2022 SUMMER

업데이터 통계학 스터디

Chapter ~~5~~ 6 – Introduction to Statistical Inference

Introduction

Parametric model ||

- 분포가 알려짐
- parameter 개수가 유한하게 고정

Estimation. { Point (데이터가 평점들의 평균 가)

Interval (〃 범위 20%.)

95% CI

{ estimate : 추정.

estimator : random variables. (ex. $\bar{Y}$)

Introduction

In many cases, populations are characterized by numerical descriptive

measure called parameters.

- Ω: parameter space = all possible values of θ
- θ : unknown parameter in a model
- $Y_1, \cdots, Y_n$: data $\Rightarrow$ $= d(Y_1, \cdots, Y_n)$: estimator of θ $\rightsquigarrow$ θ x

Many different estimators may be obtained for the same parameter.

How can we establish criteria of goodness to compare statistical estimators? $\rightsquigarrow$ MSE

Bias

Bias

- Point estimation is like firing a gunshot at a target.
- Drawing a single sample and compute an estimate for a parameter, firing a single shot at the target.

- We cannot evaluate the “goodness” of shooter based on only one shot. We must observe the results many times under the same setting. ⇒ We look at the frequency distribution of the values of estimates in repeated sampling.

Bias

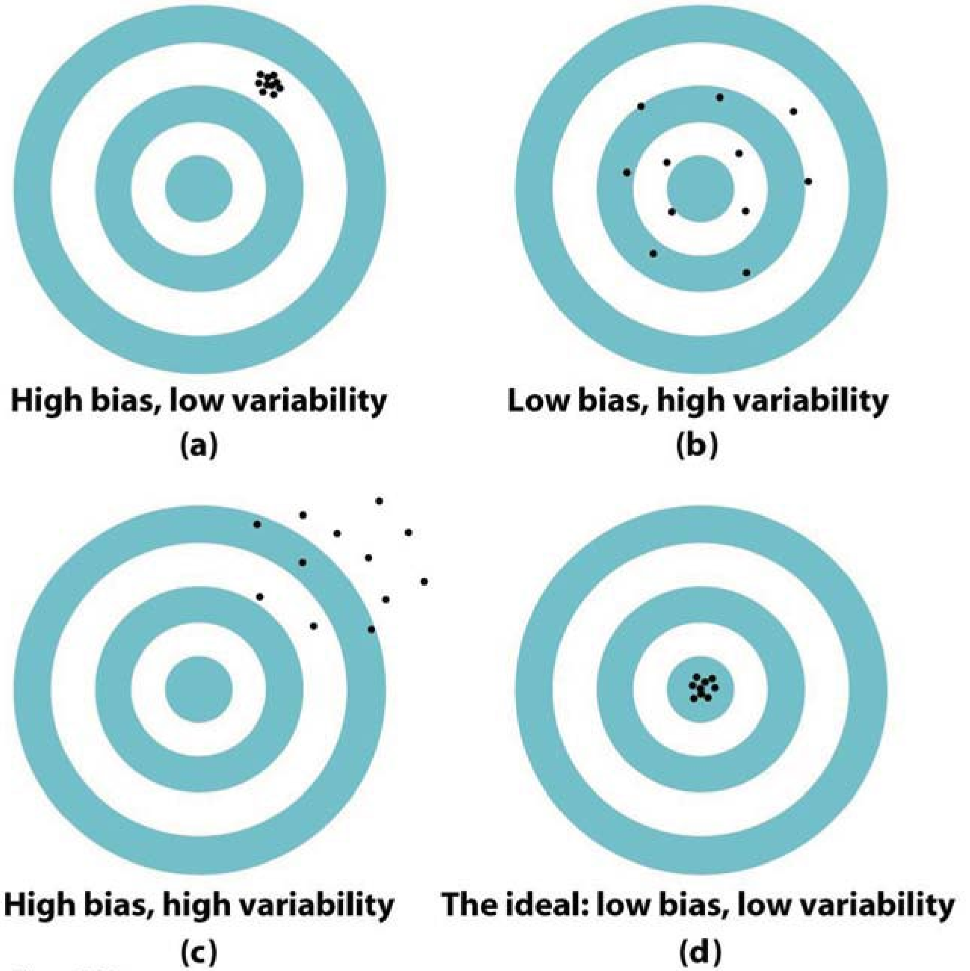


Figure 3-12
Introduction to the Practice of Statistics, Fifth Edition

Bias

θ: target parameter, $\hat{\theta}$: estimator

Definition

Bias of $\hat{\theta}$ as an estimator of θ is

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta$$

$\hat{\theta}$ is an unbiased estimator of θ if $E(\hat{\theta}) = \theta$ for all $\theta \in \Omega$.
Otherwise, biased.

비편향 추정량

Example 6.1

$Y_1, \cdots, Y_n$: *iid* with mean μ and variance σ^2.
Evaluate the following two estimator.

1) An estimator for μ : $\overline{Y}$

2) An estimator for σ^2: S^2

$\overline{Y} = \boxed{\frac{1}{n}} \sum_{i=1}^{n} Y_i$

$S^2 = \boxed{\frac{1}{n-1}} \sum_{i=1}^{n} (Y_i - \overline{Y})^2$

$E(\overline{Y}) = E\left(\frac{1}{n}(Y_1 + \cdots + Y_n)\right)$

$= \frac{1}{n} \sum_{i=1}^{n} \underbrace{E(Y)}_{\mu}$

$= \frac{1}{n} \cdot n \cdot \mu \quad = \mu$ UE

$B(\overline{Y}) = E(\overline{Y}) - \mu$

$= 0.$

$$E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} Y_i^2 - 2\bar{Y}\underbrace{\sum_{i=1}^{n} Y_i}_{=n\bar{Y}} + n\bar{Y}^2\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} Y_i^2\right) - \frac{n}{n-1} E(\bar{Y}^2)$$

$$= \frac{n}{n-1} E(Y^2) - \frac{n}{n-1} E(\bar{Y}^2)$$

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

if $s^2 = \frac{1}{n}\sum(Y_i-\bar{Y})^2$

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

Under estimate

$$E(Y^2) = V(Y) + \{E(Y)\}^2$$
$$= \sigma^2 + \mu^2.$$

$$E(\bar{Y}^2) = V(\bar{Y}) + \{E(\bar{Y})\}^2$$
$$= \frac{\sigma^2}{n} + \mu^2.$$

S^2 is an UE of σ^2.

Example 6.2

$S = \sqrt{S^2}$ is not an UE of σ.

Prove it.

$Var(Y) \geq 0$

Y 가 상수 X $\rightarrow$ $Var(Y) > 0$

$E(Y^2) - \{E(Y)\}^2 > 0$

$E(Y^2) > \{E(Y)\}^2$

$E(S^2) > \{E(S)\}^2$ (Y가 상수 X)

$\sigma^2 > \{E(S)\}^2 \rightarrow \underline{\sigma > E(S)}$ //

$E(S) \neq \sigma$.

MSE

MSE

Square distance : $\sqrt{(\hat{\theta}-\theta)^2}$

Is unbiasedness enough?

Definition

MSE = Mean Square Error

Estimator for θ. target parameter.

$$MSE(\hat{\theta}) = \underline{E_{\theta}\left\{\left(\hat{\theta} - \theta\right)^2\right\}} = V(\hat{\theta}) + \{B(\hat{\theta})\}^2$$

We want to use an estimator that has a small MSE.

$$MSE(\hat{\theta}) = E(\underset{\text{r.v.}}{\hat{\theta}} - \underset{c}{\theta})^2 = E(\hat{\theta}^2 - 2\theta\hat{\theta} + \theta^2)$$

$$= E(\hat{\theta}^2) \underbrace{- \{E(\hat{\theta})\}^2 + \{E(\hat{\theta})\}^2} - 2\theta E(\hat{\theta}) + \theta^2$$

$$= Var(\hat{\theta}) + \{\underbrace{E(\hat{\theta}) - \theta}_{B(\hat{\theta})}\}^2 = Var(\hat{\theta}) + \{B(\hat{\theta})\}^2.$$

Example 6.3

$E(Y)=\mu$
$V(Y)=\sigma^2$ $\rightarrow$ $E(\bar{Y})=\mu$. $V(\bar{Y})=\frac{\sigma^2}{n}$.

$Y_1, \cdots, Y_n$: *iid Unif*$(0, \theta)$, $\Omega = \{\theta : 0 < \theta < \infty\}$

$$\hat{\theta}_1 = 2\bar{Y}, \quad \hat{\theta}_2 = \frac{n+1}{n} max(Y_i)$$

Which estimator is the better estimator?

$Y_i \overset{iid}{\sim} Unif(0, \theta)$ $\left\{ \begin{array}{l} \text{mean}: \frac{\theta}{2} \\ Var: \frac{\theta^2}{12} \end{array} \right.$ $E(\bar{Y}) = \frac{\theta}{2}$

(i) $\hat{\theta}_1 = 2\bar{Y}$

$E(\hat{\theta}_1) = \theta$ $\therefore$ $\hat{\theta}_1$ is an UE of θ

$MSE(\hat{\theta}_1) = Var(\hat{\theta}_1) + \{\cancel{B(\hat{\theta}_1)}\}^2$ 0

$= Var(2\bar{Y})$

$= 4 \cdot \frac{\theta^2}{12n} = \frac{\theta^2}{3n}$

(ii) $\hat{\theta}_2 = \frac{n+1}{n}\underline{\max Y_i}$

Let $W := \max Y_i \longrightarrow f_W(w) = \frac{n w^{n-1}}{\theta^n} \quad (0 < w < \theta)$

$$E(\hat{\theta}_2) = E\left(\frac{n+1}{n}\cdot W\right) = \frac{n+1}{n}E(W)$$

$$= \frac{n+1}{n}\cdot\frac{n}{n+1}\cdot\theta = \theta.$$

$\therefore \hat{\theta}_2$ is an UE of θ.

$$E(W) = \int_0^\theta \frac{n w^n}{\theta^n}\,dw = \frac{n}{n+1}\cdot\frac{w^{n+1}}{\theta^n}\Big|_0^\theta = \frac{n\theta}{n+1}$$

$$E(W^2) = \int_0^\theta \frac{n w^{n+1}}{\theta^n}\,dw = \frac{n}{n+2}\cdot\frac{w^{n+2}}{\theta^n}\Big|_0^\theta = \frac{n}{n+2}\cdot\theta^2$$

$$MSE(\hat{\theta}_2) = Var(\hat{\theta}_2) + \{\cancel{B(\hat{\theta}_2)}\}^2$$

$$= Var\left(\frac{n+1}{n}W\right) = \frac{(n+1)^2}{n^2}\left\{\overset{\frac{n\theta^2}{n+2}}{\overset{\shortparallel}{E(W^2)}} - \overset{\frac{n^2\theta^2}{(n+1)^2}}{\overset{\shortparallel}{\{E(W)\}^2}}\right\}$$

$$\rightarrow \frac{(n+1)^2}{n^2} \left\{ \frac{n\theta^2}{(n+2)(n+1)^2} \right\}$$

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

$$\rightsquigarrow \mathrm{MSE}(\hat{\theta}_1) = \frac{\theta^2}{3n} \qquad \mathrm{MSE}(\hat{\theta}_2) = \frac{\theta^2}{n(n+2)}$$

$$\frac{\mathrm{MSE}(\hat{\theta}_1)}{\mathrm{MSE}(\hat{\theta}_2)} = \frac{n+2}{3} \qquad \begin{cases} \text{if } n=1 \rightarrow \text{같음} \\ \text{else} \rightarrow \mathrm{MSE}(\hat{\theta}_1) > \mathrm{MSE}(\hat{\theta}_2) \end{cases}$$

$\hat{\theta}_2$가 더 좋은 estimator

Confidence Intervals

Confidence Intervals

Definition

A level $1-\alpha$ confidence interval (CI) for parameter θ is an interval $[\hat{\theta}_L, \hat{\theta}_U]$, where $\hat{\theta}_L$, $\hat{\theta}_U$ are found from data s.t.

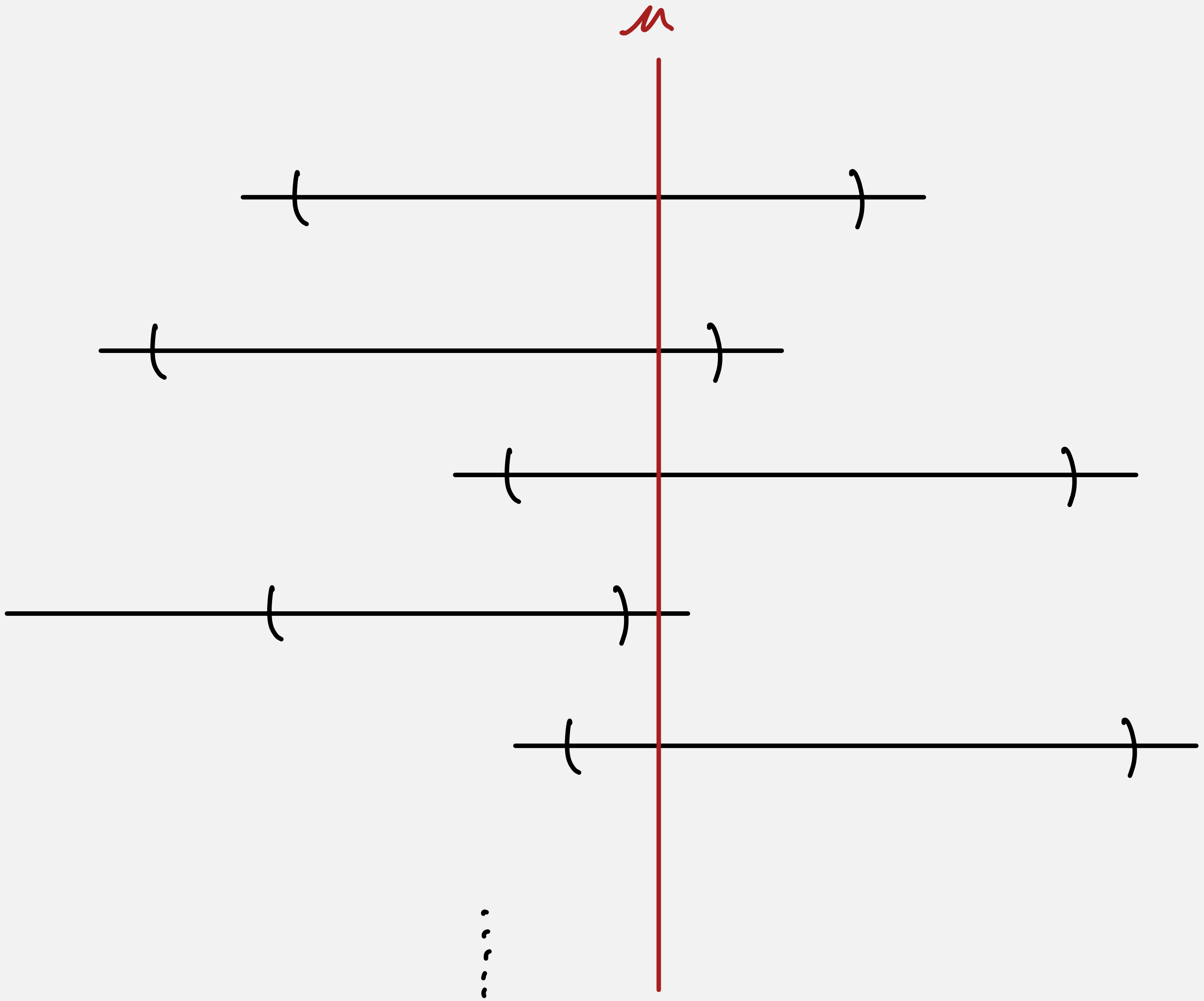
$$P_\theta(\hat{\theta}_L \le \theta \le \hat{\theta}_U) = 1 - \underbrace{\alpha}_{\text{confidence level}}$$

작은 값

0.05

0.01

$$\mu = \bar{x} \pm z\frac{\sigma}{\sqrt{n}}$$



CI for μ using $\bar{Y}$ 95%

$$P\left(\mu \in \bar{Y} \pm 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

data can vary.

Confidence Intervals

추정하려고 하는 target.

$$-z_{.95} \le \boxed{\frac{\bar{Y}-\mu}{\sigma}} \le z_{.95}$$

Pivotal Method

(pivot)

A pivotal quantity is a function of the sample measurements and unknown parameter θ, (θ is the only unknown parameter) (1) and its probability distribution does not depend on θ.

After finding a pivotal quantity and its distribution, only some algebra is needed to get CI.

Confidence Intervals

Pivotal Method

$g(y_1, \cdots, y_n; \theta)$: pivot ⇒ Find c_1 and c_2 s.t.

$$P_\theta(c_1 \le \boxed{g(y_1, \cdots, y_n; \theta)} \le c_2) = 1 - \alpha$$

0.95

$$-1.96 \le \frac{\bar{Y} - \mu}{\sigma} \le 1.96$$

⇒ Restate in the form of

$$P_\theta(\hat{\theta}_L \le \theta \le \hat{\theta}_U) = 1 - \alpha$$

$\bar{Y} - 1.96\frac{\sigma}{\sqrt{n}}$ μ $\bar{Y} + 1.96\frac{\sigma}{\sqrt{n}}$

Example 6.3

$\boxed{Y \sim Exp(\theta)}$. Use Y to form a 90% CI for θ.

$$f_Y(y) = \frac{1}{\theta} e^{-\frac{y}{\theta}} \quad (y>0)$$

$$M_Y(t) = (1-\theta t)^{-1}.$$

$$= E(e^{tY})$$

$$U := \frac{Y}{\theta}$$

sample mesurment.

몰라요

Pivot.

$$M_U(t) = E(e^{tU})$$

$$= E\left(e^{\frac{t}{\theta}Y}\right)$$

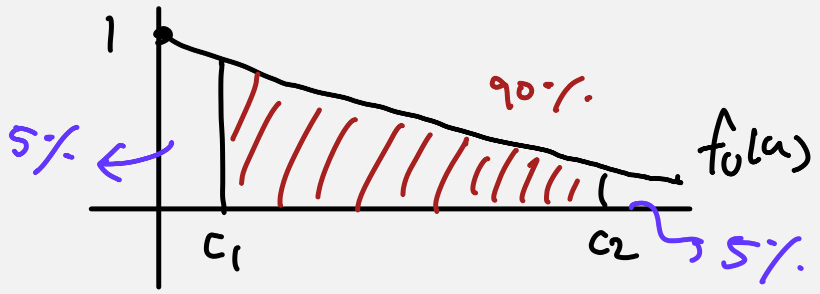
$$= (1-t)^{-1}.$$

MGF of $Exp(1)$.

$$\therefore U \sim Exp(1)$$

$$f_U(u) = e^{-u} \quad (u>0)$$

$$P(c_1 \le U \le c_2) = 0.9$$



$$\int_0^{c_1} e^{-u}\,du = 1 - e^{-c_1} = 0.05$$

$$e^{-c_1} = 0.95$$

$$\int_{c_2}^{\infty} e^{-u}\,du = e^{-c_2} = 0.05$$

$$c_1 = 0.051, \quad c_2 = 2.996.$$

$$P\left(0.051 \le \frac{Y}{\theta} \le 2.996\right)$$

$$= P\left(\frac{Y}{2.996} \le \theta \le \frac{Y}{0.051}\right)$$

$\therefore$ 90% CI of θ

$$\left(\frac{Y}{2.996}, \frac{Y}{0.051}\right)$$

Fundamental Sampling Theorem

<u>Theorem</u>

$Y_1, \cdots, Y_n$: $iid\ N(\mu, \sigma^2)$. Then,

1. $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$
2. $S^2 = \frac{1}{n-1}\sum(Y_i - \bar{Y})^2 \perp\!\!\!\perp \bar{Y}$
3. $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$

$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$

z

$\frac{(n-1)s^2}{\sigma^2} = \Sigma\left(\frac{Y_i - \bar{Y}}{\sigma}\right)^2$.

$\sim \chi^2(n-1)$.

Example 6.4

$Y_1, \cdots, Y_n$: $iid\ Unif(0, \theta)$, a level $1-\alpha$ CI for θ ?

Example 6.5

$Y_1, \cdots, Y_n$: $iid\ N(\mu, \sigma^2)$, a level $1-\alpha$ CI for μ ?

(σ: unknown parameter)

Example 6.6

$Y_1, \cdots, Y_n$: $iid\ N(\mu, \sigma^2)$, a level $1-\alpha$ CI for σ^2?

(μ, σ: unknown parameter)