Estimation(contd...)
Consistency and Sufficiency

Session 6

19/03/2021

- Eg.3 : If $X_1, X_2, \ldots.., X_n$ is a random sample obtained from the density function

$$f(x,\theta) = \begin{cases} 1, & \theta < x < \theta + 1 \\ 0, & elsewhere \end{cases}$$

Show that the sample mean $\bar{x}$ is an unbiased and consistent estimator of $\left(\theta + \frac{1}{2}\right)$.

- Solution : First we need to prove that $E(\bar{X}) = \left(\theta + \frac{1}{2}\right)$.
- In order to calculate $E(\bar{X})$, we need to calculate $E(X_i)$.
- $E(X) = \int_{\theta}^{\theta+1} x.f(x,\theta)\mathrm{dx} = \int_{\theta}^{\theta+1} x.1\mathrm{dx} = \theta + \frac{1}{2}$.
- Now $E(\bar{X}) = \frac{1}{n}\sum\left(\theta + \frac{1}{2}\right) = \theta + \frac{1}{2}$.
- Thus $\bar{x}$ is an unbiased and consistent estimator of $\left(\theta + \frac{1}{2}\right)$.

- Now using sufficient condition of consistency, As n$\rightarrow \infty$, $E(\bar{X}) \rightarrow \left(\theta + \frac{1}{2}\right)$
- Also we need to prove that As n$\rightarrow \infty$, V$(\bar{X}) = 0$.
- For calculating V$(\bar{X})$, we need to calculate V(X_i).
- V$(X) = \mathrm{E}(X^2) - [\mathrm{E(X)}]^2$
- Now $E(X^2) = \int_{\theta}^{\theta+1} x^2 . f(x,\theta)\mathrm{dx} = \int_{\theta}^{\theta+1} x^2 . 1\mathrm{dx} = \theta^2 + \theta + \frac{1}{3}$.
- Therefore V$(X) == \theta^2 + \theta + \frac{1}{3} - \left[\theta + \frac{1}{2}\right]^2 = \frac{1}{12}$.
- Thus V$(\bar{X}) = \frac{1}{n^2}\mathrm{V}(\sum X_i) = \frac{1}{12n}$ which goes to 0 as n$\rightarrow \infty$. Hence $\bar{x}$ is a consistent estimator of $\left(\theta + \frac{1}{2}\right)$.

Standard results to do Intermediate steps of earlier example

$$\int x^n dx = \frac{x^{n+1}}{n+1} + c,$$

Definite Integration

- **Definition:** Let g(x) be a differentiable function of x on some interval [a , b] of R.

  If $\frac{d}{dx}\big(g(x)\big) = f(x)$ for each x in [a, b] then the definite integral of f(x) on

  [a, b] is denoted by $\int_a^b f(x)dx$ and its value is given by

$$\int_a^b f(x)dx = [g(x)]_a^b = g(b) - g(a)$$

This is called the 'Fundamental Principle of Definite Integration'.

'a' and 'b' are 'lower limit' and 'upper limit' of DI.

e.g. $\int_1^2 x^2 dx = \left[\frac{x^3}{3}\right]_1^2 = \frac{2^3}{3} - \frac{1^3}{3} = 8/3 - 1/3 = 7/3$

Sufficiency

- An estimator is said to be sufficient for a parameter, if it contains all the information in the sample regarding the parameter.
- If $T= t(x_1, x_2, \ldots.x_n)$is an estimator of a parameter θ based on a sample $x_1, x_2, \ldots.x_n$ of size n from the population with density $f(x, \theta)$ such that the conditional distribution of $x_1, x_2, \ldots.x_n$ given T , is independent of θ, then T is sufficient estimator of θ.

- In other words, the estimator T will provide all information of θ if the joint probability mass (density) function of x_1, x_2, ….x_n under the condition that T= $t(x_1, x_2, \ldots .x_n)$ is free of θ.
- i.e., $P[X_1 = x_1, X_2 = x_2, \ldots . X_n = x_n/$ T=t]

$$= \frac{P(X_1=x_1, X_2=x_2, \ldots . X_n=x_n)}{P(T=t)}$$