Output Analysis - Part II (Criteria of Good Estimators)

W. M. Song 桑慧敏
Tsing Hua Univ. 清華大學

2015.11.25

Notations

Recall the midterm-exam problem: Identify a common property for each part.

Notations

- 1. $E(X)$ and $\overline{X}$
- 2. Var(X) and S^2
- 3. $E(X)$ and Var(X)
- 4. $\overline{X}$ and S^2
- 5. $\overline{x}$ and s^2
- 6. σ_X and σ_Y
- 7. E($X > 10$) and $\mathrm{Var}(\mathrm{X} < 10)$
- 8. $\{X > 10\}$ and $\{X < 10\}$
- 9. $\mathrm{P}(\mathrm{X} > 10)$ and $\mathrm{P}(\overline{\mathrm{X}} < 10)$
- 10. $E(X)$, Var(X), $\overline{x}$, s^2

Common Property

- 1. About centering
- 2. About variability
- 3. Population Property
- 4. Sample Property, or R.V.
- 5. Sample Property, or constants
- 6. Population Prop., or constants
- 7. Meaningless
- 8. Events
- 9. Constants in [0,1]
- 10. Constants

Notations

隨機樣本 (樣本數為 n)	估計量 $\hat{\Theta}$	統計量
$X_1, X_2, \ldots, X_n$	$\overline{X}_{(n)}$	$(\overline{X}_{(n)} - \mu)/\mathrm{se}(\overline{X})$
	S_X^2	$(n-1)S^2/\sigma^2$

抽樣前

母群體

f_X

抽樣後

觀察到的樣本	估計值 $\hat{\theta}$
$x_1, x_2, \ldots, x_n$	$\overline{x}_{(n)}$
	s_X^2

- Distinguish before and after sampling
- Distinguish capital X and lower case x

Notations

r.v.	pdf or pmf	分配的屬性 平均數	變異數	標準差或標準誤
X	f_X	E(X) 或 μ_X	V(X) 或 σ^2_X	σ_X 或 sd(X)
$\overline{X}$	$f_{\overline{X}}$	E$(\overline{X})$ 或 $\mu_{\overline{X}}$	V$(\overline{X})$ 或 $\sigma^2_{\overline{X}}$	$\sigma_{\overline{X}}$ 或 se$(\overline{X})$
S^2	f_{S^2}	E(S^2) 或 μ_{S^2}	V(S^2) 或 $\sigma^2_{S^2}$	σ_{S^2} 或 se(S^2)
$\hat{\Theta}$	$f_{\hat{\Theta}}$	E$(\hat{\Theta})$ 或 $\mu_{\hat{\Theta}}$	V$(\hat{\Theta})$ 或 $\sigma^2_{\hat{\Theta}}$	$\sigma_{\hat{\Theta}}$ 或 se$(\hat{\Theta})$

- Goal: To estimate θ
- Estimator of θ: $\hat{\Theta}$
- Distinguish θ, $\hat{\Theta}$, and $\hat{\theta}$

Criteria of Good Estimators

- Goal: To estimate θ

Estimates

- $\hat{\theta}_1$ and $\hat{\theta}_2$
- Q: Which estimate is better?
- A: No way to know

Estimators

- $\hat{\Theta}_1$ and $\hat{\Theta}_2$
- Q: Which estimator is better?
- A: Bias, Var, and MSE

- Bias, Var, and MSE of $\hat{\Theta}_1$ and $\hat{\Theta}_2$

Bias ($\hat{\Theta}$) and Var($\hat{\Theta}$)

- Goal: estimate θ
- estimator: $\hat{\Theta}$, $\hat{\Theta}_1$, $\hat{\Theta}_2$

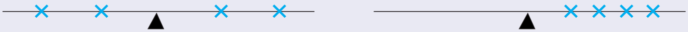
Bias (偏度)

Bias($\hat{\Theta}$): E($\hat{\Theta}$) $-\ \theta$

Variance(變異)

Var($\hat{\Theta}$) $=$ E($\hat{\Theta} - \mu_{\hat{\Theta}})^2$

Which estimator is a better?



(a) $\hat{\Theta}_1$ (b) $\hat{\Theta}_2$

- Q: Which estimator is better?
- A: In terms of Bias, $\mathrm{Bias}(\hat{\Theta}_1) < \mathrm{Bias}(\hat{\Theta}_2)$
- In terms of variance, $\mathrm{Var}(\hat{\Theta}_1) > \mathrm{Var}(\hat{\Theta}_2)$
- So, we need an “integrated criterion” to include Bias($\hat{\Theta}$) and Var($\hat{\Theta}$)

Criteria of Good Estimators

- Goal: estimating θ; estimator: $\hat{\Theta}$
- Estimator: $\hat{\Theta}$

Bias (偏度)

Bias($\hat{\Theta}$): E$(\hat{\Theta}) - \theta$

Variance(變異)

$\text{Var}(\hat{\Theta}) = \text{E}(\hat{\Theta} - \mu_{\hat{\Theta}})^2$

Mean Squared Error (MSE)(均方誤)

$\text{MSE}(\hat{\Theta}) = \text{E}(\hat{\Theta} - \theta)^2 = [\text{Bias}(\hat{\Theta})]^2 + \text{Var}(\hat{\Theta})$

- Distinguish $\text{E}(\hat{\Theta} - \theta)^2$ and $\text{E}(\hat{\Theta} - \mu_{\hat{\Theta}})^2$

Comparing X_1, $\overline{X}_{(2)}$, and $\overline{X}_{(100)}$

Goal: to estimate $\theta = \mu_X$

- $\hat{\Theta}_1 = X_1$
- $\hat{\Theta}_2 = \overline{X}_{(2)}$
- $\hat{\Theta}_3 = \overline{X}_{(100)}$

Performance

- $\mathrm{Bias}(X_1) = \mathrm{Bias}(\overline{X}_{(2)}) = \mathrm{Bias}(\overline{X}_{(100)}) = 0$
- $\mathrm{Var}(X_1) > \mathrm{Var}(\overline{X}_{(2)}) > \mathrm{Var}(\overline{X}_{(100)})$
- $\mathrm{Bias}(\overline{X}_{(n)}) = \mu_X$
- $\mathrm{Var}(\overline{X}_{(n)}) = \mathrm{Var}(X)/n$

- Which estimator (among $\hat{\Theta}_i, i = 1, 2, 3$) is the best?

Comparing S^2, $\frac{n-1}{n}S^2$, and $\frac{n-1}{n+1}S^2$

Goal: estimating $\theta = \sigma_X^2$

- $\hat{\Theta}_1 = \sum_{i=1}^{n}(X_i - \overline{X})^2/(n-1) = S^2$
- $\hat{\Theta}_2 = \sum_{i=1}^{n}(X_i - \overline{X})^2/n = \frac{n-1}{n}S^2$
- $\hat{\Theta}_3 = \sum_{i=1}^{n}(X_i - \overline{X})^2/(n+1) = \frac{n-1}{n+1}S^2$

Performance

- $\text{Bias}(\hat{\Theta}_1) = 0$
- $\text{Mse}(\hat{\Theta}_3) < \text{Mse}(\hat{\Theta}_2) < \text{Mse}(\hat{\Theta}_1)$
- $\text{mse}(\hat{\Theta}_1) = \frac{2\sigma^4}{n-1}$; $\text{mse}(\hat{\Theta}_2) = \frac{(2n-1)\sigma^4}{n^2}$; $\text{mse}(\hat{\Theta}_3) = \frac{2\sigma^4}{n+1}$

- Which estimator (among $\hat{\Theta}_i$, $i = 1, 2, 3$) is the best?

Why use CRN?

- Motivation: Estimating $\theta_1 - \theta_2$ (Recall the MM2 systems with 1 and 2 queues)
- Crude Estimator: $\hat{\Theta} \equiv \hat{\Theta}_1 - \hat{\Theta}_2$ with arbitrary random numbers
- CRN-estimator: Apply common random numbers (CRN) for

$$\hat{\Theta}_1 \text{ and } \hat{\Theta}_2 \text{ so that } \mathrm{Cov}(\hat{\Theta}_1, \hat{\Theta}_2) > 0$$

$$\text{Result 1}: \mathrm{E}(\hat{\Theta}) = \mathrm{E}(\hat{\Theta}_1) - \mathrm{E}(\hat{\Theta}_2) = \theta_1 - \theta_2, \quad \text{if } \Theta_1 \text{ and } \Theta_2 \text{ are unbiased}$$

$$\text{Result 2}: \mathrm{Var}(\hat{\Theta}) = \mathrm{Var}(\hat{\Theta}_1) + \mathrm{Var}(\hat{\Theta}_2) - 2\mathrm{Cov}(\hat{\Theta}_1, \hat{\Theta}_2) < \mathrm{Var}(\hat{\Theta}_1) + \mathrm{Var}(\hat{\Theta}_2) \text{ if } \mathrm{Cov}(\hat{\Theta}_1, \hat{\Theta}_2) > 0$$

- We expect CRN-estimator performs better than Crude Estimator

CRN Used as a Block Variable in DOE

- We expect to use "Random numbers" as a block variable, used in design of experiment (DOE)
- HW: DOE and block variable. One-page report