Statistical Inference (1)

* **Example:-**

Given $S_1^2 = \frac{1}{n}\Sigma(X_i-\bar{X})^2$ and $S_2^2 = \frac{1}{n-1}\Sigma(X_i-\bar{X})^2$ are two estimators for variance of Normal distribution $N(\theta,\sigma^2)$ (θ = known). Which of these two estimators is a better estimate using the MSe ?

Solu:

$$ES_1^2 = \frac{n-1}{n}\sigma^2$$

$\therefore S_1^2$ is a biased estimator for σ^2.

$$ES_2^2 = E\frac{1}{n-1}\Sigma(X_i-\bar{X})^2$$

$$= E\frac{1}{n-1}nS_1^2 \qquad (\Sigma(X_i-\bar{X})^2 = nS_1^2)$$

$$= \frac{n}{n-1}ES_1^2 = \frac{n}{n-1}\,\frac{n-1}{n}\sigma^2 = \sigma^2$$

$\therefore S_2^2$ is unbiased estimator of σ^2

$$var(S_1^2) = Var\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right] = \frac{1}{n^2}Var\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]$$

$$\because \sum_{i=1}^{n}(X_i-\bar{X})^2 = \sum_{i=1}^{n}\left[(X_i-\mu)-(\bar{X}-\mu)\right]^2$$

$$= \sum_{i=1}^{n}(X_i-\mu)^2 - n(\bar{X}-\mu)^2$$

$$\therefore \sum_{i=1}^{n}(X_i-\mu)^2 = \Sigma(X_i-\bar{X})^2 + n(\bar{X}-\mu)^2$$

$$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2}$$ بقسمة الطرفين على σ^2

$$\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} + \left(\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

$$\chi^2_{(n)} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} + \chi^2_{(1)}$$

$$\therefore \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} = \chi^2_{(n)} - \chi^2_{(1)}$$

$$\Rightarrow \frac{n S_1^2}{\sigma^2} = \chi^2_{(n-1)}$$

ملاحظة

if $X \sim N(\mu, \sigma^2)$

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi^2_{(1)}$$

$$\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2_{(n)}$$

وبأخذ التباين للطرفين

$$\frac{n^2}{\sigma^4} Var(S_1^2) = 2(n-1)$$

$$\therefore Var(S_1^2) = \frac{2(n-1)}{n^2}\sigma^4$$

$$\because S_2^2 = \frac{n}{n-1} S_1^2$$

$$\Rightarrow Var(S_2^2) = Var\left(\frac{n}{n-1} S_1^2\right) = \frac{n^2}{(n-1)^2} Var(S_1^2)$$

$$= \frac{n^2}{(n-1)^2} \cdot \frac{2(n-1)}{n^2}\sigma^4$$

$$\Rightarrow Var(S_2^2) = \boxed{\frac{2}{n-1}\sigma^4} = MSe(S_2^2)$$

$$\therefore MSe(S_1^2) = Var(S_1^2) + b^2(S_1^2)$$

$$b^2(S_1^2) = \left[E(S_1^2) - \sigma^2\right]^2$$

$$= \left[\frac{n-1}{n}\sigma^2 - \sigma^2\right]^2 = \sigma^4\left[\frac{n-1-n}{n}\right]^2 = \frac{\sigma^4}{n^2}$$

$$\therefore MSe(S_1^2) = \frac{2(n-1)}{n^2}\sigma^4 + \frac{\sigma^4}{n^2}$$

$$\therefore MSe(S_1^2) = \frac{\sigma^4}{n^2}\left[2(n-1)+1\right] = \frac{\sigma^4}{n^2}(2n-2+1)$$

$$= \boxed{\frac{\sigma^4}{n^2}(2n-1)}$$

لمعرفة اي متوسط مربعات الخطأ MSe اكبر لـ S_1^2 أم لـ S_2^2 نفرض فرضية

Assume $MSe(S_1^2) > MSe(S_2^2)$

$$\frac{\sigma^4}{n^2}(2n-1) > \frac{2}{n-1}\sigma^4$$

بقسمة الطرفين على σ^4

$$\frac{2n-1}{n^2} > \frac{2}{n-1}$$

$$(2n-1)(n-1) > 2n^2$$

$$2n^2 - 2n - n + 1 > 2n^2$$

$$2n^2 - 3n + 1 > 2n^2$$

$$2n^2 - 2n^2 + 1 > 3n$$

$$1 > 3n$$

وهذا لا يجوز اي أن الفرضية خاطئة وعليه فان

$$MSe(S_1^2) < MSe(S_2^2)$$

$\therefore S_1^2$ is better than S_2^2

* Example:

Let $X_1, X_2, \dots, X_n$ be a r.s. of size (n) from $N(\theta, \sigma^2)$. Use MSe to compare between three estimators for θ.

$$T_1 = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i \quad , T_2 = \frac{1}{2}(\bar{X} - X_1) \; , T_3 = \frac{1}{3}(\bar{X} + X_1 + X_2)$$

solu:

$$E(T_1) = E\left[\frac{1}{n-1}\sum_{i=1}^{n-1} X_i\right] = \frac{1}{n-1}\sum_{i=1}^{n-1} E X_i = \frac{n-1}{n-1}\theta = \theta \text{ (unbiased)}$$

$$E(T_2) = E\left[\frac{1}{2}(\bar{X} + X_1)\right] = \frac{1}{2}\left[E(\bar{X}) + E(X_1)\right] = \frac{1}{2}[\theta + \theta]$$

$$= \frac{1}{2} 2\theta = \theta \text{ (unbiased)}$$

$$E(T_3) = E\left[\frac{1}{3}(\bar{X} + X_1 + X_2)\right] = \frac{1}{3}\left[E(\bar{X}) + E(X_1) + E(X_2)\right]$$

$$= \frac{1}{3}[\theta + \theta + \theta] = \frac{3\theta}{3} = \theta \text{ (unbiased)}$$

$$\therefore MSe(T_1) = Var(T_1) = var\left[\frac{1}{n-1}\sum_{i=1}^{n-1} X_i\right] = \frac{1}{(n-1)^2}\sum_{i=1}^{n-1} var(X_i)$$

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

$$MSe(T_2) = var(T_2) = Var\left[\frac{1}{2}(\bar{X} + X_1)\right] = \frac{1}{4} Var[\bar{X} + X_1]$$

هنا X_1 و $\bar{X}$ غير مستقلين لأن $\bar{X}$ تحتوي على X_1 لذلك نقوم بالتجزئة لتحويل التباين الى تباين مقادير مستقلة حتى يمثل حاصل جمع التباينات .

$$= \frac{1}{4} var\left[\frac{1}{n}\sum_{i=2}^{n} X_i + \frac{X_1}{n} + X_1\right]$$

$$= \frac{1}{4} Var\left[\frac{1}{n}\sum_{i=2}^{n} X_i + \frac{n+1}{n} X_1\right]$$

$$= \frac{1}{4}\left[\frac{1}{n^2}\sum_{i=2}^{n} var(X_i) + \frac{(n+1)^2}{n^2} var(X_1)\right]$$

$$= \frac{1}{4}\left[\frac{n-1}{n^2}\sigma^2 + \frac{(n+1)^2}{n^2}\sigma^2\right]$$

$$= \frac{n+3}{4n}\sigma^2$$

$$MSe(T_3) = Var(T_3) = Var\left[\frac{1}{3}(\bar{X} + X_1 + X_2)\right]$$

$$= \frac{1}{9} var(\bar{X} + X_1 + X_2)$$

$$= \frac{1}{9} Var\left[\frac{1}{n}\sum_{i=3}^{n} X_i + \frac{X_1}{n} + \frac{X_2}{n} + X_1 + X_2\right]$$

$$= \frac{1}{9} var\left[\frac{1}{n}\sum_{i=3}^{n} X_i + \frac{n+1}{n} X_1 + \frac{n+1}{n} X_2\right]$$

$$= \frac{1}{9}\left[\frac{1}{n^2}\sum_{i=3}^{n} Var(X_i) + \frac{(n+1)^2}{n^2} var(X_1) + \frac{(n+1)^2}{n^2} Var(X_2)\right]$$

$$= \frac{1}{9}\left[\frac{(n-2)}{n^2} a^2 + \frac{(n+1)^2}{n^2} a^2 + \frac{(n+1)^2}{n^2} a^2\right]$$

$$= \frac{2n+5}{9n} a^2$$

$\therefore T_1$ is better estimator.