Some population

described by some distribution $\longleftrightarrow$ pdf $f(x)$

Sample of n independent, identically distributed (i.i.d)

variables $X_1, \ldots, X_n$ (all given by $f(x)$)

e.g. joint distribution: $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i)$

$f(x)$ really $f(x, \theta)$ θ unknown parameter

goal: find θ

<u>Basic method</u>: design new variables $\hat{\Theta}$ supposed to estimate θ.

$$\hat{\Theta} = g(X_1, \ldots, X_n)$$

<u>Examples</u>: $\theta = \mu$, $\hat{\Theta} = \overline{X}$

$\theta = \sigma^2$, $\hat{\Theta} = S^2$

such random variable: "estimator"

point estimator.

Def An estimator $\hat{\Theta}$ for a parameter θ is **unbiased** if $E[\hat{\Theta}] = \theta$ for all possible values of θ.

Ex: $\bar{X}$ is an unbiased estimator for μ.

Pf: $$E[\bar{X}] = E\left[\frac{\sum X_i}{n}\right] = \frac{\sum E[X_i]}{n}$$

$$= \frac{\sum_{i=1}^{n} \mu}{n} = \frac{n\mu}{n} = \mu$$

Ex: S^2 as a pt estimator of σ^2

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

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

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

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - 2\bar{X} \underbrace{\sum_{i=1}^{n} X_i}_{n\bar{X}} + \sum_{i=1}^{n} \bar{X}^2\right]$$

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right]$$

$$= \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \right]$$

$$E[X^2] = \sigma_X^2 + \mu_X^2$$

$$\frac{1}{n-1} \sum_{i=1}^{n} E[X_i^2] - n E[\bar{X}^2]$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

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

$$\mu_{\bar{X}} = \mu$$

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

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

If $\hat{\Theta}$ is an estimator for θ

Def The bias of $\hat{\Theta}$ is:

$$b_n(\theta) = E[\hat{\Theta}] - \theta$$

Ex: Bernoulli w/ parameter θ $\quad P(X=1) = \theta = 1 - P(X=0)$

good estimator: $\bar{X}$ $\quad E[\bar{X}] = \mu = \theta$

not great estimator: $\hat{\Theta} = \frac{1}{2}$ $\quad b_n(\hat{\Theta}) = \frac{1}{2} - \theta$

Ex: general populations

$$\hat{\Theta} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n} \qquad E[\hat{\Theta}] = \frac{n-1}{n} E[S^2]$$

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

as an estimator for σ^2, bias of $\hat{\Theta}$ is

$$b_n(\sigma^2) = E[\hat{\Theta}] - \sigma^2 = -\frac{1}{n}\sigma^2$$

Note $\lim_{n \to \infty} b_n(\sigma^2) = 0$

Def An estimator $\hat{\Theta}$ for a parameter θ is asymptotically unbiased if $\lim_{n \to \infty} b_n(\theta) = 0$ all θ.

"Efficiency"

Want an unbiased estimator $\hat{\Theta}$ w/ small variance

Theorem "Cramér-Rao inequality"

If $\hat{\Theta}$ is an unbiased estimator for θ, population described by $f(X,\theta)$, which is continuously differentiable then

$$\operatorname{var}\hat{\Theta} \geq \frac{1}{n\,E\left[\left(\frac{\partial \ln f(X)}{\partial \theta}\right)^2\right]}$$

"Fisher Information" represents information about θ obtainable from a measurements.

Idea: if we observe a value $X = x_0$ how much do we know about θ?

$$I(\theta) = E\left[\left(\frac{d \ln f(X,\theta)}{d\theta}\right)^2\right]$$

"Fisher information"

Def An unbiased estimator $\hat{\Theta}$ of θ is minimum variance if its variance is no larger than that of any other unbiased estimator.

Theorem: $\hat{\Theta}$ is min. variance if

$$\operatorname{var}(\hat{\Theta}) = \frac{1}{n I(\theta)}$$

efficiency

$$e(\hat{\Theta}) = \frac{\left(1/nI(\theta)\right)}{\operatorname{var}\hat{\Theta}}$$

Relative efficiency

$$\frac{\operatorname{var}\hat{\Theta}_1}{\operatorname{var}\hat{\Theta}_2}$$ efficiency of $\hat{\Theta}_2$ relative to $\hat{\Theta}_1$