# Point Estimation: Unbiased Estimators, Std Error

Engineering Statistics
Section 6.1

Josh Engwer

TTU

25 March 2016

# 1-Sample Point Estimation

**Point Estimation** is a key building block of Statistical Inference:

## Definition

Given a random sample $X_1, \dots, X_n$ from a population with parameter $\theta$.

Then a **point estimator** $\widehat{\theta}$ of parameter $\theta$ is a suitable statistic $T$ of sample:

$$\widehat{\theta} = T(X_1, \dots, X_n)$$

Recall from Chapter 5 that a statistic $T$ is a function of the random sample.

NOTE: Often there are several point estimators for a population parameter.

| POPULATION PARAMETER | POINT ESTIMATOR(S) |
|---|---|
| Mean $\mu$ | $\widehat{\mu} := \overline{X},\ \widehat{\mu} := \widetilde{X},\ \widehat{\mu} := \overline{X}_{tr(10\%)}$ |
| Proportion $p$ | $\widehat{p} := X/n$ |
| Variance $\sigma^2$ | $\widehat{\sigma}^2 := S^2$ |

# 2-Sample Point Estimation

**Point Estimation** is a key building block of Statistical Inference:

## Definition

Given a random sample $X_1, \ldots, X_n$ from a population with parameter $\theta$.

Then a **point estimator** $\widehat{\theta}$ of parameter $\theta$ is a suitable statistic $T$ of sample:

$$\widehat{\theta} = T(X_1, \ldots, X_n)$$

Recall from Chapter 5 that a statistic $T$ is a function of the random sample.

NOTE: Often there are several point estimators for a population parameter.

Random sample $X_1, \ldots, X_{n_1}$ from population with mean $\mu_1$ & variance $\sigma_1^2$
Random sample $Y_1, \ldots, Y_{n_2}$ from population with mean $\mu_2$ & variance $\sigma_2^2$

| POPULATION PARAMETER | POINT ESTIMATOR(S) |
|---|---|
| Mean Difference $\mu_1 - \mu_2$ | $\widehat{\mu}_1 - \widehat{\mu}_2 := \overline{X} - \overline{Y}$ |
| Proportion Difference $p_1 - p_2$ | $\widehat{p}_1 - \widehat{p}_2 := X/n_1 - Y/n_2$ |
| Variance Ratio $\sigma_1^2/\sigma_2^2$ | $\widehat{\sigma}_1^2/\widehat{\sigma}_2^2 := \frac{(n_1-1)S_1^2}{(n_2-1)S_2^2}$ |

# Unbiased Point Estimators (Definition)

How to determine if a point estimator is any good?

## Definition

Given a population with parameter $\theta$.

Then a point estimator $\widehat{\theta}$ is an **unbiased estimator** of $\theta$ if $\mathbb{E}[\,\widehat{\theta}\,] = \theta$

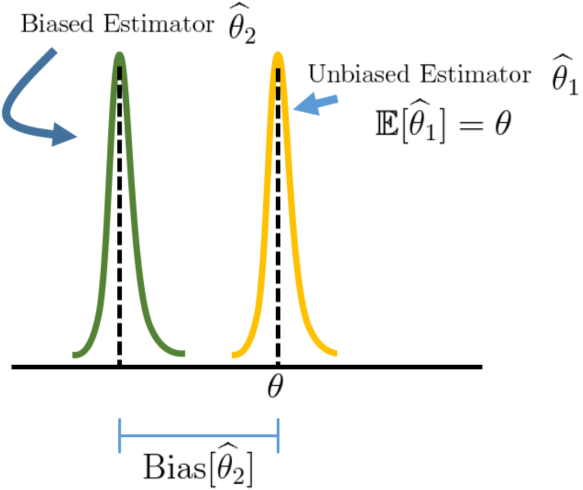
Otherwise, the point estimator is **biased** with a **bias** of $\text{Bias}[\,\widehat{\theta}\,] := \mathbb{E}[\,\widehat{\theta}\,] - \theta$

i.e. A point estimator is unbiased if its sampling distribution is always "centered" at the true value of the population parameter.

If $\text{Bias}[\,\widehat{\theta}\,] < 0$, then $\widehat{\theta}$ tends to <u>underestimate</u> the population parameter value.
If $\text{Bias}[\,\widehat{\theta}\,] > 0$, then $\widehat{\theta}$ tends to <u>overestimate</u> the population parameter value.

# Biased & Unbiased Point Estimators (Visually)

The two curves are the pdf's of the sampling distributions of $\widehat{\theta}_1$ and $\widehat{\theta}_2$



# An Unbiased Estimator of Population Mean $\mu$

Not too surprisingly, the sample mean is an unbiased estimator of the population mean:

**Proposition**

*Given a random sample $X_1, \ldots, X_n$ from a population with mean $\mu$.*

*Then $\overline{X}$ is an unbiased estimator of $\mu$.*

<u>PROOF:</u> Use the definition & properties of expected value:

$$\begin{aligned}
\mathbb{E}[\overline{X}] &= \mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] = \frac{1}{n}\cdot\mathbb{E}\left[\sum_{k=1}^{n} X_k\right] = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[X_k] \\
&= \frac{1}{n}\sum_{k=1}^{n}\mu = \frac{1}{n}\Big[n\mu\Big] = \mu
\end{aligned}$$

$\therefore$ Since $\mathbb{E}[\overline{X}] = \mu$, $\overline{X}$ is an unbiased estimator of $\mu$ $\square$

# An Unbiased Estimator of Population Mean $\mu$

Under certain conditions, the sample median and trimmed means are also unbiased estimators of the population mean:

## Proposition

*Given a random sample* $X_1, \ldots, X_n$ *from a* *symmetric continuous* *population with mean* $\mu$.

*Then* $\overline{X}$, $\widetilde{X}$, $\overline{X}_{tr(10\%)}$, $\overline{X}_{tr(25\%)}$ *are unbiased estimators of* $\mu$.

PROOF: Beyond the scope of this course.

# An Unbiased Estimator of Population Proportion $p$

## Proposition

*Let random variable $X \sim$ Binomial$(n, p)$.*

*Then sample proportion $X/n$ is an unbiased estimator of pop. proportion $p$.*

<u>PROOF:</u> Recall the expected value of a Binomial$(n, p)$ rv: $\mathbb{E}[X] = np$

$$\mathbb{E}[X/n] = \mathbb{E}\left[\frac{X}{n}\right] = \frac{1}{n} \cdot \mathbb{E}[X] = \frac{1}{n}[np] = p$$

$\therefore$ Since $\mathbb{E}[X/n] = p$, $X/n$ is an unbiased estimator of $p$ $\square$

# An Unbiased Estimator of Population Variance $\sigma^2$

The sample variance is an unbiased estimator of the population variance:

## Proposition

*Given a random sample $X_1, \ldots, X_n$ from a population with mean $\mu$ and variance $\sigma^2$.*

*Then $S^2$ is an unbiased estimator of $\sigma^2$.*

<u>PROOF:</u> $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$, $\mathbb{E}[X] = \mu$, $\mathbb{V}[X] = \sigma^2$, $\mathbb{E}[\overline{X}] = \mu$, $\mathbb{V}[\overline{X}] = \sigma^2/n$

$$\begin{aligned}
\mathbb{E}[S^2] &= \mathbb{E}\left[\tfrac{1}{n-1}\sum_{k=1}^n (X_k - \overline{X})^2\right] = \tfrac{1}{n-1}\mathbb{E}\left[\sum_{k=1}^n \left[X_k^2 - 2\overline{X}X_k + (\overline{X})^2\right]\right] \\
&= \tfrac{1}{n-1}\mathbb{E}\left[\sum_{k=1}^n X_k^2 - 2\overline{X}\sum_{k=1}^n X_k + \sum_{k=1}^n (\overline{X})^2\right] \\
&= \tfrac{1}{n-1}\mathbb{E}\left[\sum_{k=1}^n X_k^2 - 2\overline{X}(n\overline{X}) + n(\overline{X})^2\right] = \tfrac{1}{n-1}\mathbb{E}\left[\sum_{k=1}^n X_k^2 - n(\overline{X})^2\right] \\
&= \tfrac{1}{n-1}\left[\sum_{k=1}^n \mathbb{E}[X_k^2] - n\mathbb{E}[\overline{X}^2]\right] = \tfrac{1}{n-1}\left[\sum_{k=1}^n \left[\mathbb{V}[X_k] + (\mathbb{E}[X_k])^2\right] - n\mathbb{E}[\overline{X}^2]\right] \\
&\overset{iid}{=} \tfrac{1}{n-1}\left[n\sigma^2 + n\mu^2 - n\left[\mathbb{V}[\overline{X}] + (\mathbb{E}[\overline{X}])^2\right]\right] \\
&= \tfrac{1}{n-1}\left[n\sigma^2 + n\mu^2 - n\left[\sigma^2/n + \mu^2\right]\right] = \tfrac{1}{n-1}\left[n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2\right] \\
&= \tfrac{1}{n-1}\left[(n-1)\sigma^2\right] \\
&= \sigma^2
\end{aligned}$$

$\therefore\ \mathbb{E}[S^2] = \sigma^2 \implies S^2$ is an unbiased estimator of $\sigma^2$ $\square$

# The "Best" Unbiased Estimator: UMVUE

The "best" unbiased estimator is the one that varies the least:

## Definition

The unbiased estimator of a parameter with the smallest variance is called the **uniformly minimum variance unbiased estimator (UMVUE)**.

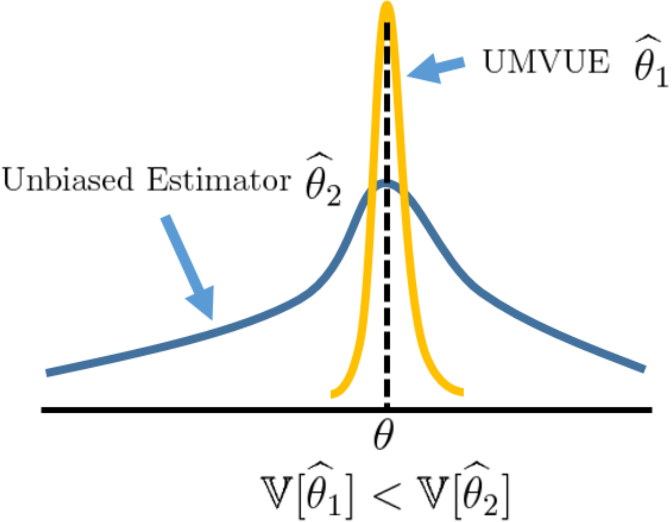
## Theorem

*Let random sample* $X_1, \ldots, X_n \stackrel{iid}{\sim}$ *Normal*$(\mu, \sigma^2)$. *Then* $\overline{X}$ *is the UMVUE for* $\mu$.

PROOF: Beyond the scope of this course.

# UMVUE (Visually)

The two curves are the pdf's of the sampling distributions of the estimators



# The Standard Error of a Point Estimator

The **standard error** of a point estimator is a crucial piece of many Statistical Inference methods:

## Definition

The **standard error** of a point estimator $\widehat{\theta}$ is $\quad \sigma_{\widehat{\theta}} := \sqrt{\mathbb{V}[\,\widehat{\theta}\,]}$

The **estimated standard error** of a point estimator, denoted $\widehat{\sigma}_{\widehat{\theta}}$, is the value of the standard error $\sigma_{\widehat{\theta}}$ when any unknown parameters involved in its expression have themselves been estimated.

Standard Error of $\overline{X}$ is $\quad \sigma_{\overline{X}} = \sigma/\sqrt{n}$

Estimated Standard Error of $\overline{X}$ is $\quad \widehat{\sigma}_{\overline{X}} = \widehat{\sigma}/\sqrt{n} = s/\sqrt{n}$

(Estimated) Standard Errors of the other common point estimators are encountered in the 6.1 Outline, 6.1 Homework, and the textbook.

# Expected Value & Variance Revisited

The following properties of expected & variance of a random sample are useful when finding standard errors:

## Proposition

*Let $X_1, \ldots, X_n$ be a random sample from a population and $c_1, \ldots, c_n \neq 0$. Then:*

*(i)* $\mathbb{E}[c_1X_1 + \cdots + c_nX_n] = c_1\mathbb{E}[X_1] + \cdots + c_n\mathbb{E}[X_n]$
*(ii)* $\mathbb{V}[c_1X_1 + \cdots + c_nX_n] = c_1^2\mathbb{V}[X_1] + \cdots + c_n^2\mathbb{V}[X_n]$
*(iii)* $\mathbb{E}[X_1 - X_2] = \mathbb{E}[X_1] - \mathbb{E}[X_2]$
*(iv)* $\mathbb{V}[X_1 - X_2] = \mathbb{V}[X_1] + \mathbb{V}[X_2]$

PROOF:

(iii) $\mathbb{E}[X_1 - X_2] = \mathbb{E}[X_1 + (-1)X_2] \overset{(i)}{=} \mathbb{E}[X_1] + (-1)\mathbb{E}[X_2] = \mathbb{E}[X_1] - \mathbb{E}[X_2]$

(iv) $\mathbb{V}[X_1 - X_2] = \mathbb{V}[X_1 + (-1)X_2] \overset{(ii)}{=} \mathbb{V}[X_1] + (-1)^2\mathbb{V}[X_2] = \mathbb{V}[X_1] + \mathbb{V}[X_2]$ $\square$

# Textbook Logistics for Section 6.1

- Difference(s) in Notation:

| CONCEPT | TEXTBOOK NOTATION | SLIDES/OUTLINE NOTATION |
|---|---|---|
| Probability of Event | $P(A)$ | $\mathbb{P}(A)$ |
| pmf with parameter | $p_X(x;\theta)$ | $p_X(k;\theta)$ |
| pdf with parameter | $f(x;\theta)$ | $f_X(x;\theta)$ |
| Expected Value of r.v. | $E(X)$ | $\mathbb{E}[X]$ |
| Variance of r.v. | $V(X)$ | $\mathbb{V}[X]$ |
| Bias of Estimator $\widehat{\theta}$ | (none) | Bias$[\ \widehat{\theta}\ ]$ |
| "Best" Unbiased Estimator | MVUE | UMVUE |

# Textbook Logistics for Section 6.1

- Skip the ”Some Complications” section (pgs 257-258)
  - TLDR: The best estimator in one scenario may not be the best in others.
  - Perhaps read the details of this section at the end of this course.

- Ignore the notion of **censoring** [EXAMPLE 6.8 (pg 258)]
  - Censoring means the value of a data point is only partially known
    - e.g. An analog produce scale can only weigh up to 20 lbs of produce, and someone places more than 20 lbs of produce on the scale. All a statistician would know is that particular data point has a value greater than 20 lbs.
    - e.g. When measuring the lifetimes of a random sample of components, sometimes it's preferred to only measure the first couple component failures instead of waiting for all the components in the sample to fail.
  - Censoring is beyond the scope of this course.

- Ignore **bootstrap sampling & estimation** [EXAMPLE 6.11 (pg 260)]
  - Bootstrap Sampling is useful when a point estimator's expression is either impossible to find or far too complicated to use.
  - Bootstrap Sampling is a numerical method in Statistics.
  - Bootstrap Sampling is beyond the scope of this course.

Fin.