PLSC 502 – Autumn 2016
Estimation and Estimators

October 11, 2016

Random Variables, Take Two

For:

$$X_i = \mu + u_i$$

and so

$$u_i = X_i - \mu,$$

Random Variables, Take Two

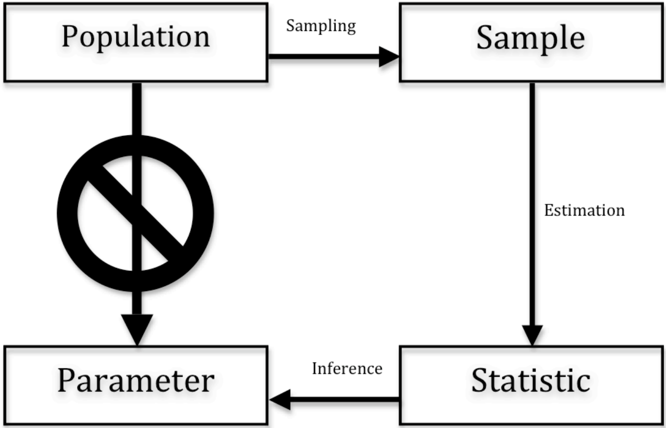
That means

$$\begin{aligned}
\mathsf{E}(u) &= \mathsf{E}(X-\mu) \\
&= \mathsf{E}(X)-\mathsf{E}(\mu) \\
&= \mathsf{E}(X)-\mu \\
&= \mu-\mu \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{Var}(X) &= \mathsf{E}[(X-\mu)^2] \\
&= \mathsf{E}(u^2)
\end{aligned}$$

$$\begin{aligned}
\mathsf{Var}(u) &= \mathsf{E}[(u-E(u))^2] \\
&= \mathsf{E}[(u-0)^2] \\
&= \mathsf{E}(u^2).
\end{aligned}$$


Population
Sampling
Sample
Estimation
Parameter
Inference
Statistic

Estimation Example: $\bar{X}$

$$
\begin{aligned}
\bar{X} &= \frac{1}{N}\sum_{i=1}^{N} X_i \\
&= \frac{1}{N}\sum_{i=1}^{N}(\mu + u_i) \\
&= \frac{1}{N}\sum_{i=1}^{N}(\mu) + \frac{1}{N}\sum_{i=1}^{N}(u_i) \\
&= \frac{1}{N}(N\mu) + \frac{1}{N}\sum_{i=1}^{N}(u_i) \\
&= \mu + \bar{u}
\end{aligned}
$$

Properties of Estimators

Small-Sample Properties

- Hold irrespective of N
- "Small sample estimators"

Large-Sample (Asymptotic) Properties

- Hold as $N \to \infty$
- "More is better"

Unbiasedness

Means:

$$\mathsf{E}(\hat{\theta}) = \theta$$

"Bias":

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

Example: $\bar{X}$

$$\begin{aligned} \mathsf{E}(\bar{X}) &= \mathsf{E}(\mu + \bar{u}) \\ &= \mathsf{E}(\mu) + \mathsf{E}(\bar{u}) \\ &= \mu + 0 \\ &= \mu \end{aligned}$$

Multiple Unbiased Estimators

For $N = 2$:

$$Z = \lambda_1 X_1 + \lambda_2 X_2.$$

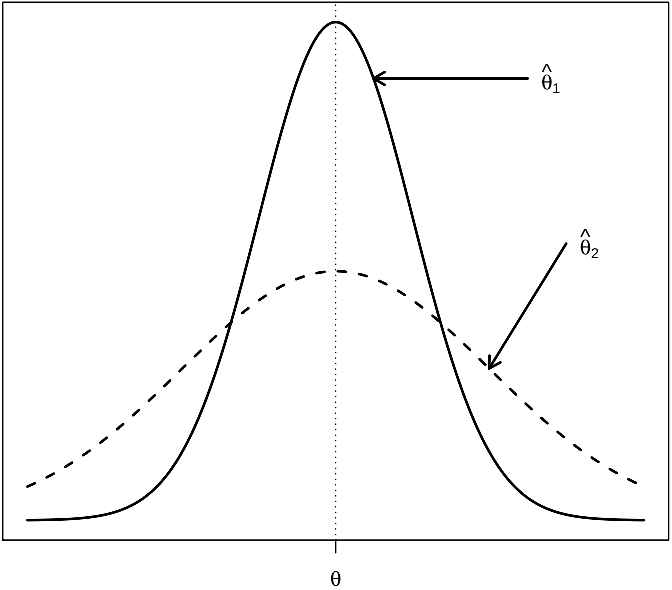
note that

$$\begin{aligned} \mathsf{E}(Z) &= \mathsf{E}(\lambda_1 X_1 + \lambda_2 X_2) \\ &= \mathsf{E}(\lambda_1 X_1) + \mathsf{E}(\lambda_2 X_2) \\ &= \lambda_1 \mathsf{E}(X_1) + \lambda_2 \mathsf{E}(X_2) \\ &= \lambda_1 \mu + \lambda_2 \mu \\ &= (\lambda_1 + \lambda_2)\mu \end{aligned}$$

Means

$$\mathsf{E}(Z) = \mu \iff (\lambda_1 + \lambda_2) = 1.0$$

Efficiency



Efficiency (continued)

Note:

$$\begin{aligned} \mathrm{Var}(Z) &= \mathrm{Var}(\lambda_1 X_1 + \lambda_2 X_2) \\ &= (\lambda_1^2 + \lambda_2^2)\sigma^2 \end{aligned}$$

and:

$$\begin{aligned} \lambda_1^2 + \lambda_2^2 &= \lambda_1^2 + (1 - \lambda_1)^2 \\ &= \lambda_1^2 + (1 - 2\lambda_1 + \lambda_1^2) \\ &= 2\lambda_1^2 - 2\lambda_1 + 1. \end{aligned}$$

Minimize:

$$\frac{\partial 2\lambda_1^2 - 2\lambda_1 + 1}{\partial \lambda_1} = 4\lambda_1 - 2$$

$$\begin{aligned} 4\lambda_1 - 2 &= 0 \\ \lambda_1 &= 0.5 \end{aligned}$$

Mean Squared Error

$$\begin{aligned} \mathsf{MSE}(\hat{\theta}) &= \mathsf{E}[(\hat{\theta} - \theta)^2] \\ &= \mathsf{E}[B(\hat{\theta})^2] \\ &= \mathsf{Var}(\hat{\theta}) + [B(\hat{\theta})^2] \end{aligned}$$

Among unbiased estimators, the efficient estimator will always have the smallest MSE (because $B(\hat{\theta}) = [B(\hat{\theta})]^2 = 0$).

Comparing Estimators via MSE

$\bar{X}$ has:

- $B(\bar{X}) = 0$
- $\mathrm{Var}(\bar{X}) = \sigma^2/N$, so
- $\mathrm{MSE}(\bar{X}) = \sigma^2/N + (0)^2 \; = \; \sigma^2/N$.

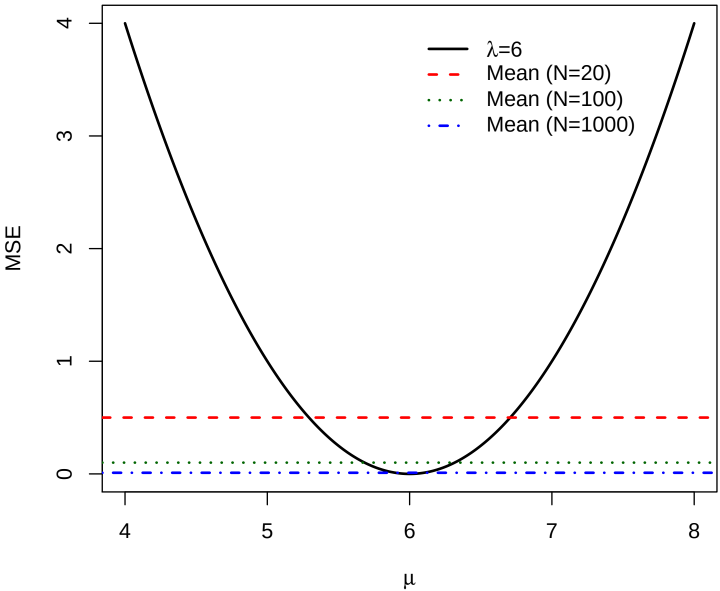
$\lambda = 6$ has:

$$\begin{aligned} B(\lambda) &= \mathrm{E}(\lambda - \mu) \\ &= \mathrm{E}(6) - \mathrm{E}(\mu) \\ &= 6 - \mu, \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(\lambda) &= \mathrm{Var}(6) \\ &= 0 \end{aligned}$$

and so:

$$\begin{aligned} \mathrm{MSE}(\lambda) &= \mathrm{Var}(\lambda) + [B(\lambda)]^2 \\ &= 0 + (6 - \mu)^2 \\ &= 36 - 12\mu + \mu^2 \end{aligned}$$



The black line is the MSE of λ, expressed as a function of the "true" population mean μ. The other colored lines are the MSEs for $\bar{X}$, under the assumption that $\sigma^2 = 10$ and $N = \{20, 100, 1000\}$, respectively.

Large-Sample Properties: Consistency

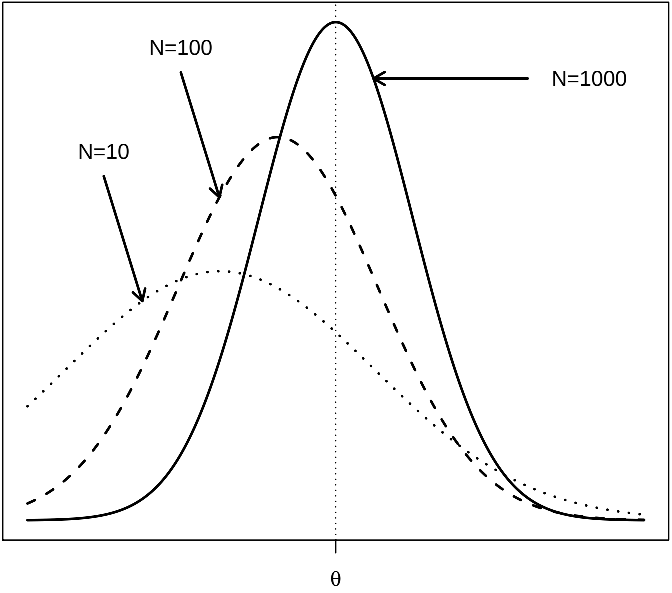
An estimator $\hat{\theta}$ is *consistent* if:

$$\lim_{N \to \infty} \Pr[|\hat{\theta} - \theta| < \epsilon] = 1.0$$

for an arbitrarily small $\epsilon > 0$

Equivalently:

$$\mathrm{E}(\hat{\theta}_N) \to \theta \text{ as } N \to \infty$$


N=100
N=1000
N=10
θ

Estimation, Generally

- Unbiased $>$ Consistent $>$ Biased
- Fully Efficient $>$ Asymptotically Efficient $>$ Inefficient
- MSE is <u>one</u> way to trade off bias vs. efficiency