Unbiased Estimators

Gene Quinn

Estimators

We may give a very broad definition of an **estimator** as a function of a random sample $x = (x_1, \ldots, x_n)$:

$$\hat{\theta} \; = \; h(x_1, x_2, \ldots, x_n)$$

Estimators

We may give a very broad definition of an **estimator** as a function of a random sample $x = (x_1, \ldots, x_n)$:

$$\hat{\theta} \;=\; h(x_1, x_2, \ldots, x_n)$$

Over time, many ideas about what properties are desirable in an estimator have evolved.

One of the simplest notions is that of *unbiasedness*.

Unbiasedness

Ideally, we would like an estimator $\hat{\theta}$ to neither consistently overestimate nor consistently underestimate the population parameter θ.

Unbiasedness

Ideally, we would like an estimator $\hat{\theta}$ to neither consistently overestimate nor consistently underestimate the population parameter θ.

We can state this in terms of the expected value of the estimator, $\mathsf{E}(\hat{\theta})$:

We would like to have:

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

regardless of the actual value of the population parameter θ.

Unbiasedness

Definition: Let $x_1, \ldots, x_n$ be a random sample from a population with continuous pdf $f_X(x;\theta)$ where θ is an unknown parameter.

An estimator

$$\hat{\theta} \;=\; h(x_1, \ldots, x_n)$$

is said to be *unbiased* [for θ] if

$$\mathsf{E}(\hat{\theta}) \;=\; \theta \quad \forall \theta$$

regardless of the actual value of the population parameter θ.

Unbiasedness

Definition: Let $x_1, \dots, x_n$ be a random sample from a population with continuous pdf $f_X(x; \theta)$ where θ is an unknown parameter.

An estimator

$$\hat{\theta} = h(x_1, \dots, x_n)$$

is said to be *unbiased* [for θ] if

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

regardless of the actual value of the population parameter θ.

A similar definition can be stated for discrete random variables.

Unbiasedness

Example: Let $x_1, \ldots, x_n$ be a random sample from a $N(\mu, \sigma^2)$ population.

We will show that the sample mean

$$\overline{x} = \sum_{i=1}^{n} x_i$$

is an unbiased estimator for the population mean μ.

Unbiasedness

Example: Let $x_1, \ldots, x_n$ be a random sample from a $N(\mu, \sigma^2)$ population.

We will show that the sample mean

$$\overline{x} = \sum_{i=1}^{n} x_i$$

is an unbiased estimator for the population mean μ.

The moment-generating function $M_X(t)$ for each of the x_i is:

$$M_X(t) = \exp\left(t\mu + \frac{1}{2}t^2\sigma^2\right)$$

Unbiasedness

Now $\overline{x}$ can be thought of as the sum of the transformed random variables $y_i = x_i/n$.

Unbiasedness

Now $\overline{x}$ can be thought of as the sum of the transformed random variables $y_i = x_i/n$.

From a theorem on moment generating functions, we know that

$$M_Y(t) = M_X(t/n) = \exp\left(\mu\frac{t}{n} + \frac{1}{2}\frac{t^2\sigma^2}{n^2}\right)$$

Unbiasedness

Now $\overline{x}$ can be thought of as the sum of the transformed random variables $y_i = x_i/n$.

From a theorem on moment generating functions, we know that

$$M_Y(t) = M_X(t/n) = \exp\left(\mu\frac{t}{n} + \frac{1}{2}\frac{t^2\sigma^2}{n^2}\right)$$

A second theorem states that moment-generating function of the sum of the n independent random variables y_i is the product of their individual mgfs:

$$M_{\overline{x}}(t) = \prod_{i=1}^{n} M_Y(t) = \prod_{i=1}^{n} \exp\left(\mu\frac{t}{n} + \frac{1}{2}\frac{t^2\sigma^2}{n^2}\right)$$

Unbiasedness

We can write the product on the right hand side as a single exponential by adding the exponents,

$$M_{\overline{x}}(t) \;=\; \exp\left(\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right)$$

Unbiasedness

We can write the product on the right hand side as a single exponential by adding the exponents,

$$M_{\overline{x}}(t) \;=\; \exp\left(\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right)$$

By inspection, we see that the right hand side is the moment-generating function of a normal random variable with expected value μ and variance σ^2/n.

Unbiasedness

We can write the product on the right hand side as a single exponential by adding the exponents,

$$M_{\overline{x}}(t) \;=\; \exp\left(\mu t + \frac{1}{2}\frac{\sigma^2}{n}t^2\right)$$

By inspection, we see that the right hand side is the moment-generating function of a normal random variable with expected value μ and variance σ^2/n.

This establishes that

$$\mathsf{E}(\overline{x}) \;=\; \mu$$

so the sample mean $\overline{x}$ is an unbiased estimator for the population mean μ.

Unbiasedness

Now suppose $x = (x_1, \ldots, x_n)$ is a random sample from a $N(\mu, \sigma^2)$ population and consider whether the estimator

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

is unbiased for σ^2.

Unbiasedness

Now suppose $x = (x_1, \ldots, x_n)$ is a random sample from a $N(\mu, \sigma^2)$ population and consider whether the estimator

$$\hat{\sigma}^2 \;=\; \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

is unbiased for σ^2.

By definition,

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right]$$

Unbiasedness

Expanding the square,

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right)\right]$$

Unbiasedness

Expanding the square,

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\overline{x} + \overline{x}^2\right)\right]$$

Since the expectation of a sum is the sum of the expectations, we can write

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n}x_i^2\right] - \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n}2x_i\overline{x}\right] + \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\overline{x}^2\right]$$

Unbiasedness

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} 2x_i\overline{x}\right] + \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} \overline{x}^2\right]$$

For the purpose of summation, $\overline{x}$ is a constant so we can move it outside the sum (and move $1/n$ inside):

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[2\overline{x}\sum_{i=1}^{n} \frac{x_i}{n}\right] + \mathsf{E}\left[\overline{x}^2\sum_{i=1}^{n} \frac{1}{n}\right]$$

Unbiasedness

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} 2x_i\overline{x}\right] + \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} \overline{x}^2\right]$$

For the purpose of summation, $\overline{x}$ is a constant so we can move it outside the sum (and move $1/n$ inside):

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[2\overline{x}\sum_{i=1}^{n} \frac{x_i}{n}\right] + \mathsf{E}\left[\overline{x}^2\sum_{i=1}^{n} \frac{1}{n}\right]$$

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[2\overline{x}\left(\overline{x}\right)\right] + \mathsf{E}\left[\overline{x}^2\left(1\right)\right]$$

Unbiasedness

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - 2\mathsf{E}\left[\overline{x}^2\right] + \mathsf{E}\left[\overline{x}^2\right]$$

Unbiasedness

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - 2\mathsf{E}\left[\overline{x}^2\right] + \mathsf{E}\left[\overline{x}^2\right]$$

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

Unbiasedness

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - 2\mathsf{E}\left[\overline{x}^2\right] + \mathsf{E}\left[\overline{x}^2\right]$$

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

$$\mathsf{E}(\hat{\sigma}^2) = \frac{1}{n}\mathsf{E}\left[\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

Unbiasedness

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - 2\mathsf{E}\left[\overline{x}^2\right] + \mathsf{E}\left[\overline{x}^2\right]$$

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \mathsf{E}\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \frac{1}{n}\mathsf{E}\left[\sum_{i=1}^{n} x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \frac{1}{n}\sum_{i=1}^{n}\mathsf{E}\left[x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

Unbiasedness

Now

$$x_i \sim N(\mu, \sigma^2), \quad i = 1, \ldots, n$$

Unbiasedness

Now

$$x_i \sim N(\mu, \sigma^2), \quad i = 1, \ldots, n$$

This means that

$$\sigma^2 \; = \; \mathsf{E}(x_i^2) - [\mathsf{E}(x_i)]^2 \; = \; E(x_i^2) - \mu^2$$

from which it follows that

$$\mathsf{E}(x_i^2) \; = \; \sigma^2 + \mu^2$$

Unbiasedness

Also

$$\overline{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Unbiasedness

Also

$$\overline{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

This means that

$$\frac{\sigma^2}{n} = \mathsf{E}(\overline{x}^2) - [\mathsf{E}(\overline{x})]^2 = E(\overline{x}^2) - \mu^2$$

from which it follows that

$$\mathsf{E}(\overline{x}^2) = \frac{\sigma^2}{n} + \mu^2$$

Unbiasedness

Now by substitution since

$$\mathsf{E}(x_i^2) \;=\; \sigma^2 + \mu^2 \quad \text{and} \quad \mathsf{E}(\overline{x}^2) \;=\; \frac{\sigma^2}{n} + \mu^2$$

we can rewrite the expression

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[x_i^2\right] - \mathsf{E}\left[\overline{x}^2\right]$$

as

$$\mathsf{E}(\hat{\sigma}^2) \;=\; \frac{1}{n}\sum_{i=1}^{n} (\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} - \mu^2\right) \;=\; (\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} - \mu^2\right)$$

Unbiasedness

Then

$$\mathsf{E}(\hat{\sigma}^2) \ = \ \left(1 - \frac{1}{n}\right)\sigma^2$$

or

$$\mathsf{E}(\hat{\sigma}^2) \ = \ \frac{n-1}{n}\sigma^2$$

Unbiasedness

Then

$$\mathsf{E}(\hat{\sigma}^2) = \left(1 - \frac{1}{n}\right)\sigma^2$$

or

$$\mathsf{E}(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

Because

$$\mathsf{E}(\hat{\sigma}^2) \neq \sigma^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

is *not* an unbiased estimator for σ^2.

Unbiasedness

However,

$$\mathsf{E}\left(\frac{n}{n-1}\hat{\sigma}^2\right) \;=\; \frac{n}{n-1}\mathsf{E}\left(\hat{\sigma}^2\right) \;=\; \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 \;=\; \sigma^2$$

Unbiasedness

However,

$$\mathsf{E}\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \frac{n}{n-1}\mathsf{E}\left(\hat{\sigma}^2\right) = \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 = \sigma^2$$

so

$$\frac{n}{n-1}\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

is an unbiased estimator for σ^2.

Unbiasedness

However,

$$\mathsf{E}\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \frac{n}{n-1}\mathsf{E}\left(\hat{\sigma}^2\right) = \left(\frac{n}{n-1}\right)\left(\frac{n-1}{n}\right)\sigma^2 = \sigma^2$$

so

$$\frac{n}{n-1}\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

is an unbiased estimator for σ^2.

As n becomes large, the difference between $\hat{\sigma}^2$, the maximum likelihood estimate, and S^2, the method of moments estimate, becomes negligible.