29 January 2024 08:04 PM

$t \rightarrow$ statistic

$\theta \rightarrow$ parameter

Statistical Inference

(1) POINT ESTIMATION

Properties of a good estimator:

(a) Unbiasedness.

$$E(t) = \theta$$

or, $E(t) - \theta = 0$

Ex: if $\bar{x}$ is sample mean and μ is population mean then if $E(\bar{x}) = \mu$ then $\bar{x}$ is unbiased estimator of μ.

(i) t is unbiased ie $E(t) = \theta$ for whatever value of θ maybe

(ii) and among all the unbiased estimators, t has the least variance ie

$$var(t) \leq var(t')$$

where t' is another statistic of a sample from same population

condition (i) and (ii) together is called a minimum-variance unbiased estimator (MVUE).

(2) Consistency: let t_n be an estimator for parameter θ based on a sample size n. Then t_n is called a consistent estimator for θ if

Then t_n is called a consistent estimator for θ if t_n converges stochastically (or in probability) to θ as $n \to \infty$, ie for any $\varepsilon > 0$

$$P\{|t_n - \theta| > \varepsilon\} \to 0 \text{ as } n \to \infty$$

∴ The sufficient condition for t_n to be a consistent estimator for θ are

(i) $E(t_n) \to \theta$ and $Var(t_n) \to 0$ as $n \to \infty$

4. EFFICIENCY:

There may exist several consistent estimators for θ.

Then in that case we will choose asymptotically normally distributed (ie tend to normality for large n) so that the rapidity of convergence may be indicated by the inverse of the variance of an asymptotic distribution.

Let 'avar' denote asymptotic variance.

Then the statistic t is called the efficient estimator of θ if t is consistent, asymptotically normally distributed and $avar(t) < avar(t')$.

consistent, ... distributed and $avar(t) \le avar(t')$.

5 SUFFICIENCY:

A necessary and sufficient condition for t to be a sufficient estimator for θ is that, the joint probability mass function of a sample obs $x_1, x_2, \dots, x_n$ should be expressible in the form

$$f(x_1, x_2, \dots, x_n ; \theta) = g(t;\theta) \cdot h(x_1, x_2, \dots, x_n)$$

Say, let's take an example of normal distribution with population mean (μ) and population variance (σ^2)

Let $x_1, x_2 \dots x_n$ be a sample from normal distribution and ($\bar{x}$) is the sample mean.

pdf of N.D is $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

$$\underbrace{f(x_1, x_2 \dots, x_n ; \theta)} = \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \right]$$

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

$(\sigma\sqrt{2\pi})$

Let us write $\Sigma(x_i-\mu)^2 = \Sigma\left[(x_i-\bar{x})+(\bar{x}-\mu)\right]^2$

$$= \Sigma(x_i-\bar{x})^2 + \Sigma(\bar{x}-\mu)^2 + 2(\bar{x}-\mu)\Sigma(x_i-\bar{x})$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$$

$$f(x_1,x_2,\dots,x_n;\theta) = \frac{1}{(\sigma\sqrt{2\pi})^n}\, e^{-\frac{1}{2\sigma^2}\left\{\underbrace{\Sigma(x_i-\bar{x})^2}_{\text{sample}} + \underbrace{n(\bar{x}-\mu)^2}_{t,\theta}\right\}}$$

$$= \frac{1}{(\sigma\sqrt{2\pi})^n}\, e^{-\frac{n(\bar{x}-\mu)^2}{2\sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\bar{x})^2}$$

$$= g(\overset{\bar{x},\mu}{t,\theta}) \cdot h(x_1,x_2,\dots,x_n)$$

$\therefore$ $\bar{x}$ is a sufficient estimator for μ.

Method of Point Estimation → Maximum Likelihood Estimator (M.L.E).

Let $x_1, x_2 \dots x_n$ be a random sample from a population with pmf $f(x;\theta)$

For a fixed θ, the function

$$L(x_1, \dots, x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

For a ...

$$f(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

The above function may be looked upon as a function of θ, which is called the likelihood function of θ and is denoted by $L(\theta)$.

The principle of maximum likelihood suggests us to take that value of estimate of θ for which $L(\theta)$ is maximum. That is we need

(i) $\dfrac{dL(\theta)}{d\theta} = 0$ (ii) $\dfrac{d^2L(\theta)}{d\theta^2} < 0$

Q: Let us consider a Bernoullian Probability mass function with 'p' as probability of success in a trial. such that,

$$x_i = \begin{cases} 1 & \text{if there is success} \\ 0 & \text{otherwise)} \end{cases}$$

then the pmf is $f(x_i; p) = p^{x_i}(1-p)^{1-x_i}$ for $x = 0, 1$.

find out the maximum Likelihood Estimator.

Soln: Likelihood function is

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

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

ie $$L(p) = \prod_{i=1}^{n} \left[p^{x_i} (1-p)^{1-x_i} \right]$$

$$L(p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{\sum_{i=1}^{n}(1-x_i)}$$

$$L(p) = p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i} \quad \text{——} \quad (1)$$

Taking log on both sides,

$$\log_e L(p) = \sum_{i=1}^{n} x_i \log_e(p) + \left(n - \sum x_i\right) \log_e(1-p)$$

F.O.C requires $\dfrac{d \log_e L(p)}{dp} = 0$

or, $$\frac{\sum x_i}{p} + \frac{(n-\sum x_i)(-1)}{1-p} = 0$$

or, $$\frac{(1-p)\sum x_i - np + p\sum x_i}{p(1-p)} = 0$$

or, $$\frac{\sum x_i - \cancel{p\sum x_i} - np \cancel{+ p\sum x_i}}{p(1-p)} = 0$$

or, $$\sum x_i - np = 0$$

or, $$np = \sum x_i$$

or, $$p = \frac{1}{n} \sum x_i$$

or, $\boxed{\hat{p} = \bar{x}}$

$\therefore$ The MLE estimate of p is

$$\hat{p} = \frac{1}{n}\sum x_i = \bar{x}$$

<u>TRY:</u> MLE for Binomial Distribution and Poison Distribution.

MLE for normal distribution:

Suppose $x_1, x_2, x_3, \ldots, x_n$ is a random sample from $N(\mu, \sigma^2)$ where both μ and σ are unknown then find the unknown parameter using MLE.

The likelihood function is

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

$$= \frac{1}{(\sigma\sqrt{2\pi})^n}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\log_e L(\mu,\sigma) = -n\log_e\sigma - n\log_e\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum(x_i-\mu)^2$$

For estimate of μ, f.o.c requires

$$\frac{d\log_e L(\mu,\sigma)}{} = 0$$

$$\frac{d \log_e L(\mu,\sigma)}{d\mu} = 0$$

or, $-\frac{1}{\cancel{2}\sigma^2} \cancel{2} \sum (x_i - \mu) \cdot (-1) = 0$

or, $\sum_{i=1}^{n} (x_i - \mu) = 0$

or, $\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \mu = 0$

or, $\sum_{i=1}^{n} x_i - n\mu = 0$

or, $\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$

or, $\boxed{\hat{\mu} = \bar{x}}$ $\Rightarrow$ MLE for $\mu = \hat{\mu} = \bar{x}$.

Similarly for σ;

$$\frac{d \log_e L(\mu,\sigma)}{d\sigma} = 0$$

or, $-\frac{n}{\sigma} - \frac{(-\cancel{2})}{\cancel{2}\sigma^3} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$

or, $-\frac{n}{\sigma} + \frac{\sum (x_i - \mu)^2}{\sigma^3} = 0$

or, $\sum (x_i - \mu)^2 \quad n$

or, $$\frac{\sum(x_i-\mu)^2}{\sigma^3} = \frac{n}{\sigma}$$

or, $$\sum(x_i-\mu)^2 = \frac{n.\cancel{\sigma^3}\sigma^2}{\cancel{\sigma}}$$

or, $$\frac{1}{n}\sum(x_i-\mu)^2 = \sigma^2$$

Let $\hat{\sigma}$ be the estimate, and we substitute μ with $\hat{\mu} = \bar{x}$, then

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

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum(x_i-\bar{x})^2}$$

_____ * _____