1 Evaluating Estimators

Recall:

Definition: A *statistic* is a function of some or all of the data collected that does not rely on any unknown parameter values.

Definition: An *estimator* is a statistic That is meant to estimate an unknown parameter value *NOTE:* Since an estimator is a stistic, it cannont be a function of any unknown values (typically, parameters)

1.1 Examples

Let $X_1, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$ where $-\infty < \mu < \infty$ and $0 < \sigma < \infty$ are unknown.

1. $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an estimator for μ
2. $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an estimator for σ^2
3. $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is another estimator for σ^2
4. X_1 is an estimator for μ
5. $\frac{X_1+X_2}{2}$ is an estimator for μ
6. $\frac{(X_1-X_2)^2}{2}$ is an estimator for σ^2

We can see that there are many estimators for the same parameters, so how do we determine the best estimator?

Definition: Let $\hat{\theta}$ be an estimator for the unknown parameter, θ. Then if $E[\hat{\theta}] = \theta$ then $\hat{\theta}$ is said to be *unbiased*. If $E[\hat{\theta}] \neq \theta$ then $\hat{\theta}$ is said to be *biased*.

Definition: The *bias* of a point estimator $\hat{\theta}$ for unknown parameter θ is given as $B(\hat{\theta}) = E[\hat{\theta}] - \theta$

Definition: the *Mean square error* of a point estimator $\hat{\theta}$ for unknown parameter θ is given as $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$

1.2 Examples

Let $X_1, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$ where $-\infty < \mu < \infty$ and $0 < \sigma < \infty$ are unknown.

1. We said that $\bar{X}$ was an estimator for μ, so

$$\begin{aligned} E[\bar{X}] &= \mu \leftarrow \text{ Since } \bar{X} \sim N(\mu, \tfrac{\sigma^2}{n}) \\ \Rightarrow B(\bar{X}) &= E[\bar{X}] - \mu \\ &= \mu - \mu \\ &= 0 \end{aligned}$$

2. We said that S^2 was an estimator for σ^2, so

$$\begin{aligned} E[\frac{n-1}{\sigma^2}S^2] &= n-1 \leftarrow \text{ Since } \tfrac{n-1}{\sigma^2}S^2 \sim \chi^2_{n-1} \equiv \Gamma(\tfrac{n-1}{2}, 2) \\ \Rightarrow E[S^2] &= \sigma^2 \\ \Rightarrow B(S^2) &= E[S^2] - \sigma^2 \\ &= \sigma^2 - \sigma^2 \\ &= 0 \end{aligned}$$

3. We said that X_1 is an estimator for μ, so

$$\begin{aligned} E[X_1] &= \mu \leftarrow \text{ Since } X_1 \sim N(\mu, \sigma^2) \\ \Rightarrow B(X_1) &= E[X_1] - \mu \\ &= \mu - \mu \\ &= 0 \end{aligned}$$

4. We said that $\frac{(X_1-X_2)^2}{2}$ is an estimator for σ^2. So,

$$\begin{aligned} E[\frac{(X_1-X_2)^2}{2}] &= E[\frac{1}{2}(X_1^2 - 2X_1X_2 + X_2^2)] \\ &= \frac{1}{2}(E[X_1^2] - 2E[X_1X_2] + X_2^2]) \\ &= \frac{1}{2}(\sigma^2 + \mu^2 - 2\mu^2 + \sigma^2 + \mu^2]) \leftarrow \text{ because } X_1 \text{ and } X_2 \text{ are i.i.d. } N(\mu, \sigma^2) \\ &= \frac{1}{2}(2\sigma^2) \\ &= \sigma^2 \\ \Rightarrow B(\frac{(X_1-X_2)^2}{2}) &= E[\frac{(X_1-X_2)^2}{2}] - \sigma^2 \\ &= \sigma^2 - \sigma^2 \\ &= 0 \end{aligned}$$

Note, both $\bar{X}$ and X_1 are unbiased for μ. so which estimator is better? The one that has a lower MSE, which is means to measure how *precise* the estimator is (the bias is a measure of how *accurate* an estimator is).

5. The MSE of $\bar{X}$ is

$$\begin{aligned} \bar{X} - \mu &\sim N(0, \frac{\sigma^2}{n}) \leftarrow \text{ since } \bar{X} \sim N(\mu, \tfrac{\sigma^2}{n}) \\ \Rightarrow E[(\bar{X} - \mu)^2] &= V[(\bar{X} - \mu)] + E^2[\bar{X} - \mu] \\ &= \frac{\sigma^2}{n} + 0 \\ &= \frac{\sigma^2}{n} \end{aligned}$$

The MSE of X_1 is

$$\begin{aligned} X_1 - \mu &\sim N(0, \sigma^2) \leftarrow \text{ since } X_1 \sim N(\mu, \sigma^2) \\ \Rightarrow E[(X_1 - \mu)^2] &= V[(X_1 - \mu)] + E^2[X_1 - \mu] \\ &= \sigma^2 + 0 \\ &= \sigma^2 \end{aligned}$$

Since $MSE(\bar{X}) < MSE(X_1)$ we conclude that $\bar{X}$ is a better estimator for μ, because it is more likely to be closer to the value of μ

1.3 Exercises

1. Let $X_1 ... X_n$ be i.i.d. $Exp(\delta)$
   a) Find $B(X_1)$
   b) Find $MSE(X_1)$
   c) Find $B(\bar{X})$
   d) Find $MSE(\bar{X})$
   e) Which Estimator is better?
2. Let $X_1 ... X_n$ be i.i.d. $Poisson(\lambda)$
   a) Find $B(X_1)$
   b) Find $MSE(X_1)$
   c) Find $B(\bar{X})$
   d) Find $MSE(\bar{X})$
   e) Which Estimator is better?

1.4 Solutions

1. Let $X_1 ... X_n$ be i.i.d. $Exp(\delta)$

a) Find $B(X_1)$
Solution:

$$\begin{aligned} B(X_1) &= E[X_1] - \delta \\ &= \delta - \delta \\ &= 0 \end{aligned}$$

b) Find $MSE(X_1)$
Solution:

$$\begin{aligned} MSE(X_1) &= E[(X_1 - \delta)^2] \\ &= E[X_1^2 - 2\delta X_1 + \delta^2] \\ &= E[X_1^2] - 2\delta E[X_1] + \delta^2 \\ &= V[X_1] + E^2[X_1] - 2\delta E[X_1] + \delta^2 \\ &= \delta^2 + \delta^2 - 2\delta^2 + \delta^2 \\ &= \delta^2 \end{aligned}$$

c) Find $B(\bar{X})$
Solution:

$$\begin{aligned} B(\bar{X}) &= E[\bar{X}] - \delta \\ &= \delta - \delta \leftarrow \text{ because } \textstyle\sum_{i=1}^{n} X_1 \sim \Gamma(n, \delta) \\ &= 0 \end{aligned}$$

d) Find $MSE(\bar{X})$
Solution:

$$\begin{aligned} MSE(\bar{X}) &= E[(\bar{X} - \delta)^2] \\ &= E[\bar{X}^2 - 2\delta\bar{X} + \delta^2] \\ &= E[\bar{X}^2] - 2\delta E[\bar{X}] + \delta^2 \\ &= V[\bar{X}] + E^2[\bar{X}] - 2\delta E[\bar{X}] + \delta^2 \\ &= \frac{\delta^2}{n} + \delta^2 - 2\delta^2 + \delta^2 \\ &= \frac{\delta^2}{n} \end{aligned}$$

e) Which Estimator is better?
Solution:
Since both $\bar{X}$ and X_1 are unbiased, we can compare MSEs to decide which is better. Since $MSE(\bar{X}) < MSE(X_1)$ we conclude that $\bar{X}$ is the better estimator.

2. Let $X_1 ... X_n$ be i.i.d. $Poisson(\lambda)$

a) Find $B(X_1)$
Solution:

$$\begin{aligned} B(X_1) &= E[X_1] - \lambda \\ &= \lambda - \lambda \\ &= 0 \end{aligned}$$

b) Find $MSE(X_1)$
Solution:

$$\begin{aligned} MSE(X_1) &= E[(X_1 - \lambda)^2] \\ &= E[X_1^2 - 2\lambda X_1 + \lambda^2] \\ &= E[X_1^2] - 2\lambda E[X_1] + \lambda^2 \\ &= V[X_1] + E^2[X_1] - 2\lambda E[X_1] + \lambda^2 \\ &= \lambda + \lambda^2 - 2\lambda^2 + \lambda^2 \\ &= \lambda \end{aligned}$$

c) Find $B(\bar{X})$
Solution:

$$\begin{aligned} B(\bar{X}) &= E[\bar{X}] - \lambda \\ &= \frac{1}{n} E[\sum_{i=1}^{n} X_i] - \lambda \\ &= \frac{1}{n} n\lambda - \lambda \leftarrow \text{ Since } \textstyle\sum_{i=1}^{n} X_1 \sim Poiss(n\lambda) \\ &= \lambda - \lambda \\ &= 0 \end{aligned}$$

d) Find $MSE(\bar{X})$
Solution:

$$\begin{aligned} MSE(\bar{X}) &= E[(\bar{X}-\lambda)^2] \\ &= E[\bar{X}^2 - 2\lambda\bar{X} + \lambda^2] \\ &= E[\bar{X}^2] - 2\lambda E[\bar{X}] + \lambda^2 \\ &= V[\bar{X}] + E^2[\bar{X}] - 2\lambda E[\bar{X}] + \lambda^2 \\ &= \frac{1}{n^2} V[(\sum_{i=1}^{n} X_i)] + \lambda^2 - 2\lambda^2 + \lambda^2 \\ &= \frac{1}{n^2} n\lambda \\ &= \frac{\lambda}{n} \end{aligned}$$

e) Which Estimator is better?
Solution:
Since both $\bar{X}$ and X_1 are unbiased, we can compare MSEs to decide which is better. Since $MSE(\bar{X}) < MSE(X_1)$ we conclude that $\bar{X}$ is the better estimator.

2 Common Unbiased Estimators

We have already discussed the most common unbiased estimators one will work with when our population is $N(\mu, \sigma^2)$, namely $\bar{X}$ and S^2. what are some other common situations we might get in to?

1. Instead of measuring a quantity in a population we want to measure a proportion of a population
   - i.e. what proportion, $0 < p < 1$, of UCONN students have a car?
   - Again we take a sample from this population of size n
   - $X_1, ..., X_n$ can be thought of as i.i.d. random variables with distribution $Bern(p)$
   - If $X_i = 1$ then the person has a car, and if $X_i = 0$ then the person doesn't have a car
   - Our best estimate of p will be $\frac{1}{n}\sum_{i=1}^{n} X_i$, the sample proportion of people who have a car
2. What if we want to compare two independent populations, each distributed $N(\mu_i, \sigma_i^2)$ for $i = 1, 2$
   - i.e. What is the difference between the average heights of male and female UCONN students
   - We take samples from both populations of sizen_1, n_2

- $X_1, ..., X_{n_1}$ can be thought of as the n_1 female students sampled, and $Y_1, ..., Y_{n_2}$ can be thought of as the n_2 male students sampled. we assume that all random variables measured are independent and all the $X's$ are distributed $N(\mu_1, \sigma_1^2)$ and the $Y's$ are distributed $N(\mu_2, \sigma_2^2)$
- Our best estimate of the difference in average heights $\mu_1 - \mu_2$ will be $\bar{X} - \bar{Y}$

3. Another situation that can often come up is measuring the difference between two independent population proportions, $p_1 - p_2$, where both p_1 and p_2 range from 0 to 1.
   - i.e. What is the difference between the proportion of male and female UCONN students with cars
   - We take samples from both populations of sizen_1, n_2
   - $X_1, ..., X_{n_1}$ can be thought of as the n_1 female students sampled, and $Y_1, ..., Y_{n_2}$ can be thought of as the n_2 male students sampled. We assume that all random variables measured are independent and all the $X's$ are distributed $Bern(p_1)$ and the $Y's$ are distributed $Bern(p_2)$
   - Our best estimate of the difference in average heights $p_1 - p_2$ will be $\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i$

2.1 Exercises

Show that the following are unbiased statistics for their corresponding parameter and find their MSE

1. $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ for p where $X_1, ..., X_n$ are i.i.d. $Bern(p)$
2. $\widehat{\mu_1 - \mu_2} = \bar{X} - \bar{Y}$ for $\mu_1 - \mu_2$ where $X_1, ..., X_{n_1}$ are distributed $N(\mu_1, \sigma_1^2)$, $Y_1, ..., Y_{n_2}$ are distributed $N(\mu_2, \sigma_2^2)$, and all $X's$ and $Y's$ are assumed independent
3. $\widehat{p_1 - p_2} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i$ for $p_1 - p_2$ where $X_1, ..., X_{n_1}$ are distributed $Bern(p_1)$, $Y_1, ..., Y_{n_2}$ are distributed $Bern(p_2)$, and all $X's$ and $Y's$ are assumed independent

2.2 Solutions

Show that the following are unbiased statistics for their corresponding parameter and find their MSE

1. $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ for p where $X_1, ..., X_n$ are i.i.d. $Bern(p)$
   Solution:

$$B(\hat{p}) = E[\frac{1}{n}\sum_{i=1}^{n} X_i] - p$$

$$
\begin{aligned}
&= \frac{1}{n}E[\sum_{i=1}^{n} X_i] - p \\
&= \frac{1}{n}(np) - p \leftarrow \text{ since } \textstyle\sum_{i=1}^{n} X_i \sim Bin(n,p) \\
&= 0 \\
MSE(\hat{p}) &= E[(\frac{1}{n}\sum_{i=1}^{n} X_i - p)^2] \\
&= V[\frac{1}{n}\sum_{i=1}^{n} X_i - p] + E^2[\frac{1}{n}\sum_{i=1}^{n} X_i - p] \\
&= \frac{1}{n^2}V[\sum_{i=1}^{n} X_i] + (\frac{1}{n}E[\sum_{i=1}^{n} X_i] - p)^2 \\
&= \frac{1}{n^2}(np(1-p)) + (\frac{1}{n}(np) - p)^2 \\
&= \frac{p(1-p)}{n}
\end{aligned}
$$

2. $\widehat{\mu_1 - \mu_2} = \bar{X} - \bar{Y}$ for $\mu_1 - \mu_2$ where $X_1, ..., X_{n_1}$ are distributed $N(\mu_1, \sigma_1^2)$, $Y_1, ..., Y_{n_2}$ are distributed $N(\mu_2, \sigma_2^2)$, and all $X's$ and $Y's$ are assumed independent
Solution:

$$
\begin{aligned}
B(\widehat{\mu_1 - \mu_2}) &= E[\bar{X} - \bar{Y}] - (\mu_1 - \mu_2) \\
&= E[\bar{X}] - E[\bar{Y}] - \mu_1 + \mu_2 \\
&= \mu_1 - \mu_2 - \mu_1 + \mu_2 \leftarrow \text{ Since } \bar{X} \sim N(\mu_1, \tfrac{\sigma_1^2}{n_1}) \text{ and } \bar{Y} \sim N(\mu_2, \tfrac{\sigma_2^2}{n_2}) \\
&= 0 \\
MSE(\widehat{\mu_1 - \mu_2}) &= E[(\bar{X} - \bar{Y} - (\mu_1 - \mu_2))^2] \\
&= V[\bar{X} - \bar{Y} - (\mu_1 - \mu_2)] + E^2[\bar{X} - \bar{Y} - (\mu_1 - \mu_2)] \\
&= V[\bar{X} - \bar{Y}] + (E[\bar{X} - \bar{Y}] - (\mu_1 - \mu_2))^2 \\
&= V[\bar{X}] + V[\bar{Y}] + (E[\bar{X}] - E[\bar{Y}] - (\mu_1 - \mu_2))^2 \leftarrow \text{Since all the } X's \text{ and } Y's \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{are independent} \\
&= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} + (\mu_1 - \mu_2 - \mu_1 + \mu_2)^2 \\
&= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}
\end{aligned}
$$

3. $\widehat{p_1 - p_2} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i$ for $p_1 - p_2$ where $X_1, ..., X_{n_1}$ are distributed $Bern(p_1)$, $Y_1, ..., Y_{n_2}$ are distributed $Bern(p_2)$, and all $X's$ and $Y's$ are assumed independent
Solution:

$$
\begin{aligned}
B(\widehat{p_1 - p_2}) &= E[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i] - (p_1 - p_2) \\
&= E[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i] - E[\frac{1}{n_2}\sum_{i=1}^{n_2} Y_i] - p_1 + p_2 \\
&= \frac{1}{n_1}E[\sum_{i=1}^{n_1} X_i] - \frac{1}{n_2}E[\sum_{i=1}^{n_2} Y_i] - p_1 + p_2 \\
&= \frac{1}{n_1}(n_1 p_1) - \frac{1}{n_2}(n_2 p_2) - p_1 + p_2 \leftarrow \text{ Since } \textstyle\sum_{i=1}^{n_1} X_i \sim Bin(n_1, p_1) \text{ and} \\
&\qquad\qquad \textstyle\sum_{i=1}^{n_2} Y_i \sim Bin(n_2, p_2) \\
&= 0 \\
MSE(\widehat{p_1 - p_2}) &= E[(\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i - (p_1 - p_2))^2] \\
&= V[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i - (p_1 - p_2)] + E^2[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i - (p_1 - p_2)] \\
&= V[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i] + (E[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i] - (p_1 - p_2))^2 \\
&= V[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i] + V[\frac{1}{n_2}\sum_{i=1}^{n_2} Y_i] + (E[\frac{1}{n_1}\sum_{i=1}^{n_1} X_i] - E[\frac{1}{n_2}\sum_{i=1}^{n_2} Y_i] - (p_1 - p_2))^2 \\
&\qquad\qquad \uparrow \text{Since all the } X's \text{ and } Y's \text{ are independent} \\
&= \frac{1}{n_1^2}V[\sum_{i=1}^{n_1} X_i] + \frac{1}{n_2^2}V[\sum_{i=1}^{n_2} Y_i] + (\frac{1}{n_1}E[\sum_{i=1}^{n_1} X_i] - \frac{1}{n_1}E[\sum_{i=1}^{n_2} Y_i] - (p_1 - p_2))^2 \\
&= \frac{n_1 p_1(1-p_1)}{n_1^2} + \frac{n_2 p_2(1-p_2)}{n_2} + (\frac{n_1 p_1}{n_1} - \frac{n_2 p_2}{n_2} - p_1 + p_2)^2 \\
&= \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}
\end{aligned}
$$

3 Error of Estimation

Definition: The *Error* of an estimator $\hat{\theta}$ is defined to be

$$\varepsilon = |\hat{\theta} - \theta|$$

Because $\hat{\theta}$ is a random variable, we see that ε is also a random variable. Usually what we are interested in is the probability that the error of our estimate will be of a certain size. Let $b > 0$, then

$$P(\varepsilon < b) = P(|\hat{\theta} - \theta| < b)$$

$$
\begin{aligned}
&= P(-b < \hat{\theta} - \theta < b) \\
&= P(\theta - b < \hat{\theta} < \theta + b)
\end{aligned}
$$