$\begin{aligned} p(\mathbf{x}, \theta) &= \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} \\ &= (1-\theta)^n exp\left[ln(\frac{\theta}{1-\theta})\sum_{i=1}^n x_i\right] \\ &= u(\theta)exp(q(\theta)T(\mathbf{x})v(\mathbf{x}) \end{aligned}$$

Where,

- $c(\theta) = \theta^n$
- $q(\theta) = ln(\frac{\theta}{1-\theta})$
- $T(\mathbf{x}) = \sum_{i=1}^n x_i$
- $v(\mathbf{x}) = 1$

p(x, θ) has MLR in $T(\mathbf{x}) = \sum_{i=1}^{n} x_i$

2. Monotone likelihood ratio based on Normal Population

Let $X = (X_1, X_2, \ldots, X_n)$, be a random sample of size n from $N(0, \theta_2)$ population.

Therefore,

$$\begin{aligned} p(\mathbf{x}, \theta) &= (2\pi)^{-n/2}\theta^{-n} exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) \\ &= u(\theta)exp(q(\theta)T(\mathbf{x}))v(\mathbf{x}) \end{aligned}$$

Where,

- $u(\theta) = \theta^{-n}$
- $q(\theta) = -\frac{1}{2\theta^2}$
- $T(\mathbf{x}) = \sum_{i=1}^{n} x_i$
- $v(\mathbf{x}) = (2\pi)^{-n/2}$

p(x, θ) has MLR in $T(\mathbf{x}) = \sum_{i=1}^{n} x_i$

3. Monotone likelihood ratio based on Exponential Population

Let $X = (X_1, X_2, \ldots, X_n)$, be a random sample of size n from the exponential distribution with p.d.f.

$$p(x, \theta) = \theta\, exp[-\theta x],\ x > 0,\ \theta > 0$$

Now

$$\begin{aligned} p(\mathbf{x}, \theta) &= \theta^n exp\left[-\theta\sum_{i=1}^{n} x_i\right] \\ &= u(\theta)exp(q(\theta)T(\mathbf{x})v(\mathbf{x}) \end{aligned}$$

Where,

- $u(\theta) = \theta^n$
- $q(\theta) = \theta$
- $T(\mathbf{x}) = -\sum_{i=1}^{n} x_i$
- $v(\mathbf{x}) = 1$

p(x, θ) has MLR in $T(\mathbf{x}) = -\sum_{i=1}^{n} x_i.$

Uniformly Most Powerful Tests

A critical region w of size α for testing $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$ is said to be uniformly most powerful critical region if for every value of $\theta \neq \theta_0$ the power of the critical region w must be greater than or equal to the critical region w must be greater than or equal to power of any other critical region w^* of same size α any test based on uniformly most powerful critical region is called uniformly most powerful test.

Applications to Standard Statistical Distributions on Uniform Most Powerful Test

Given a random sample $x_1, x_2, \ldots, x_n$ from the distribution with the pdf

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

Show that there exist no UMPT for testing $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$.

Solution

Let $x_1, x_2, \ldots, x_n$ be a random sample from exponential distribution then the likelihood function is given by

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

$$\frac{L_0}{L_1} \leq k$$

$$\prod_{i=1}^{n} f(x_i, \theta_0) = \prod_{i=1}^{n} \theta_0 e^{-\theta_0}$$

$$\frac{L_0}{L_1} = \frac{\theta_0^{\,n} e^{-\theta_0 \sum_{i=1}^{n} x_i}}{\theta_1^{\,n} e^{-\theta_1 \sum_{i=1}^{n} x_i}} \leq k$$

$$\left(\frac{\theta_0}{\theta_1}\right)^n e^{-\sum_{i=1}^{n} x_i (\theta_0 - \theta_1)} \leq k$$

Taking log on both sides,

$$n \log\left(\frac{\theta_0}{\theta_1}\right) - \sum_{i=1}^{n} x_i (\theta_0 - \theta_1) \leq \log k$$

$$n[\log \theta_0 - \log \theta_1] - \sum_{i=1}^{n} x_i (\theta_0 - \theta_1) \leq \log k$$

$$-\sum_{i=1}^{n} x_i (\theta_0 - \theta_1) \leq \log k - n[\log \theta_0 - \log \theta_1]$$

Case I: $\theta_0 > \theta_1$

$$\Rightarrow \theta_0 - \theta_1 > 0$$

$\log \theta_0 - \log \theta_1$ is a positive quantity

$$\sum_{i=1}^{n} x_i \geq \frac{\log k - n[\log \theta_0 - \log \theta_1]}{\theta_0 - \theta_1}$$

The BCR is given by

$$\sum x_i \leq \frac{k_1}{\theta_0 - \theta_1} = \lambda_1 \,(say)$$

Case I: $\theta_0 < \theta_1$

$\log \theta_0 - \log \theta_1$ is a positive quantity

$$\sum_{i=1}^{n} x_i \leq \frac{\log k + n[\log \theta_0 - \log \theta_1]}{\theta_0 - \theta_1}$$

the BCR is given by

$$\sum x_i \geq \frac{k_1}{\theta_0 - \theta_1} = \lambda_2 \,(say)$$

The constant λ_1 and λ_2 are determined such that

$$p\left[\sum x_i \leq \lambda_1 / H_0\right] = \alpha$$

$$p\left[\sum x_i \geq \lambda_2 / H_0\right] = \alpha$$

Note that if $x \sim E(\theta)$ then $2\theta \sum x_i \sim \chi^2_{2h}$

$$p\left(2\theta \sum x_i\right) = p\left[2\theta \sum x_i \leq 2\theta\lambda_1 / H_0\right] = \alpha$$

$$p\left(2\theta \sum x_i\right) = p\left[2\theta \sum x_i \geq 2\theta\lambda_2 / H_0\right] = \alpha$$

Using this result,

$$p\left[2\theta \sum x_i \leq \mu_1\right] = p\left[\chi^2_{2h} \leq \mu_1\right] = \alpha$$

$$\chi^2_{2h} = \mu_1$$

Hence the BCR for testing H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ $(> \theta_0)$ is given by

$$w_0 = \left\{x_i : 2\theta \sum x_i \leq \chi^2_{1-\alpha, 2n}\right\}$$

$$w_0 = \left\{x_i : \sum x_i \leq \chi^2_{1-\alpha, 2n/2\theta}\right\}$$

Since w_0 is independent of w_0, θ_0 is UMPCR for testing H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ $(> \theta_0)$ similarly

$$p\left[2\theta\sum x_i \geq 2\theta\lambda_2\right]$$
$$\alpha = p\left[\chi^2_{2n} \geq \mu_1\right] \text{ where} -$$

Hence the BCR for testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ $(< \theta_0)$ is given by

$$w_1 = \left\{x_i : 2\theta\sum x_i \geq \chi^2_{1-\alpha,2n}\right\}$$
$$w_1 = \left\{x_i : \sum x_i \geq \chi^2_{1-\alpha,2n/2\theta}\right\}$$

Since w_1 is independent of w_1, θ_1 is UMPCR for testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ $(< \theta_0)$ similarly. Since the two CR w_0 and w_1 are different there exists no CR of size α which is UMP for $H_0: \theta = \theta_0$ against $H_1: \theta \neq \theta_0$.