Stat 400, section 6.1b  Point Estimates of Mean and Variance
notes by Tim Pilachowski

What we have so far:
Researchers often know a lot about a population, including the probability distribution, but the value of the
population parameter remains unknown. Examples of common parameters are mean \((\mu)\), variance \((\sigma^2)\), median
\((\tilde{\mu})\), and proportion \((p)\). A population parameter has a value, however we usually don’t know what that value is.

“A point estimate of a parameter \(\theta\) is a single number that can be regarded as a sensible value for \(\theta\) … The
selected statistic is called the point estimator of \(\theta\)” The symbol \(\hat{\theta}\) is used for both the random variable and the
calculated value of the point estimate.

Ideally, the point estimator \(\hat{\theta}\) is unbiased, i.e. \(E(\hat{\theta}) = \theta\). In words, the sampling distribution based on the
statistic has an expected value equal to the actual (but unknown) population parameter.

Task 1: Show that the point estimator \(\hat{\mu} = \bar{X}\) (sample mean) is an unbiased estimator of the population
parameter \(\mu\). That is, show \(E(\bar{X}) = \mu\).

We already did this in Lecture 5.4b, as part of the development of the Central Limit Theorem.

Random variables \(X_1, X_2, …, X_n\) form a (simple) random sample of size \(n\) if they meet two (important)
requirements:
1. The \(X_i\)'s are independent random variables.
2. Every \(X_i\) has the same probability distribution.

Given a linear transformation/change of variables which is a sum of \(n\) independent random variables,

\[
Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n \implies E(Y) = \mu_Y = a_1\mu_{X_1} + a_2\mu_{X_2} + \ldots + a_n\mu_{X_n}.
\]

notes on the proof:

\[
\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}
\]

\[
E(\bar{X}) = \frac{1}{n}[E(X_1) + E(X_2) + \ldots + E(X_n)]
\]

\[
= \frac{1}{n}[\mu + \mu + \ldots + \mu]
\]

\[
= \frac{1}{n}[n \cdot \mu]
\]

\[
= \mu
\]

Note that, if we were to pick a single, randomly chosen element of the population, it would be true that our
statistic would be unbiased, i.e. \(E(X) = \mu\). Why don’t we just do that? Why is it better to choose a random
sample of size \(n\)?

Principle: Among all unbiased estimators of a population parameter \(\theta\), choose one that has the minimum
variance.
Task 2: Show that the point estimator $\hat{\sigma}^2 = S^2$ (sample variance) is an unbiased estimator of the population parameter $\sigma^2$. That is, show $E(\hat{S}^2) = \sigma^2$.

First, we take a short side trip, using the formula for sample mean.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \Rightarrow \quad n\bar{X} = \sum_{i=1}^{n} X_i$$

We begin with the “sum of squares” formula.

notes on the proof:

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} (X_i^2 - 2X_i\bar{X} + \bar{X}^2)$$

$$= \sum_{i=1}^{n} X_i^2 - 2\sum_{i=1}^{n} X_i\bar{X} + \sum_{i=1}^{n} \bar{X}^2$$

$$= \sum_{i=1}^{n} X_i^2 - 2\bar{X} \sum_{i=1}^{n} X_i + \bar{X}^2 \sum_{i=1}^{n} 1$$

$$= \sum_{i=1}^{n} X_i^2 - 2\bar{X}(n\bar{X}) + n\bar{X}^2$$

$$= \sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2$$

$$= \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$$

$$= \sum_{i=1}^{n} X_i^2 - n\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right)^2$$

$$= \sum_{i=1}^{n} X_i^2 - \frac{n}{n^2} \left(\sum_{i=1}^{n} X_i\right)^2$$

$$= \sum_{i=1}^{n} X_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} X_i\right)^2$$

We now substitute into the formula for random variable $S^2$, sample variance.

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

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

How is the formula for $S^2$ (sample variance) different from the formula for $V(X)$ (population variance)?
Next is another short side trip, using the shortcut formula for population variance.

\[
V(y) = E(y^2) - [E(y)]^2
\]

\[
V(y) + [E(y)]^2 = E(y^2)
\]

\[
\sigma^2 + \mu^2 = E(y^2)
\]

Finally, we substitute into the shortcut formula for sample variance and simplify.

notes on the proof:

\[
S^2 = \frac{1}{n-1}\left[\sum X_i^2 - \frac{1}{n}\left(\sum X_i\right)^2\right]
\]

\[
E(S^2) = \frac{1}{n-1}\left\{\sum E(X_i^2) - \frac{1}{n}E\left[\sum X_i\right]^2\right\}
\]

\[
= \frac{1}{n-1}\left\{\sum E(X_i^2) - \frac{1}{n}\left\{V(\sum X_i) + E(\sum X_i)^2\right\}\right\}
\]

\[
= \frac{1}{n-1}\left\{\sum (\sigma^2 + \mu^2) - \frac{1}{n}\left\{n\sigma^2 + (n\mu)^2\right\}\right\}
\]

\[
= \frac{1}{n-1}\left\{n\sigma^2 + n\mu^2 - \frac{n\sigma^2}{n} - \frac{n^2\mu^2}{n}\right\}
\]

\[
= \frac{1}{n-1}\left\{n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2\right\}
\]

\[
= \frac{1}{n-1}\left\{(n-1)\sigma^2\right\}
\]

\[
E(S^2) = \sigma^2
\]