1st lecture

20.04.25

Statistical inference II → Estimation
→ Hypothesis

Book

1. Introduction to Theory of statistics - Mood and Graybill
2. Statistical Inference → Casella an Berger
3. Introduction to Mathematical statistics — Hogg an Craig

Modal and Median Unbiased estimator:

$\bar{x}_n$ $\frac{\sum_{i=1}^{n} x_i}{n}$ → Estimator

Statistic = 5.5 → estimate

Qn. difference among estimator, statistic & estimate

H.W. Die → toss, find the median or modal unbiased estimator.

Outcomes = 1, 2, 3, 4, 5, 6

X → random variable

Hints: $X \Rightarrow x_1, x_2, x_3, x_4, x_5, x_6$

$P(X = x_1) = 1/6, \ldots P(X = x_6) = 1/6$ → follows discrete uniform dist^n

According to question, the distⁿ of random variable is defined as discrete uniform distribution.

$$P(x;\theta) = \frac{1}{\theta} \quad ; \quad x = 0, \ldots, \theta$$

Here, $\theta = 6$

Mean $= E(x)$

$$= \sum x\, P(x)$$

$$= 1 \cdot P(x=1) + 2 \cdot P(x=2) + \cdots + 6 \cdot P(x=6)$$

$$= \frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 1$$

$$= \frac{1}{6} + \frac{1}{3} + \frac{1}{2} + \frac{2}{3} + \frac{5}{6} + 1$$

$$= \frac{1+2+3+4+5+6}{6}$$

$$= \frac{21}{6}$$

$$= 3.5$$

(*) Let x is a r.v for the value $1, 2, \ldots 6$

$P(X = x_1) = 1/6, \quad P(X = x_6) = 1/6$

which follows discrete dist^n

Mean, $E(X) = \sum_{x=1}^{6} x \cdot P(x) = 1 \times \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + 1$

$= \frac{21}{6}$

$= 3.5$

$\therefore$ Sample mean is an unbiased estimator of pop^n mean. $E(\bar{x}) = \mu_{..} = 3.5$

Median: Since the outcomes are 1, 2, 3, 4, 5, 6 The median is the average of 3rd and 4th numbers

$\text{median} = \frac{3+4}{2} = 3.5$

Mode: As the mode is the most frequent value and for a fair die toss all outcomes are equally likely (uniform dist^t). So there is no unique mode.

Median or mode unbiased estimator

Die → {1, 2, 3, 4, 5, 6} → discrete

median = ? $E[X] = ?$

If $g(x)$ = medan → median unbiased

If $E(x)$ = mode → mode unbiased

Estimator = $\bar{x} = \frac{\sum x_i}{n}$

Statistic

$$\mu = \bar{X} = \frac{X_1 + X_2 + \cdots + X_N}{N}$$

Estimator but statistic নয়

কারণ এবসুলেট pop^n

If sample নিয়ে কাজ করি, তাহলে

সেটা estimator and statistic

হবে

Properties of good estimators:

1. Unbiasedness → $E[\text{statistic}] = \text{Pop}^n$ Parameter
2. Consistency
3. Sufficiency
4. Efficiency

estimator → unbiased estimator

$E[\bar{x}] = \mu$

$E(s^2) = \sigma^2$

HW $s^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})}{n-1}$ is a biased estimator of pop^n variance σ^2

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

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

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

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

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

$$= \frac{1}{n}\sum \text{var}(x_i) - \text{var}(\bar{x})$$

$$= \frac{1}{n}\sum_{i=1}^{n}\sigma^2 - \frac{\sigma^2}{n}$$

$$= \sigma^2 - \frac{\sigma^2}{n}$$

$$= \left(\frac{n-1}{n}\right)\sigma^2$$

$\therefore S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is biased estimator of σ^2

$E(S^2) = \left(\frac{n-1}{n}\right)\sigma^2$

$\Rightarrow E\left[\frac{nS^2}{n-1}\right] = \sigma^2$

$E\left[\frac{n}{n-1}\cdot\frac{1}{n}\sum(x_i - \bar{x})^2\right] = \sigma^2$

$\Rightarrow E\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right] = \sigma^2$

So, $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ is an unbiased estimate of σ^2

- Amount of bias = Estimate - Parameter

$$= \left(\frac{n-1}{n}\right)\sigma^2 - \sigma^2$$

$$= -\frac{\sigma^2}{n}$$

$\therefore -\frac{\sigma^2}{n}$ is the amount of the bias

Lecture 5th & 6th

22-06-25

* Properties of good estimator

(i) Unbiasedness
(ii) Consistency
(iii) Efficiency
(iv) Sufficiency

Ex: If X_i is a Bernoulli random variable with parameter p, then:

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is the maximum likelihood estimator (MLE) of p. Is the MLE of p an unbiased estimator of p?

$$f(x;\lambda) = \log\left[\frac{e^{-\lambda}\lambda^{x}}{x!}\right]$$

$$\frac{\partial}{\partial\lambda}\left[\log\left(\frac{e^{-\lambda}\lambda^{x}}{x!}\right)\right] = \exp\left[\frac{\partial}{\partial\lambda}\left[\log\left(\frac{e^{-\lambda}\lambda^{x}}{x!}\right)\right]\right]$$

$$\text{Var}(\hat{\lambda}) = \frac{1}{n^2}\,\text{Var}\left[\sum_{i=1}^{n} x_i\right]$$

$$= \frac{n\lambda}{n^2} = \frac{\lambda}{n}$$

same করে Binomial করবো।

$$P[(0,1,1)/S] = ?$$

$$P[(0,1,1)/T] = ?$$

(2)

$$P = [X_1 = 0, X_2 = 1, X_3 = 1 \mid S = 2]$$

$$= \frac{P^2(1-P)}{P(S=2)} = \frac{P^2(1-P)}{\binom{3}{2}\cdot P^2(1-P)} = \frac{1}{3}$$

$$P(0,1,1 \mid 1) = \frac{P[X_1=0,\ X_2=1,\ X_3=1;\ T=2]}{P[T=2]}$$

$$= \frac{p^2(1-p)}{(1-p)^2 + p^2}$$

$$= \frac{p}{1+p}$$

* Median & Unbiased Estimator:

Let, $X_1, X_2, \ldots, X_n$ be iid random variables with pdf $f(x;\theta)$ and let $t(x_1, \ldots, x_n)$ be a statistic such that the median of the pdf of the statistic t is θ. Then the statistic $t(X_1, \ldots, X_n)$ is defined to be an median estimate of θ.

For discrete dist^n. median will be

$$\text{Median} = \sum_{x=0} P_n(X=x) = \frac{1}{\theta}$$

and for continuous dist^n.

$$\text{Median} = \int_{\text{min Range}}^{\text{Median}} f(x;\theta) = \frac{1}{2}$$

$$= \int_{\text{median}}^{\text{Max Range}} f(x;\theta) = 1/2$$

To estimate the pop^n. parameter θ by statistic t, we need to know about bias.

So Bias = $\theta - E(t)$

To make the statistic t unbiased, bias should be zero.

$\therefore$ Bias = 0 $\Rightarrow$ $\theta - E(t) = 0$

$\Rightarrow E(t) = \theta$

$P_\theta\{E(t) > \theta\} = P_\theta\{E(t) < \theta\}$

so, if these two inequalities holds for both sides, then the statistic t will be median unbiased estimator of popⁿ parameter θ.

Lecture 7th & 8th

30-06-25

Modal Unbiased Estimator:

Let, $x_1, x_2, \ldots, x_n$ be iid distributed random variables with probability density function $f(x_i; \theta)$ and let $t(x_1, x_2, \ldots, x_n)$ be a statistic. Such that the mode of the density function of statistic t is θ. Then the statistic t is said to be a modal unbiased estimate of θ.

$E(\bar{x}) = \theta$ mean
↳ mode
$E(\bar{x}) = \theta$ median

H-W

Example:

Let $X_1, \ldots, X_n$ be iid random samples from a continuous uniform dist^n

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}; & 0 \le x \le \theta \\ 0; & \text{otherwise} \end{cases}$$

Find the modal unbiased estimator of the pop^n parameter θ.

Ex: Let $X_1, \ldots, X_n$ be iid random samples from an exponential dist^n.

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

Let the two statistic,

$Y_1 = \frac{X_n}{\log_n}$, where $X_n = \max\{X_1, X_2, \ldots, X_n\}$

and $Y_2 = \frac{\delta_n}{n-1}$, where $\delta_n = \sum_{i=1}^{n} X_i$

Prove that both statistics Y_1 and Y_2 are the modal unbiased estimate of the

pop^n parameter θ.

Sol^n: the prob^y density function of the highest order statistic, x_n is

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

$$= n\left[\int_0^{x_n} \frac{1}{\theta} e^{-x/\theta}\, dx\right]^{n-1} \frac{1}{\theta} e^{-x_n/\theta} \quad ; x_n > 0$$

$$= n \cdot \frac{1}{\theta}(-\theta)\left[e^{-x/\theta}\right]_0^{x_n} \cdot \frac{1}{\theta} e^{-x_n/\theta}$$

$$= \frac{n}{\theta}\left[\frac{n}{\theta}\left[\frac{1}{\theta} e^{-x_n/\theta}\, dx\right]\right]$$

$$= \frac{n}{\theta}\left[\left(\frac{1}{\theta} \cdot e^{-x_n/\theta} \cdot \frac{-1}{1/\theta}\right)_0^{x_n}\right]^{n-1} e^{-x_n/\theta}$$

$$= \left[\frac{n}{\theta}\left[-e^{-x_n/\theta} - 0 + e^{0}\right]\right]^{n-1} \cdot e^{x_n/\theta}$$

$$= n\left[\frac{n}{\theta}\left[1 - e^{-x_n/\theta}\right]\right]^{n-1} \cdot e^{-x_n/\theta}$$

Given that,

$$y_2 = \frac{x_n}{\log n}$$

$$\Rightarrow x_n = y_1 \log n$$

$$\Rightarrow \frac{\partial x_n}{\partial y_1} = \log n$$

$$\Rightarrow |J| = \left| \frac{\partial x_n}{\partial y_1} \right| = \log n$$

$$\therefore g(y_1; \theta) = g(x_n; \theta) \cdot |J|$$

$$= \frac{n}{\theta}\left[1 - e^{-x_n/\theta}\right]^{n-1} e^{-x_n/\theta} \cdot \log n$$

$$= \frac{n \log n}{\theta}\left[1 - e^{-x_n/\theta}\right]^{n-1} \cdot e^{-x_n/\theta}$$

$$e^{-\frac{y_1 \log n}{\theta}}$$

$$= \frac{n \log n}{\theta}\left[1 - e^{-\frac{y_1 \log n}{\theta}}\right]^{n-1} e^{-\frac{y_1 \log n}{\theta}}$$

$$= \frac{n \log n}{\theta}\left[1 - n^{-y_1/\theta}\right]^{n-1} n^{-y_1/\theta}$$

$$= \frac{n \log n}{\theta} \left[1 - \frac{n y_1}{\theta}\right]^{n-1} \cdot \left(-\frac{n y_1}{\theta}\right) \Rightarrow$$

Now,

Let,

$$n^{-1/\theta} = m$$

$$\therefore \frac{n \log n}{\theta} \left[1 - m^{y_1}\right]^{n-1} \cdot m^{y_1}$$

Differentiate both sides w.r.t to y_1 and put equal to zero

$$\frac{\partial g(y_1; \theta)}{\partial y_1} = 0$$

$$\Rightarrow \frac{\partial}{\partial y_1} \left[\frac{n \log n}{\theta} \left\{1 - m^{y_1}\right\}^{n-1} \cdot m^{y_1}\right] = 0$$

$$\Rightarrow \frac{\partial}{\partial y_1} \left[\left(1 - m^{y_1}\right)^{n-1} \cdot m^{y_1}\right] = 0$$

$$\Rightarrow (1-m^{y_1})^{n-1} m^{y_1} \log n + m^{y_1} (n-1)(1-m^{y_1})^{n-2} \times (0 - m^{y_1} \log m) = 0$$

$$\left[\because \frac{\partial}{\partial x} a^x = a^x \log a\right]$$

$$\Rightarrow (1-m^{y_1})^{n-1} m^{y_1} \log m - m^{2y_1} (n-1)(1-m^{y_1})^{n-2} \log m = 0$$

$$\Rightarrow \log m \left[(1-m^{y_1})^{n-1} m^{y_1} - m^{2y_1} (n-1)(1-m^{y_1})^{n-2}\right] = 0$$

$$\Rightarrow (1-m^{y_1})^{n-1} m^{y_1} = m^{2y_1} (n-1)(1-m^{y_1})^{n-2}$$

$$\Rightarrow \frac{(1-m^{y_1})^{n-1}}{(1-m^{y_1})^{n-2}} = \frac{(n-1)\, m^{2y_1}}{m^{y_1}}$$

$$\Rightarrow (1-m^{y_1})^{n-1-n+2} = (n-1)\, m^{2y_1 - y_1}$$

$$\Rightarrow (1-m^{y_1}) = (n-1) m^{y_1}$$

$$\Rightarrow 1 - m^{y_1} = n m^{y_1} - m^{y_1}$$

$$\Rightarrow n m^{y_1} = 1$$

$$\Rightarrow y_1 \log nm = 0$$

$$\Rightarrow m^{y_1} = \frac{1}{n}$$

$$\Rightarrow (n^{-1/\theta})^{y_1} = \frac{1}{n}$$

$$\Rightarrow n^{-y_1/\theta} = n^{-1}$$

$$\Rightarrow \hat{y}_1 = \theta$$

For y_2,

$$y_2 = \frac{s_n}{n-1}$$

$$\therefore s_n = y_2 (n-1)$$

We know that the prob^y density

$$f(x; \alpha, \beta) = \frac{\beta^{\alpha} \cdot e^{-\beta x} \cdot x^{\alpha-1}}{\Gamma\alpha}$$

$$\therefore f(s_n; n; \theta) = \frac{1}{\Gamma n} \left(\frac{1}{\theta}\right)^n \cdot e^{-s_n/\theta} \cdot s_n^{n-1} \; ; \; s_n > 0, \theta > 0$$

Given that

$$y_2 = \frac{s_n}{n-1}$$

$$\Rightarrow \frac{2s_n}{2y_2} = n-1 \qquad \Rightarrow s_n = y_2 (n-1)$$

$$\Rightarrow |J| = \left| \frac{\partial s_n}{\partial y_2} \right| = (n-1)$$

Prob^y of y_2 is

~~$g_2(y_2;\theta) \cdot g(s_n; n, \theta)|J|$~~

$$f(s_n; n, \theta) = \frac{1}{\lfloor n} \left(\frac{1}{\theta}\right)^n \cdot e^{\frac{-(n-1)y_2}{\theta}} \left[y_2(n-1)\right]^{(n-1)}$$

$$= \frac{1}{\lfloor n} \left(\frac{1}{\theta}\right)^n \cdot e^{\frac{-(n-1)y_2}{\theta}} \left[y_2(n-1)\right]^{(n-1)}$$

$$= \frac{(n-1)^{n-1+1}}{\theta^n \lfloor n} \cdot e^{-\frac{(n-1)y_2}{\theta}} y_2^{n-1}$$

$$\frac{\partial}{\partial y_2} \{g(y_2;\theta)\} = 0$$

$$\Rightarrow \frac{\partial}{\partial y_2} \left[e^{\frac{-(n-1)y_2}{\theta}} \cdot y_2^{n-1} \right] = 0$$

$$\therefore \hat{y}_2 = \theta$$

H.W → Download book — Stuart ~~and~~ Kendal and Ord. Keith

⊛ Resampling techniques ⊛

$$\bar{t}_n = n\bar{t}_n + (n-1)\bar{t}_{n-1}$$

$$t_1 = \hat{\theta}_1 = 4 \times 2.5 + (4-1) \times 2$$

$$\hat{\theta}_2 = 4 \times 2.5 + 3 \times 2.33$$

$$\hat{\theta}_3 = 4 \times 2.5 + 3 \times 2.67$$

$$\hat{\theta}_4 = 4 \times 2.5 + 3 \times 3$$

$$\bar{x} = \frac{2+3+4}{4} = 2.5$$

$$J_1 = \frac{1+2+3}{3} = 2$$

$$J_2 = \frac{1+2+4}{3} = 2.33$$

$$J_3 = 2.67$$

$$J_4 = 3$$

L-9th & 10th

10-07-25

Resampling Technique:

a) Jacknife

Ex. We have 4 original samples

1, 2, 3, 4

$$\bar{x} = \frac{1+2+3+4}{4} = 2.5 = \frac{\sum X_i}{N}$$

~~Jacknife estimate = $n\bar{x} - (n-1)\bar{x}_{i-1}$~~

the four delete - one value are

$$\hat{x}_1 = \frac{1+2+3}{3} = 2$$

$$\hat{x}_2 = \frac{2+3+4}{3} = 3$$

$$\hat{x}_3 = \frac{3+4+1}{3} = 2.67$$

$$\hat{x}_4 = \frac{4+1+2}{3} = 2.33$$

The four pseudo values / Jacknife samples are

$\hat{\theta}_1 = n\bar{X} - (n-1)\bar{X}_1 = 4 \times 2.5 - 3 \times 2 = 4$

similarly

$\hat{\theta}_2 = 1 \; ; \quad \hat{\theta}_3 = 2 \; , \quad \hat{\theta}_4 = 3$ ← absolute value

In general, Jacknife estimate $= n\bar{X} - (n-1)\bar{X}_{n-k}$

$k \leq n$

$$\text{Mean} = \frac{\sum \hat{\theta}_i}{n} = \hat{\theta}_Q$$

$$\text{var}(\hat{\theta}_Q) = \frac{\text{var}(\hat{\theta}_i)}{n} = \frac{\sum (\hat{\theta}_i - \hat{\theta}_Q)^2}{n(n-1)}$$

$$Sd(\hat{\theta}_Q) = \sqrt{\text{var}(\hat{\theta}_Q)}$$

$$C.I - \hat{\theta}_Q \pm t_{\alpha/2,\, n-1} \sqrt{\text{var}(\hat{\theta}_Q)}$$

H.W

Ex: 3.2

17.23, 13.93, 15.78, 14.91, 18.21, 14.28, 18.83, 13.45, 18.71, 18.81, 11.29, 13.39, 11.57, 10.94, 15.52, 15.25,

$S^2 = 7.171, \quad \bar{x} = 1.9701$

Find out the Jacknife samples with estimated mean variance and 95% C.I

r (*)

What is Jacknife estimation method ?

Let, $S^2 = \frac{\Sigma(x_i - \bar{x})^2}{N}$: it is consistent estimator if $n \to \infty$. Like S^2, if we have a consistent estimator but biased estimator t_n of θ then we can add a simple adjustment factor to make an unbiased estimator.

Maurice - Quenouille (1956) gives a method of overcoming this type of biasness. This method is called Jackknife estimation method.

For example,

$$E(S^2) = \frac{N-1}{N}\mu_2 - \frac{1}{n}\mu_2$$

$$= \mu_2 - \frac{1}{n}\mu_2 + \frac{1}{n}\mu_2$$

Let, t_n be a biased estimator of θ based on $x_1, x_2, \ldots, x_n$ observations.

$$E(t_n) = \theta + \sum_{n=1}^{\infty} \frac{a_n}{n^n}$$

Let t_{n-1} be an another biased estimate of θ based on $(n-1)$ observations.

So, according to jackknif method of estimation the Jackknife estimat is

$$t_n' = nt_n - (n-1)\bar{t}_{n-1}$$

$$= t_n - t_n + nt_n - (n-1)\bar{t}_{n-1}$$

$$= t_n + t_n(n-1) - (n-1)\bar{t}_{n-1}$$

$$= t_n + (n-1)\left[t_n - \bar{t}_{n-1}\right]$$

$$E[t_n'] = E[t_n] + (n-1)E\left[t_n - \bar{t}_{n-1}\right]$$

$$= \theta + \sum_{n=1}^{\infty} \frac{a_n}{n^n} + (n-1)\left[\theta + \sum_{n=1}^{\infty} \frac{a_n}{n^n} - E\left(\theta - \sum_{n=1}^{\infty} \frac{a_n}{(n-1)^n}\right)\right]$$

$$\Rightarrow E(t_n') = \theta + \sum_{n=1}^{\infty} \frac{a_n}{n^n} + n \cdot \sum \frac{a_n}{n^n} - \sum_{n=1}^{\infty} \frac{a_n}{n^n}$$

$$(n-1) \sum \frac{a_n}{(n-1)^n}$$

$$= \theta + n \sum_{n=1}^{\infty} \frac{a_n}{n^n} - (n-1) \sum_{n=1}^{\infty} \frac{a_n}{(n-1)^n}$$

$$= \theta + \sum_{n=1}^{\infty} \frac{a_n}{n^{n-1}} - \sum_{n=1}^{\infty} \frac{a_n}{(n-1)^{n-1}}$$

If $n=1$

$$= \theta + 0 + \left\{ \frac{a_2}{n} - \frac{a_2}{(n-1)} \right\} + \left\{ \frac{a_3}{n^2} - \frac{a_2}{(n-2)^2} \right\}$$

$$+ \dots$$

$$\therefore E(t_n') = \theta + 0 - a_2 \left(\frac{1}{n(n-1)} \right) + \text{higher orders}$$

$$E(t_n) = \theta + \sum \frac{a_n}{n^n}$$

$$E(t_n') = n t_n' - (n-1) t_{(n-1)}$$

H.W

Example:

Find out the unbiased estimator of θ^2 in Binomial dist^n according to the Jackknife method of estimation.

Sol^n:

The prob^y density f^n of Binomial is

$$P_r\{x; \theta, n) = {}^nC_r \cdot \theta^r (1-\theta)^{n-r}$$

where, $\theta = r/n$, $1-\theta = 1-\frac{r}{n} = \frac{n-r}{\theta}$ ——— (1)

Let, $\frac{r}{n} = \theta$ is an unbiased estimator but $\left(\frac{r}{n}\right)^2 = \theta^2$ is not an unbiased estimator.

Let the intuitive estimator is

$$t_n = \theta^2 = \left(\frac{r}{n}\right)^2$$

$$E[t_n] = E\left[\left(\frac{r}{n}\right)^2\right]$$

$$= \text{var}\left(\frac{r}{n}\right) + \left\{E\left(\frac{r}{n}\right)\right\}^2$$

$V(x) = E(x^2) - \{E(x)\}^2$

$E(x^2) = V(x) + \{E(x)\}^2$

$$= \frac{1}{n^2} \text{var}(r) + \theta^2$$

$$\frac{1}{n^2} n\theta(1-\theta) + \theta^2$$

$$= \frac{\theta(1-\theta)}{n} + \theta^2$$

$$\neq \theta^2$$

According to Jackknife method the Jackknife estimate is

$$t_n' = n t_n - (n-1)\bar{t}_{n-1} \quad \text{---} (1)$$

Now, $\bar{t}_{n-1}$ can be taken as $\left(\frac{r-1}{n-1}\right)^2$ if a success is omitted from the sample with prob^y $\theta = \frac{r}{n}$.

Or, t_{n-1} can be taken as $\left(\frac{r}{n-1}\right)^2$ if a failure is omitted from the sample with prob $= 1 - \theta = 1 - \frac{r}{n} = \frac{n-r}{n}$

$$F_{n-1} = s + f$$

$$= \left(\frac{n-1}{n-1}\right)^2 \frac{n}{n} + \left(\frac{n}{n-1}\right)^2 \cdot \frac{n-n}{n}$$

$$= \frac{1}{n(n-1)^2}\left[n + n^2(n-2)\right]$$

① $\Rightarrow$

$$t_n' = nt_n - (n-1)t_{n-1}$$

$$= n\left(\frac{n}{n}\right)^2 - (n-1)\left[\frac{1}{n(n-1)^2}\left\{n + n^2(n-2)\right\}\right]$$

$$= \frac{n^2}{n} - \frac{1}{n(n-1)}\left\{n + n^2(n-2)\right\}$$

$$= \frac{n(n-1)}{n(n-1)}$$

$$E[t_n'] = E\left[\frac{n(n-1)}{n(n-1)}\right]$$

$$= \theta^2$$

HW

(8) Given,

Number of observations:

$n = 16$

$$\sum x_i = \frac{17.23 + 13.93 + \cdots + 10.94}{16} = 242.20$$

$$\text{Sample mean, } \bar{x} = \frac{242.20}{16} = 15.13$$

(b)

Jackknife sample means: $\bar{x}_{(i)}$

$$\bar{x}_{(i)} = \frac{\sum x - x_i}{15}$$

We remove each observation one by one, then calculate the sample mean of the remaining 15 observations:

Formula:

$$\bar{x}_{(i)} = \frac{\sum x - x_i}{15}$$

i	x_i	$\bar{x}_{(i)} = \frac{24220 - x_i}{15}$
1	17.23	$\frac{224.97}{15} = 14.33$
2	13.93	$\frac{228.27}{15} = 15.05$
3	15.78	$\frac{226.42}{15} = 15.08$

i	x_i	$x_{(i)} = \frac{242.20 - x_i}{15}$
4	14.91	15.18
5	18.21	14.79
6	14.28	15.19
7	18.83	14.54
8	13.45	15.25
9	18.71	14.57
10	18.81	14.57
11	11.29	15.39
12	13.39	15.20
13	11.57	15.42
14	10.94	15.44
15	15.52	15.11
16	~~15.85~~ 15.25	15.13

Jackknife variance

$$Var_{jack} = \frac{n-1}{n} \sum_{i=1}^{n} \left(\bar{x}_{(i)} - \bar{x}\right)^2$$

where $\bar{x} = 15.13$

Now,

i	$x_{(i)}$	$(x_{(i)} - \bar{x})^2$
		$(14.33 - 15.13)^2 = 0.6468$
1	14.33	
2	15.05	0.0068
3	15.08	0.0032
4	15.18	0.0023
5	14.79	0.1153
6	15.19	0.0038
7	14.54	0.3505
8	15.25	0.0126
9	14.57	0.3115
10	14.57	0.3198
11	15.39	0.0663
12	15.20	0.0048
13	15.42	0.0807
14	15.44	0.0952
15	15.11	0.0006
16	15.13	0.000

$$\sum (\bar{x}_{(i)} - \bar{x})^2 = 2.0199$$

$$\text{Var}_{\text{Jack}} = \frac{15}{16} \times 2.0199 = 1.8937$$

$$\therefore \text{SE} = \sqrt{1.8937} = 1.3769$$

& $t_{0.025,15} = 2.131$

95% CI is given by

$$\text{CI} = \bar{x} \pm t \times \text{SE} = 15.1375 \pm 2.131 \times 1.3769$$

$$= 15.1375 \pm 2.9345$$

$$= (12.20, 18.07)$$

Book Effron : An introduction to Bootstrap

Lecture - 11 & 12

13-07-25

Bootstrap method:

5, 9, 7, 3, 11 → resample করতে হবে।
↳ $\bar{x}$

5 times.

S = n_1 = Bootstrap sample 01 = 9, 7, 9, 9, 11 = $B_1 = x_1^* = \bar{x}_1^*$ mean

S = n_2 = " " 02 =

S = n_3 = " " 03 =

S = n_4 = " " 04 = $\bar{x}_5^*$

S = n_5 = " " 05 = x_5^*

Bootstrap mean, $\hat{\mu}_B = \sum_{i=1}^{B} \frac{\bar{X}_i^*}{B} = \frac{1}{B} \sum_{i=1}^{B} \bar{X}_i^*$

" S.d =

" S. Error = $\sqrt{\frac{\sum_{i=1}^{B} (\bar{X}_i^* - \bar{x})^2}{B-1}}$

" C.I ⇒

↳ = $\hat{\mu}_B \pm t_{\alpha/2} \sqrt{\text{var}(\hat{\mu}_B)}$

Ex → 4.1

Q1 Consider the following dataset:

81, 70, 63, 54, 74, 65, 98, 71

(i) Generate 10 Bootstrap sample

(ii) Calculate the Bootstrap Mean, S.E and 95% C.I using Mean.

Solution:

(i) We will randomly select 8 numbers with replacement from the dataset 10 times.

The bootstrap samples and their means are:

Sample No.	Bootstrap Sample	Mean ($\bar{x}_i^*$)
1	70, 74, 81, 54, 98, 70, 65, 74	73.25
2	54, 54, 81, 65, 63, 70, 71, 74	66.50
3	98, 98, 63, 65, 74, 71, 70, 70	76.13
4	63, 63, 65, 65, 70, 71, 74, 70	67.63
5	98, 81, 98, 74, 81, 98, 63, 54	80.88
6	54, 54, 54, 54, 81, 70, 63, 74	63.00
7	81, 81, 70, 65, 74, 71, 81, 70	74.13
8	98, 98, 98, 98, 81, 81, 81, 81	89.50
9	54, 54, 54, 54, 54, 54, 54, 54	54
10	63, 63, 63, 63, 63, 63, 63, 63	63

(ii) Bootstrap Mean

$$\bar{x}_{boot} = \frac{1}{10}\sum_{i=1}^{10}(\bar{x}_i^*)$$

$$= \frac{1}{10}(73.25 + 66.5 + 76.13 + 67.63 + 80.88 + 63 + 74.13 + 89.5 + 54 + 63)$$

$$= \frac{708.02}{10}$$

$$= 70.80$$

Standard error,

$$S.E = \sqrt{\frac{\sum_{i=1}^{10}(\bar{x}_i^* - \bar{x}_{boot})^2}{B-1}}$$

$$= \sqrt{\frac{(73.25-70.8)^2 + (66.5-70.8)^2 + \cdots + (63-70.8)^2}{10-1}}$$

$$= \sqrt{\frac{927.56}{9}}$$

$$\approx 10.15$$

∴ 95% Confidence interval

$CI = \bar{x}_{boot} \pm 1.96 \times S.E$

$= 70.80 \pm 1.96 \times 10.15$

$= 70.80 \pm 19.89$

$\therefore CI = [50.91, 90.69]$

∴ Bootstrap mean = 70.80

Bootstrap standard error = 10.15

95% C.I : [50.91, 90.69]

Bootstrap sample:

Suppose we have observed n independent random samples, say, $X = x_1, x_2, x_3, \ldots, x_n$

We compute a statistic, $s(x)$

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

A bootstrap sample $x^* = x_1^*, x_2^*, \ldots, x_B^*$ is obtained by randomly selecting n times with replacement from the original data, $x_1, x_2, \ldots, x_n$: where each of sample size is n.

The Bootstrap estimate of the standard error is the standard deviation of the Bootstrap samples.

$$S.E = \sqrt{\frac{\sum_{i=1}^{B}\left[\left(s(x^*)_b^* - \bar{s}(x)\right)\right]^2}{(B-1)}}$$

) where, $S(x^*)_b = \sqrt{\dfrac{\sum_{i=1}^{B} (s(x_i^*) - \bar{x}_i^*)^2}{B-1}}$

$$\bar{x}_i^* = \frac{\sum_{i=1}^{B} x_i^*}{B} \quad \text{and}$$

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

Bias = statistic - parameter

= statistic - E(statistic)

= 0

$= S(x) - E[S(x)]$

$= S(x) - t(F)$

where $t(F)$ is the estimated value of the respective statistic, which is known as plug-in estimate.

⊛ H.W → How can you improve the bias by using Bootstrap sample.

Exam. 2MARK

Also. find out 95% C.I using different Statistic including median, standard

deviation, mode, instead of mean and variance.

Solution

we will now find bootstrap 95% CI

$X = \{4, 5, 6, 8, 9, 10, 10, 12\}$

Sample no	Bootstrap sample	median	std	mode
1	{5, 5, 6, 8, 10, 10, 12, 9}	8.5	2.4	5, 10
2	{6, 6, 6, 8, 9, 10, 12, 12}	8.5	2.2	6
3	{4, 5, 8, 8, 9, 10, 10, 12}	8.5	2.5	8, 10
4	{5, 6, 6, 8, 9, 9, 10, 10}	8	1.7	6, 9, 10
5	{4, 4, 6, 8, 9, 10, 12, 12}	8.5	2.8	4, 12

medians

$\{8.5, 8.5, 8.5, 8, 8.5\}$

sorted : $\{8.0, 8.5, 8.5, 8.5, 8.5\}$

CI = $\{8, 8.5\}$

For s.d

{2.4, 2.2, 2.5, 1.7, 2.8}

sorted = {1.7, 2.2, 2.4, 2.5, 2.8}

C.I = (1.7, 2.8)

mode CI = (6, 10)

Lecture -13 & 14

20-7-25

Vector of Parameter:

~~$f(X_1, \theta)$~~

$X_1, X_2, \ldots, X_n$ are random variables with parameters $\theta_1, \theta_2, \ldots, \theta_k$

$\underline{\theta} = \{\theta_1, \theta_2, \ldots, \theta_k\}$

$T_1, T_2, \ldots, T_k$ are estimators

$$E\{T_1, T_2, \ldots, T_k\} \approx \{T_1, T_2, \ldots, T_k\} = \{\theta_1, \theta_2, \ldots \theta_k\}$$

Assuming that $x_1, x_2, \ldots, x_n$ be a random sample with size n from the prob^y density

function

$f(x; \theta_1, \theta_2, \ldots, \theta_k)$. where

$\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ are the parameters and $\underline{\theta} = \{\theta_1, \theta_2, \ldots, \theta_k\}$ is the parameter space or vector of parameters

We are interested to estimate, $T_1(\theta) = \theta_1$, $T_2(\theta) = \theta_2, \ldots, T_k(\theta) = \theta_k$

If we let $k = 1$, we get the point estimate But a very special case is $k = n$. In this case, we will get n number of parameters. A point estimator of $T_1(\theta), T_2(\theta), \ldots T_n(\theta) = \theta_n$ is a vector of statistic as $T_1, T_2, \ldots, T_n$.

$T_1 = t_1(x_1, \ldots, x_n) = T_1(\theta)$

$T_2 = t_2(x_1, \ldots, x_n) = T_2(\theta)$

$\vdots$

$T_n = t_n(x_1, \ldots, x_n) = T_n(\theta)$

where $T_1, T_2, \ldots, T_n$ are the n sample observations and $\tau_1(\theta), \tau_2(\theta), \ldots, \tau_n(\theta)$ are the function of parameter θ.

Our next task is to find out how the dist^n of the estimators $T_1, T_2, \ldots, T_n$ be be concentrated around $\tau_1(\theta), \tau_2(\theta), \ldots, \tau_n(\theta)$

We can find out the concentrations of the estimators $T_1, T_2, \ldots, T_n$ to the parameters $\tau_1(\theta), \tau_2(\theta), \ldots, \tau_n(\theta)$ by using the following four methods.

(a) Vector of variances.

(b) Linear combination of variances

(c) Ellipsoid of concentration.

(d) Wilk's generalised variance.

H.W
Ellipsoid

Ellipsoid of Concentrations:

The ellipsoid of concentration is a region around the mean that contains the most likely values of the vectors of estimators, $T = (T_1, T_2, \ldots, T_n)$. Let $T_1, T_2, \ldots, T_n$ be an unbiased estimator of the popⁿ parameters, $\tau_1(\theta), \tau_2(\theta), \ldots, \tau_n(\theta)$ be an unbiased estimators of the popⁿ parameters $\tau_1(\theta), \tau_2(\theta), \ldots, \tau_n(\theta)$. Also let $\sigma^{ij}(\theta)$ be the ijth element of the inverse of the covariance matrix of the estimators $(T_1, T_2, \ldots, T_n)$ where $\sigma_{ij}(\theta) = \text{cov}[T_i, T_j]$

H.W → (do practical example of Ellipsoid of concentration and Wilk's generalised variance)

Person	Height	weight
A	160	60
B	165	65
C	170	70
D	175	75

$$\bar{x} = \begin{bmatrix} 167.5 \\ 67.5 \end{bmatrix}$$

$$\text{var}(h) = \frac{(160-167.5)^2 + \dots + (175-167.5)^2}{n-1} = 41.67$$

$$\text{cov}(h,w) = \frac{(160-167.5)(60-67.5) + \dots + (175-167.5)(75-67.5)}{3}$$

$$= 41.67$$

$$\therefore S = \begin{bmatrix} 41.67 & 41.67 \\ 41.67 & 41.67 \end{bmatrix}$$

Exam ৭ প্রস্তুতি শুরু

Lecture 15 & 16

27-07-25

Minimum Variance Band Bound Estimator

$X_1, X_2, \ldots, X_n \rightarrow f(x_i; \theta)$

$\theta_1, \theta_2, \ldots \theta_m$

$\text{var}(\theta_1) \leq \text{var}(\theta_i) \quad ; \; i = 2, \ldots, m$

↳ Minimum variance bound estimator

Book - Stuart Kendal

* Prove that,
the minimum variance Bound estimator is an unique estimator.

Proof. In many situation when although the condition of attending of MVB estimator satisfy but still condition there ~~many~~ may be a MV estimator exists, it is unique irrespective of whether a MVB exists or not

Let t_1 and t_2 be two distinct estimators. Also let t_1 and t_2 be two MV unbiased estimators of $\tau(\theta)$, each with variance V. let t_3 a new estimator such that

$$t_3 = \frac{1}{2}(t_1 + t_2)$$

$$E(t_3) = \frac{1}{2}\left[E(t_1) + E(t_2)\right] = \frac{\tau(\theta) + \tau(\theta)}{2}$$

$$= \tau(\theta)$$

$$\text{Var}(t_3) = \text{Var}\left[\frac{1}{2}(t_1 + t_2)\right]$$

$$= \frac{1}{4}\text{Var}(t_1 + t_2)$$

$$= \frac{1}{4}\text{Var}(t_1) + \frac{1}{4}\text{Var}(t_2) + 2\text{cov}(t_1, t_2)$$

$$= \frac{1}{4} \cdot V + \frac{1}{4}V + 2\,\text{cov}(t_1, t_2) \quad \text{---(1)}$$

According to the Cauchy Schwarts inequality we get,

$$\text{cov}(t_1, t_2) = \iint \cdots \int \left[t_1 - \tau(\theta)\right]\left[t_2 - \tau(\theta)\right] L \, dx_1 x_2 \cdots x_n$$

$$\leq \left[\int \cdots \int \{[t_1 - \tau(\theta)]^2 L\, dx_1, x_2, \ldots, x_n\} \times \int \cdots \int [t_2 - \tau(\theta)]^2 [dx_1, x_2, \ldots, x_n]\right]^{1/2}$$

$$\text{cov}(t_1, t_2) \leq \sqrt{\text{var}(t_1) \times \text{var}(t_2)}$$

$$\leq \sqrt{V \times V}$$

$$\leq V$$

From eqn (1) we get

$$\text{var}(t_3) \leq \frac{1}{4}(V+V) + \frac{1}{4} \cdot 2V$$

$$\leq V \quad \text{---} \quad \text{(ii)}$$

the eqn (ii) contradicts the assumption that t_1 and t_2 have MV unless the equality holds

This means that the equality sign holds in eqn (ii)

$$\text{cov}(t_1, t_2) = V \quad \text{——(iii)}$$

We can also say that

$$[t_1 - \tau(\theta)] \propto [t_2 - \tau(\theta)]$$

$$\Rightarrow [t_1 - \tau(\theta)] = k(\theta)[t_2 - \tau(\theta)] \quad \text{——(iv)}$$

$$\therefore \text{cov}(t_1, t_2) = k(\theta) \cdot \text{var}(t_2) \text{ var}(t_2)$$

$$t_1 - \tau(\theta) = 1\,[t_2 - \tau(\theta)]$$

$$\therefore t_1 = t_2$$

UMVUE:

Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(\cdot\,; \theta)$. An estimator $T^* = t(X_1, \ldots, X_n)$ of $\tau(\theta)$ is defined to be an UMVUE of $\tau(\theta)$ if:

① =

Lecture - 17 & 18

03-08-25

Lower bound for variance:

1st tutorial syllabus:

1. Median and Model unbiased estimator
2. Jacknife and Bootstrap
3. Properties of good estimators
4. Cramer-e Rao inequality, Bhattacharyea Bound.
5. Pitman estimator for location and scale.

Lower bound for variance:

Let, $X_1, X_2, \ldots, X_n$ be a random sample drawn from $f(x_i; \theta)$, $\theta \in \overline{\Theta}$. Also let $T = t(X_1, X_2, \ldots, X_n)$ be an unbiased estimator of $\tau(\theta)$. Here we will consider that $f(x_i; \theta)$ is a prob^y density f^n, is the analogous for discrete density f^n.

We make the following assumptions, called regularity conditions:

(i) $\frac{d}{d\theta} \log f(x_i;\theta)$; $\theta \in \underline{\theta}$ and for all x exists

(ii) $$\frac{d}{d\theta} \int \cdots \int \prod_{i=1}^{n} f(x_i;\theta)\, dx_1, \ldots, dx_n =$$

$$\int \cdots \int \frac{\partial}{\partial \theta} \prod_{i=1}^{n} f(x_i;\theta)\, dx_1, \ldots, dx_n$$

(iii) $$\frac{d}{d\theta} \int \cdots \int t(x_1, x_2, \ldots, x_n) \prod_{i=1}^{n} f(x_i;\theta)\, dx_1 \cdots dx_n$$

$$= \int \cdots \int t(x_1, x_2, \ldots, x_n) \frac{\partial}{\partial \theta} \prod_{i=1}^{n} f(x_i;\theta)\, dx_1, \ldots dx_n$$

(iv) $0 < E_\theta\left[\frac{\partial}{\partial \theta}\left[\log f(x,\theta)\right]^2\right] < \infty$ for all $\theta \in \underline{\theta}$

Prove that the Cramer Rao inequality :

$$\text{Var}[T] \geq \frac{[\psi'(\theta)]^2}{nE_\theta\left[\left[\frac{\partial}{\partial\theta}\log f(x;\theta)\right]^2\right]}$$

Proof:

If the frequency function of discrete or continuous is $f(x_i;\theta)$, we may define the likelihood function of a sample of n independent samples by

$$L(x_1, x_2, \ldots, x_n/\theta) = f(x_1/\theta) f(x_2/\theta) \cdots f(x_n/\theta)$$

where, L is the joint frequency distribution of

$$\int \cdots \int L\, dx_1, dx_2, \ldots, dx_n = 1 \quad \text{---} (1)$$

Now differentiating both sides w.r.to θ we get,

$$\int \cdots \int \frac{\partial L}{\partial\theta} dx_1, dx_2, \ldots, dx_n = 0$$

which we may rewrite,

$$E\left(\frac{\partial \log L}{\partial \theta}\right) = \int \cdots \int \left(\frac{1}{L} \cdot \frac{\partial L}{\partial \theta}\right) L \, dx_1 dx_2 \cdots dx_n = 0 \quad \text{---(2)}$$

$$\Rightarrow \int \cdots \int \left\{\left(\frac{1}{L} \frac{\partial L}{\partial \theta}\right) \frac{\partial L}{\partial \theta} + L \cdot \frac{\partial}{\partial \theta}\left[\frac{1}{L} \cdot \frac{\partial L}{\partial \theta}\right]\right\} dx_1 \cdot dx_2 \cdots dx_n = 0$$

$$\Rightarrow \int \cdots \int \left\{\left(\frac{1}{L} \cdot \frac{\partial L}{\partial \theta}\right)^2 + \frac{\partial^2 \log L}{\partial \theta^2}\right\} L \, dx_1, dx_2, \ldots, dx_n = 0$$

$$\Rightarrow \int \cdots \int \left(\frac{1}{L} \frac{\partial L}{\partial \theta}\right)^2 L \, dx_1, dx_2, \ldots, dx_n =$$

$$- \int \cdots \int \frac{\partial^2 \log L}{\partial \theta^2} \times L \, \partial x_1, \ldots, \partial x_n$$

from (2)

$$\Rightarrow E\left[\left(\frac{\partial \log L}{\partial \theta}\right)^2\right] = - E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right] \quad \text{---(3)}$$

Now considering an unbiased estimator of t of $\tau(\theta)$, we have

$$E(t) = \int \cdots \int t \, L \, dx_1, \ldots, dx_n = \tau(\theta) \quad \text{---(4)}$$

Taking 1st difference w.r.t to θ,

$$\int \cdots \int t \cdot \frac{\partial L}{\partial \theta} \, dx_1, \ldots, dx_n = \tau(\theta)$$

$$\Rightarrow \int \cdots \int t \left(\frac{1}{L} \cdot \frac{\partial L}{\partial \theta} \right) L \, dx_1, \ldots, dx_n = \tau'(\theta)$$

$$\Rightarrow \int \cdots \int t \, \frac{d \log L}{\partial \theta} \, L \, dx_1 \, dx_2, \ldots dx_n = \tau'(\theta)$$

Now, applying cauchy schwartz inequality,

$$\{\tau'(\theta)\}^2 \leq \int \cdots \int [t - \tau(\theta)]^2 L \, dx_1, \ldots, dx_n \times$$

$$\int \cdots \int \left(\frac{\partial \log L}{\partial \theta} \right)^2 L \, dx_1, \ldots dx_n$$

$$\Rightarrow [\tau'(\theta)]^2 \leq \text{var}(t) \times E \left(\frac{\partial \log L}{\partial \theta} \right)^2$$

$$\Rightarrow \text{var}(t) \geq \frac{[\tau'(\theta)]^2}{E \left(\frac{\partial \log L}{\partial \theta} \right)^2}$$ which is

Cramer - Rao Inequality

Lecture 19 & 20

04-08-15

Chapman Robbins and kiefer inequality for lower bound of variance

Cramer-Rao inequality

$$\text{Var}(t) \geq \frac{\{\tau'(\theta)\}}{E\left[\frac{\partial \log L}{\partial \theta}\right]^2}$$

CRK:

CRK gives a lower bound for the variance of an estimate but does not require regularity conditions of the cramer-rao inequality.

Theorem:

Let $X_1, X_2, \ldots, X_n$ be a random variables drawn from the prob^y density f^n $f(x_i; \theta)$, $\theta \in \Theta$.

Also let $x_1, x_2, \ldots, x_n$ be the random variables with prob^y density f^n

$f(x;\varphi)$, $\varphi \in \Theta$. Also let T be an unbiased estimator of both $\tau(\theta)$ and $\tau(\varphi)$ with $E[T^2] < \infty, \forall\ \theta \in \Theta$

Theorem:

If $\theta \neq \varphi$, assume that $f(x;\theta)$ and $f(x;\varphi)$ are different

Also, if $S(\theta) \supset S(\varphi)$

$$\Rightarrow \left[f(x;\theta) > 0\right] \supset \left[f_\varphi(x;\varphi) > 0\right]$$

Then

$$\text{var}(T) \geq \sup\left\{\varphi : S(\varphi) \subset S(\theta), \varphi = \theta\right\} \frac{[\tau(\varphi) - \tau(\theta)]^2}{\text{var}\left(f(x;\varphi)/f(x;\theta)\right]}$$

Proof:

Since T is an unbiased estimator for $\tau(\theta)$.

So, we may write, $E[T(X)] = \tau(\theta)$; $\forall \phi \in \Theta$

Here, $\phi = \theta$,

$$\int_{S(\theta)}^{T(x)} \frac{f(x, \phi) - f(x; \theta)}{f(x, \theta)} \times f(x; \theta)\, dx$$

$$= \tau(\phi) - \tau(\theta), \quad \text{which yield}$$

$$\text{cov}\left\{T(x), \left[\frac{f(x; \phi)}{f(x; \theta)} - 1\right]\right\} = \tau(\phi) - \tau(\theta)$$

Now, applying Cauchy-Schwartz inequality, we get

$$\text{cov}^2\left\{T(x), \left[\frac{f(x; \phi)}{f(x; \theta)} - 1\right]\right\} \leq \text{var}(T(x)) \times \text{var}\left\{\frac{f(x, \phi)}{f(x, \theta)} - 1\right\}$$

$$\Rightarrow [\tau(\phi) - \tau(\theta)]^2 \leq \text{Var}(T(x)) \times \text{Var}\left\{\frac{f(x_i;\phi)}{f(x;\theta)} - 1\right\}$$

$$\Rightarrow \text{Var}(T(x)) \geq \frac{[\tau(\phi) - \tau(\theta)]^2}{\text{Var}\left\{\frac{f(x;\phi)}{f(x;\theta)} - 1\right\}}$$

⊛ Bhattacharyya Bound:

We can find better lower bound than the minimum variance bound estimator, i.e.,

$$\text{Var}(t_n) \geq \frac{\{\tau'(\theta)\}^2}{E\left[\frac{\partial \log L}{\partial \theta}\right]^2}$$ for the variance of an estimator. In this case MVB is not attainable.

From the condition of MVB we have an estimator t_n for which $t_n - \tau(\theta)$ is a linear function of $\frac{\partial \log L}{\partial \theta} = \frac{1}{L} \cdot \frac{\partial L}{\partial \theta}$

But even if no such estimator exists, still there may be one for which $t_n - \tau(\theta)$ is a linear f^n of $\frac{\partial \log L}{\partial \theta}$,

$$\frac{\partial^2 \log L}{\partial \theta^2}, \ldots; \frac{\partial^n \log L}{\partial \theta^n}$$

Using (B4) we can write such a linear function that mostly clus closer to $t_n - \tau(\theta)$

Let us consider t_n be an estimation of $\tau(\theta)$, we may write

$$L^{(n)} = \frac{\partial^n L}{\partial \theta^n} \quad \text{and} \quad \tau(\theta)^{(n)} = \frac{\partial^{(n)} \tau(\theta)}{\partial \theta^{(n)}}$$

$$D_S = t_n - q(\theta) - \sum a_n \frac{L^n}{L} \quad \text{---(1)}$$

where the constains to be determined

If we differenciati differentiate $\int \cdots \int L \, dx_1 \, dx_2 \cdots dx_n$ n'n n times we get,

$$\int \cdots \int \frac{\partial L^n}{\partial L} L \, dx_1 \, dx_2, \cdots, dx_n = 0$$

$$\Rightarrow E\left[\frac{L^{(n)}}{L}\right] = 0 \quad \text{---(2)}$$

$$E[D_S] = E\left[t_n - q(\theta) - \sum_{n=1}^{\infty} a_n \frac{L^n}{L}\right]$$

$$\therefore E[D_S^2] = \int \cdots \int \left\{ t_n - q(\theta) - \sum a_n \frac{L^n}{L} \right\}^2 \times L \, dx_1 \, dx_2, \ldots, dx_n$$

Now Diff^n p times we get

$$\int \cdots \int \left[t_n - \tau(\theta) - \sum a_n \times \frac{L^{(n)}}{L} \right] \times \frac{L^{(p)}}{L} \times L\, dx_1, dx_2, \ldots dx_n = 0$$

As summation and integration are interchangable, this gives

$$\int \cdots \int \left[t_n - \tau(\theta) \right] \times \frac{L^{(p)}}{L} \, L\, dx_1, dx_2, \ldots, dx_n =$$

$$\sum_{n=0}^{\infty} a_n \cdot \frac{L^{(n)}}{L} \times \frac{L^{(p)}}{L} \times L\, dx_1, dx_2 \ldots dx_n$$

L.H.S

$$= \int \cdots \int \left[t_n - \tau(\theta) \right] \times \frac{L^{(p)}}{L} \times L\, dx_1, dx_2 \ldots dx_n$$

$$\tau^{(p)}(\theta)$$

$$R.H.S. = \sum_{n=0}^{\infty} a_n \int \cdots \int \left[\frac{L^n}{L} \times \frac{L^{(p)}}{L} \right] L\, dx_1, dx_2 \ldots dx_n$$

$$= \sum_{n=1}^{\infty} a_n E\left[\frac{L^n}{L} \times \frac{L^p}{L}\right]$$

$$= \sum_{n=1}^{\infty} a_n \quad \text{cov}\left[\frac{L^n}{L} \times \frac{L^p}{L}\right]$$

$$= \sum_{n=1}^{\infty} a_n J_{np}$$

$$\therefore \quad \gamma^{(p)}(\theta) = \sum_{n=1}^{\infty} a_n \cdot J_{np}$$

$$\Rightarrow a_n = \sum_{p=1}^{\infty} \gamma(\theta)^{(p)} \; J_{np}^{-1}$$

Now setting the value of a_n in eqn. no (1)

$$Ds = f_n - \gamma(\theta) - \sum_{n=1}^{5} \sum_{p=1}^{5} \gamma(\theta)^{p} \; J_{np}^{-1} \; \frac{L^{n'}}{L}\Big] \times$$

$$L \, dx_1, dx_2 \ldots dx_n$$

1, 2, 4 Chapters

Lecture - 21 & 22

6-8-25

1st ~~tutorial~~ chapter

Chapter-1 : Median and Modal unbiased estimates.

Chapter-2: Jacknife and Bootstrap

Chapter-3: Cramer Rao Inequality; Bhattacharyya lower bound

Invariance property:

estimator $\pm, \times, \div$ const

~~Location~~

Location invariant:

As estimator $T = t(X_1, X_2, \ldots, X_n)$ is defined to be location invariant iff

$t(X_1 + c, X_2 + c, \ldots, X_n + c) = t(x_1, x_2, \ldots, x_n) + c$

for all values $x_1, x_2, \ldots, x_n$ and all c

Prove that

(1) $\frac{Y_1 + Y_n}{2}$ is a location invariant.

(2) Prove that S^2 is not a location invariant estimator.

(3) Prove that $\frac{Y_n - Y_1}{2}$ is not a " " estimator

(4) Give a practical example of location invariant estimator.

(5) $\bar{x}$ is a location invariant estimator

* According to the definition,

$$t(x_1 + c, x_2 + c, \ldots, x_n + c) = \bar{x}_n$$

$$= \frac{x_1 + x_2 + \ldots + x_n}{n}$$

$$= \frac{(x_1 + c) + (x_2 + c) + \ldots + (x_n + c)}{n}$$

$$= \frac{x_1 + x_2 + \cdots + x_n}{n} + c$$

$$= \bar{x}_n + c$$

$$= t(x_1, x_2, \ldots, x_n) + c$$

Scale Invariant:

An estimator $T = t(x_1, x_2, \ldots, x_n)$ is defined to be scale invariant iff $t(cx_1, cx_2, \ldots, cx_n) = c\,t(x_1, x_2, \ldots, x_n)$ for all values of $x_1, x_2, \ldots, x_n$ for all values of c

HW Prove that, $\bar{x}_n$, $\sqrt{s^2}$, $\frac{Y_1 + Y_n}{2}$ and $Y_n - Y_1$ are scale invariant.

⊛ Location parameter:

Let $\{f(x; \theta) ; \theta \in \Theta\}$ be a family of densities p with parameter θ. The parameter θ is defined to be a

location parameter if and only if the density $f(x;\theta)$ can be written as a f^n of $x-\theta$ that is $f(x;\theta) = h(x-\theta)$ for some f^n of $h(\cdot)$ equivalently.

Pitman estimator for location:

Let $x_1, x_2, \ldots, x_n$ denote a random sample drawn from the density f^n $f(x;\theta)$, where θ is a location parameter and $\theta \in \underline{\theta}$ then the pitman estimator is written by

$$\hat{\theta}(x_1, x_2, \ldots, x_n) = \frac{\int \theta \prod_{i=1}^{n} f(x;\theta)\, d\theta}{\int \prod_{i=1}^{n} f(x;\theta)\, d\theta} \quad \text{---(1)}$$

The estimator in eq^n (1) is called pitman estimator for location parameter, which has uniformly smallest mean squared error within the class of location-invariant estimators

Examples → chapter - 4 — mod

37, 38, 39

Page (333, 334, 375)

Scale parameter:

Let $\{f(x;\theta), \theta \in \Theta\}$ be a family of densities with parameter θ. The parameter θ is defined to be scale parameter iff the density $f^n f(x;\theta) = \frac{1}{\theta} h(x/\theta)$ for some density f^n & $h(\cdot)$.

Equivalently, θ is called scale parameter of the density $f^n f(x;\theta)$ and the dist^n of $f(x/\theta)$ doest not depend on θ.

Pitman estimator for scale:

Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from the density $f^n f(x;\theta)$; $\theta > 0$ is a scale parameter. Assume that $f(x;\theta) = 0$ for $x \leq 0$ that means the random variable X_i only takes positive values. within the class of scale-invariant estimators, the Pitman estimator for scale is

$$t(X_1, X_2, \ldots, X_n) = \frac{\int_0^\infty \frac{1}{\theta^2} \prod_{i=1}^n f(x;\theta)\, d\theta}{\int_0^\infty \frac{1}{\theta^3} \prod_{i=1}^n f(x;\theta)\, d\theta}$$

which has smallest risk the loss L^n

$$L(t;\theta) = (t-\theta)^2/\theta$$

H.W: Ex (41, 42)

(C37-42)

(*) When we apply Robust Estimator?

In practice, we make many assumptions in parametric inference, and any one or all of this may seldom have accurate knowledge about the underlying dist^{n}. Similarly the assumption of event i.i.d may not hold. Any test or estimator that perform well under modical modification or under assumption is usually referenced as robust estimator.

(*) How can you construct the robust estimator?

Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from $f_0 \sim N(\mu, \sigma^2)$ with 100% and from $f_1 \sim N(\mu, k^2\sigma^2)$

$k>1$ with $(1-\alpha)100\%$ of the time where both N and σ^2 is unknown. We are interested to estimate μ.

We may assume that really the sample is drawn from,

$f(x) = \alpha f_0(x) + (1-\alpha) f_1(x)$; clearly $\bar{x}$ is still unbiased estimator of μ

~~Still unbiased estimator~~

If α is nearly 1 there is no problem, since the underlying dist^n is nearly $N(\mu, \sigma^2)$ and $\bar{x}$ is uniformly minimum variance unbiased estimator of μ with $V(\bar{x}_n) = \frac{\sigma^2}{n}$.

If $(1-\alpha)$ large, then we are sampling from f then the variance of $V(X)$ is σ^2 with prob^y α and $k^2\sigma^2$ with prob^y $(1-\alpha)$ and we have,

$$V(\bar{x}) = \frac{\text{var}(x)}{n}$$

$$= \frac{\alpha\sigma^2}{n} + \frac{(1-\alpha)k\sigma^2}{n}$$

$$= \frac{\sigma^2}{n}[\alpha + (1-\alpha)k]$$ ~~for~~ heterogen

for heterogeneous characteristic

But, if it is homogeneous, $\alpha = 1$

If $\alpha > 1$, then $V(\bar{x})$ ~~will~~ will be more large

If $k(1-\alpha)$ is large, $\text{var}(\bar{x})$, will also be large.

Even for a single wild obs^n makes $\bar{x}$ a considerable errorness.

It is known that sample median is more better estimator than mean in process of extreme values.

If $Z_{1/2} = M$ be the sample median drawn from $X_1, \ldots, X_n$ as an estimate of central value μ, which is known as population median.

Then for large n, it is true that

$$E(m-\mu)^2 = \text{var}(m)$$

$$= \frac{1}{4m} \cdot \frac{1}{4n[f(M)^2]}$$

$$\approx \frac{\pi\sigma^2}{2n}$$

where $f(\mu)$ is the dist^n of pop^n median.

Since

$$f(\mu) = \alpha f_0(\mu) + (1-\alpha) f_1(\mu)$$

$$= \alpha \cdot \frac{1}{\sigma\sqrt{2\pi}} + (1-\alpha) \cdot \frac{1}{\sigma\sqrt{2\pi k}}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \left(\alpha + \frac{1-\alpha}{\sqrt{k}} \right)$$

$$\therefore E(m-\mu)^2 = \frac{1}{4n\left\{\left[\alpha + \frac{1-\alpha}{\sqrt{k}}\right] \frac{1}{\sigma\sqrt{2\pi}}\right\}^2}$$

$$\Rightarrow \text{var}(m) = \frac{\pi\sigma^2}{2n} \cdot \frac{1}{\left(\alpha + \frac{1-\alpha}{\sqrt{k}}\right)^2} \quad \text{②}$$

If $k \to \alpha$, then the eqn ② becomes

$$\text{var}(m) = \frac{\pi\sigma^2}{2n} \times \frac{1}{\alpha^2}$$

If $\alpha = 1$, $\text{var}(m) = \frac{\pi\sigma^2}{2n}$ that means all the observations are taking from second pdf i.e., homogeneity will be

high i.e., var(m) is not affected by k

This indicate that the estimator m will not be greatly affected by how large k is.

$$\therefore \frac{\text{var}(\bar{x})}{\text{var}(m)} = \frac{\frac{\sigma^2}{n}[\alpha + (1-\alpha)k]}{\frac{\pi\sigma^2}{2n}\left[\frac{1}{\alpha + \frac{1-\alpha}{\sqrt{k}}}\right]^2}$$

$$= \frac{\sigma^2}{n}[\alpha + (1-\alpha)k] \times \frac{2n\left\{\alpha + \frac{1-\alpha}{\sqrt{k}}\right\}^2}{\pi\sigma^2}$$

$$= \frac{2}{\pi}\left[\alpha + (1-\alpha)k\right]\left[\alpha + \frac{1-\alpha}{\sqrt{k}}\right]^2$$

If $k \to \infty$

$$\therefore \frac{\text{var}(\bar{x})}{\text{var}(m)} = \alpha$$

Whereas $\text{var}(\bar{x}) \to \infty$ as $k \to \infty$

$$\therefore \text{var}(m) = \frac{\pi\sigma^2}{2n\alpha^2}$$

So we may write

$$\therefore \text{var}(m) < \text{var}(\bar{x})$$

So, the estimator m is called a robust estimator