Lesson 023
Point Estimation

Wednesday, November 1

General Feedback from Continuing Survey

- **Paraphrasing some themes:**
  - Stats should lift up the scores of engineering students.
  - PollEverywhere would be good for course engagement.
  - Do we really have to do the surveys?

Stop-Start-Continue Feedback

- **Stop:**
  - "[...] the wording on the assignments and the test questions are incredibly vague [...] moving forward, questions should be a little more on the nose with what they are asking for."
  - "[...] it would be easier to study for the exam/ tests moving forward if the material was organized by chapter or topic on d2l rather than lesson number [...]"
  - "The surveys for each class seem slightly unnecessary [...]"

Stop-Start-Continue Feedback

- **Start:**
  - "Faster assignment feedback [...]"
  - "Smaller tests to make up for the midterm [...]"
  - "More detailed practice problems in class similar to the types of questions we would be asked on exams. The problems covered in class are often much simpler than questions we see on assessments."

    "I would like to do more examples on the board each class that are similar to the types of questions that we will see on the exam."

    "If we could do some more examples or questions in class, I feel as though that would be beneficial for many."

Stop-Start-Continue Feedback

- **Continue:**
  - Resubmissions
  - Feedback and openness
  - PollEverywhere
  - Flipped Classroom & Videos

Any other questions, comments, or concerns at this point?

Point Estimation

- Suppose that θ is a population parameter of interest.
  - Mean length, median volume, maximum proportion of defects, etc.
- Using a sample we want to **estimate the value of θ.**
  - We will denote this statistic $\widehat{\theta}$.
  - This will be a single value, based on the underlying sample. We hope it is "good".

Estimates and Estimators

- Once we have a sample and computed the single value we refer to this as a **point estimate**.
  - Sometimes, we will just say "estimate".
- As a function, $\widehat{\theta}$ is referred to as an **estimator**.
- Estimates are single values actually computed on samples.
- Estimators are random variables in the form of a function.

Estimates and Estimators: Example

- If we wish to estimate the population mean, μ, we may decide to use $\widehat{\mu} = \overline{X}$ (the sample mean).
- $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ is an **estimator**. It is a **random variable.**
- If we observe $\{1,2,3,4,5\}$, then $\bar{x} = \dfrac{15}{5} = 3$ is an **estimate.** It is an **observed value**.

Suppose that a sample is taken resulting in $\{1, 1, 2, 2\}$. We are interested in the population mean, μ. Which of the following is true?

$\frac{1+1+2+2}{4} = 1.5$ is an estimator for μ.

0%

$\overline{X}$ is an estimate for μ.

0%

From the sample, we know that $\mu = 1.5$.

0%

The sample median is an estimator for μ.

0%

Estimators

- Because estimators are random, they have some distribution.
  - This is the **sampling distribution** from last time.
- Generally, many different estimators can be found for any parameter θ.
  - Using an estimator we produce an estimate on a sample.
  - Our estimate will depend on the estimator and the sample.

Suppose that a sample of $\{1, 1, 2, 2\}$ is obtained. What is the estimate produced by the sample mean?

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$

0%

$\overline{x} = 1.5$

0%

The sample mean is not an estimator, and therefore does not produce estimates.

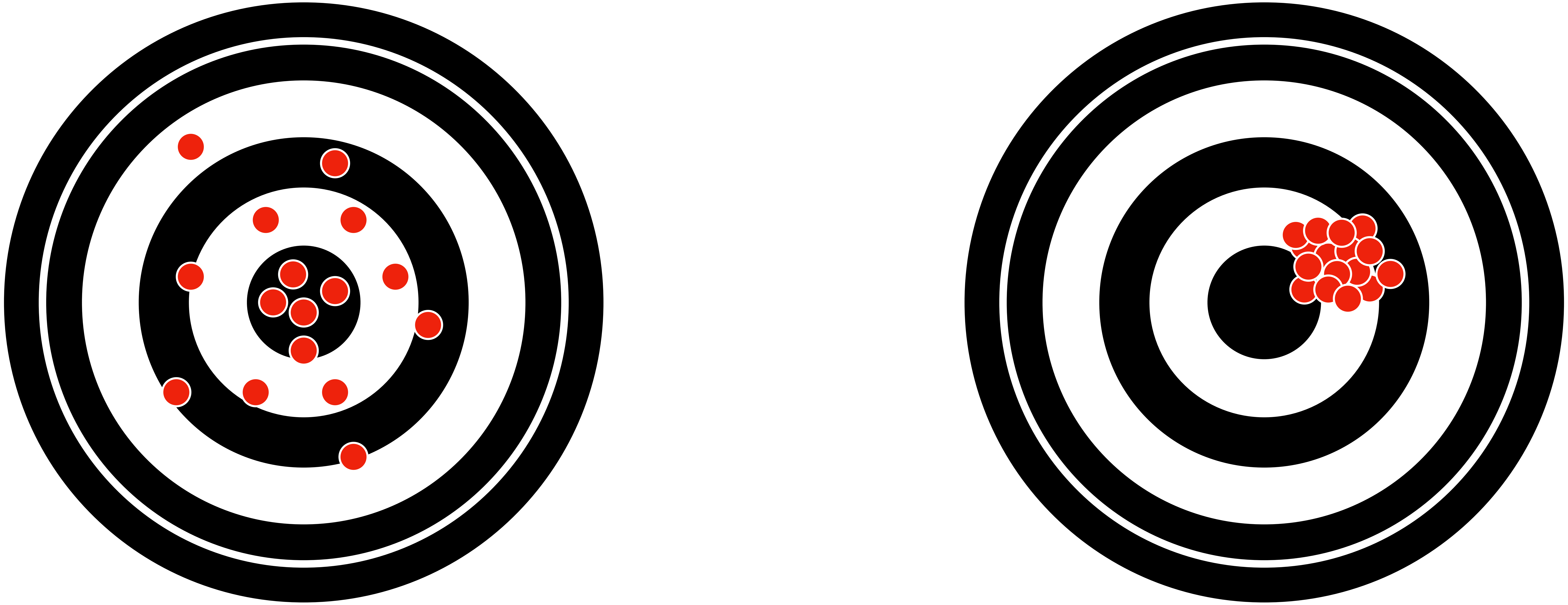
0%

Which of the following are estimators for the population mean if the population follows a Poisson distribution?

The sample mean.	0%
The sample variance.	0%
The fifth observation from a sample.	0%
All of the above.	0%

Assessing Estimators

- We do not want *any* estimator, we want a *good* estimator.
- If we have many options for estimators, we want to select the **best** one.
- Defining best can be a challenging task, in general.



Mean Squared Error (MSE)

- The most common metric to assess the quality of an estimator is the **mean squared error (MSE).**

- If $\widehat{\theta}$ is an estimator for θ, then the MSE is given by:

$$\text{MSE}(\widehat{\theta}) = E\left[\left(\widehat{\theta} - \theta\right)^2\right]$$

- The lower an estimator's MSE, the better it is considered.

Bias-Variance Decomposition

- The MSE is comprised of two separate components: the **bias** and the **variance** of the estimator.
- The bias represents how far, on average, the estimator is from the truth.

$$\text{Bias}(\widehat{\theta}) = E\left[\widehat{\theta} - \theta\right] = E[\widehat{\theta}] - \theta.$$

- The variance we have seen before.

$$\text{var}(\widehat{\theta}) = E\left[(\widehat{\theta} - E[\theta])^2\right]$$

Bias-Variance Decomposition

- We prefer smaller bias and smaller variance, and typically combine them into the MSE

$$\text{MSE}(\widehat{\theta}) = \text{Bias}(\widehat{\theta})^2 + \text{var}(\widehat{\theta})$$

- For some settings we may care more about the bias or more about the variance.
  - The bias and variance are both obtainable via the sampling distribution of $\widehat{\theta}$.

Which of the following estimators would be preferred in general?

An estimator with a 0 bias and a variance of 15.

0%

An estimator with a bias of 9 and a variance of 1.

0%

An estimator with a bias of 5 and a variance of 1.

0%

An estimator with a bias of 2 and a variance of 8.

0%

Bias of an Estimator

- Whenever $E[\widehat{\theta}] = \theta$ then we have Bias$(\widehat{\theta}) = 0$ and we say that $\widehat{\theta}$ is **unbiased**.
- Generally, a smaller bias is preferable.
- Some estimators have bias that depends on the sample size, n, with $\lim_{n \to \infty}$ Bias$(\widehat{\theta}_n) = 0$. These are **asymptotically unbiased**.
- If $\widehat{\theta}$ is unbiased for θ, we will not generally have $g(\widehat{\theta})$ unbiased for $g(\theta)$.

Example of Bias

- We defined the sample standard deviation to be
$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$
- However, $E[S^2] \neq \text{var}(X)$.
- Instead, most people use an **unbiased sample variance estimator**.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Unbiased Sample Variance

- Going forward we will use the **unbiased sample variance estimator**.
- This will correspond to what you will find in statistical software, etc.
- The sample variance will be the square root of this new estimator.
    - Note: the sample variance is **not** unbiased.

Suppose that we wish to estimate the population mean, μ. We consider both the first observation, X_1 and the sample mean $\overline{X}$ as the estimators. Which of the following is true?

$\overline{X}$ is unbiased and X_1 is biased.

0%

$\overline{X}$ is biased and X_1 is unbiased.

0%

$\overline{X}$ is unbiased and X_1 is unbiased.

0%

$\overline{X}$ is biased and X_1 is biased.

0%

Which estimator for the population mean is preferable between X_1 and $\overline{X}$?

$\overline{X}$

0%

X_1

0%

They are both unbiased, so there is no difference.

0%

I am not sure.

0%

Variance of Estimators

- Note that, this previous example illustrates, bias alone is not enough.
- Among estimators with the same bias, the smaller the variance the more **precise** the estimator will be.
- The variance is defined in the same way as for any random variable var$(\widehat{\theta}) = E[(\widehat{\theta} - E[\widehat{\theta}])^2]$.
- When Bias$(\widehat{\theta}) = 0$, then MSE$(\widehat{\theta}) =$ var$(\widehat{\theta})$.

Suppose that we wish to estimate the population mean, μ. We consider both the first observation, X_1 and the sample mean $\overline{X}$ as the estimators. Suppose that $\mathrm{var}(X) = \sigma^2$. Which of the following is true?

$\mathrm{var}(\overline{X}) = \sigma^2$ and $\mathrm{var}(X_1) = \sigma^2$.

0%

$\mathrm{var}(\overline{X}) = \frac{\sigma^2}{n}$ and $\mathrm{var}(X_1) = \sigma^2$.

0%

$\mathrm{var}(\overline{X}) = \sigma^2$ and $\mathrm{var}(X_1) = \frac{\sigma^2}{n}$.

0%

$\mathrm{var}(\overline{X}) = \frac{\sigma^2}{n}$ and $\mathrm{var}(X_1) = \frac{\sigma^2}{n}$.

0%

Why do we prefer $\overline{X}$ as an estimator compared to X_1?

Because $\overline{X}$ is unbiased.

0%

Because the variance of $\overline{X}$ is smaller than that of X_1.

0%

Because the bias of $\overline{X}$ is the same as the bias of X_1.

0%

Because the bias of $\overline{X}$ is the same as the bias of X_1 and the variance of $\overline{X}$ is smaller than that of X_1.

0%

Minimum Variance Estimators

- If we consider only unbiased estimators, then the best estimator is the one with the minimum variance.
- This is called the **minimum variance unbiased estimator (MVUE)**.
- The MVUE will depend on both the parameter being estimated, and on the population distribution.
  - If the population is normally distributed, $\overline{X}$ is MVUE for μ.
  - If the population is binomial, $\overline{X} = \widehat{p}$ is MVUE for p.
  - In general, $\overline{X}$ is **not** MVUE for $E[X]$.

Which of the following estimators could possibly be the MVUE of θ?

$\hat{\theta}_1$ with $\text{Bias}(\hat{\theta}_1) = 0$ and $\text{var}(\hat{\theta}_1) = 5$.

0%

$\hat{\theta}_2$ with $\text{Bias}(\hat{\theta}_2) = 3$ and $\text{var}(\hat{\theta}_2) = 1$.

0%

$\hat{\theta}_3$ with $\text{Bias}(\hat{\theta}_3) = 1$ and $\text{var}(\hat{\theta}_1) = 1$.

0%

$\hat{\theta}_4$ with $\text{Bias}(\hat{\theta}_4) = 0$ and $\text{var}(\hat{\theta}_1) = 4$.

0%

Suppose that $\widehat{\theta}_4$ is the MVUE. Which of the following estimators is preferable, in general?

$\widehat{\theta}_1$ with $\mathrm{Bias}(\widehat{\theta}_1) = 0$ and $\mathrm{var}(\widehat{\theta}_1) = 5$.

0%

$\widehat{\theta}_2$ with $\mathrm{Bias}(\widehat{\theta}_2) = 3$ and $\mathrm{var}(\widehat{\theta}_2) = 1$.

0%

$\widehat{\theta}_3$ with $\mathrm{Bias}(\widehat{\theta}_3) = 1$ and $\mathrm{var}(\widehat{\theta}_1) = 1$.

0%

$\widehat{\theta}_4$ with $\mathrm{Bias}(\widehat{\theta}_4) = 0$ and $\mathrm{var}(\widehat{\theta}_1) = 4$.

0%

Standard Error of an Estimator

- The standard deviation of an estimator is called the **standard error**.
  - We write $SE(\widehat{\theta}) = \sqrt{\text{var}(\widehat{\theta})}$.
- Often the standard error cannot be computed exactly, and so we will rely on an estimated standard error.
- You should **always** report the standard error with a point estimate.