BSTA 551: Statistical Inference

Lesson 2: Point Estimation; Bias, Variance, and MSE

Jessica Minnier

2026-01-07

Lesson 2: Point Estimation

Review: Where We Left Off

Key concepts from Lesson 1:

- Population vs. sample; parameters vs. statistics
- Sampling distributions
- $E(\bar{X}) = \mu$ and $\mathrm{Var}(\bar{X}) = \sigma^2/n$
- Numerical optimization with `optimize()`

Point Estimation: Core Concepts

What is a Point Estimator? (Devore 7.1)

Definitions

- A **parameter** is a fixed (but unknown) characteristic of a population (e.g., μ, σ, p)
- An **estimator** is a rule/formula for calculating an estimate from sample data
- An **estimate** is the actual number you calculate from a specific sample

Key Distinction: Estimator vs. Estimate

Estimator: A random variable (before data is collected)

- $\bar{X}$ is a function of random variables $X_1, \ldots, X_n$
- Has a sampling distribution
- Can calculate $E(\bar{X}), \mathrm{Var}(\bar{X})$

The Sampling Distribution (Revisited)

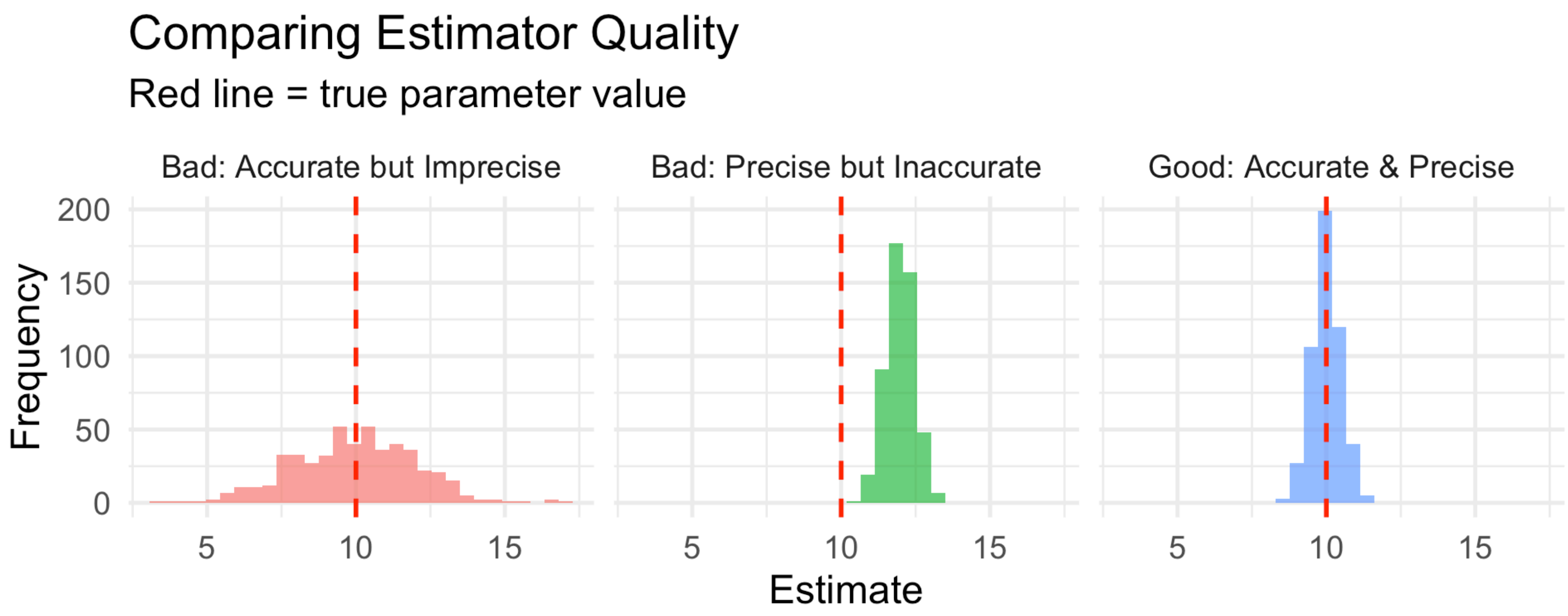
Different samples give different estimates. The **sampling distribution** describes this variability.

Sample	Sample Mean ($\bar{x}$)
1	9.90
2	10.31
3	10.03
4	10.85
5	9.17
6	9.31

What Makes a Good Estimator?

We want estimators that are:

1. **Accurate** (unbiased): On average, hits the true value
2. **Precise** (low variance): Estimates are clustered together
3. **Efficient**: Best combination of accuracy and precision



Bias: Measuring Accuracy

Bias: Formal Definition

Definition

The **bias** of an estimator $\hat{\theta}$ is:

$$\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$$

An estimator is **unbiased** if $\text{Bias}(\hat{\theta}) = 0$, i.e., $E(\hat{\theta}) = \theta$.

Worked Example: Proving Sample Mean is Unbiased

Claim: The sample mean $\bar{X}$ is an unbiased estimator of μ.

Proof: We need to show $E(\bar{X}) = \mu$.

Worked Example: Sample Proportion

Setup: In a vaccine trial, X patients out of n develop immunity. The estimator is $\hat{p} = X/n$.

Claim: $\hat{p}$ is unbiased for the true immunity rate p.

Concrete Calculation: Bias of Sample Proportion

Data: In a study of 80 patients, 52 showed improvement.

```r
n <- 80
x <- 52
p_hat <- x / n

cat("Sample proportion:", p_hat, "\n")
```

```
Sample proportion: 0.65
```

```r
cat("If true p = 0.65, what is the bias of this single estimate?\n")
```

```
If true p = 0.65, what is the bias of this single estimate?
```

```r
cat("Observed - True =", p_hat - 0.65)
```

```
Observed - True = 0
```

Simulation: Verifying Unbiasedness

```r
# Verify that sample proportion is unbiased
true_p <- 0.65
n_patients <- 80
n_simulations <- 10000

proportion_simulation <- tibble(sim = 1:n_simulations) |>
  mutate(
    successes = rbinom(n_simulations, size = n_patients, prob = true_p),
    p_hat = successes / n_patients
  )

proportion_simulation |>
  summarize(
    true_p = true_p,
    mean_of_estimates = mean(p_hat),
    empirical_bias = mean(p_hat) - true_p
  )
```

```
# A tibble: 1 × 3
  true_p mean_of_estimates empirical_bias
   <dbl>             <dbl>          <dbl>
1   0.65             0.650      -0.000183
```

The bias is essentially zero (just simulation noise)!

A Biased Estimator: The Maximum

Problem: Estimate the upper bound θ of a Uniform$[0, \theta]$ distribution.

Natural idea: Use the largest observation: $\hat{\theta} = \max(X_1, \ldots, X_n)$

Calculating the Bias Mathematically

For $X_1, \ldots, X_n \sim \text{Uniform}[0, \theta]$, it can be shown that:

$$E(\max(X_1, \ldots, X_n)) = \frac{n}{n+1}\theta$$

Your Turn: Calculate Bias

Exercise: A lab instrument has a maximum detection limit θ. We take $n = 9$ measurements from Uniform[0, θ] and use the maximum as our estimate.

1. If $\theta = 100$, what is $E(\hat{\theta})$?
2. What is the bias?
3. By what percentage does this estimator underestimate on average?

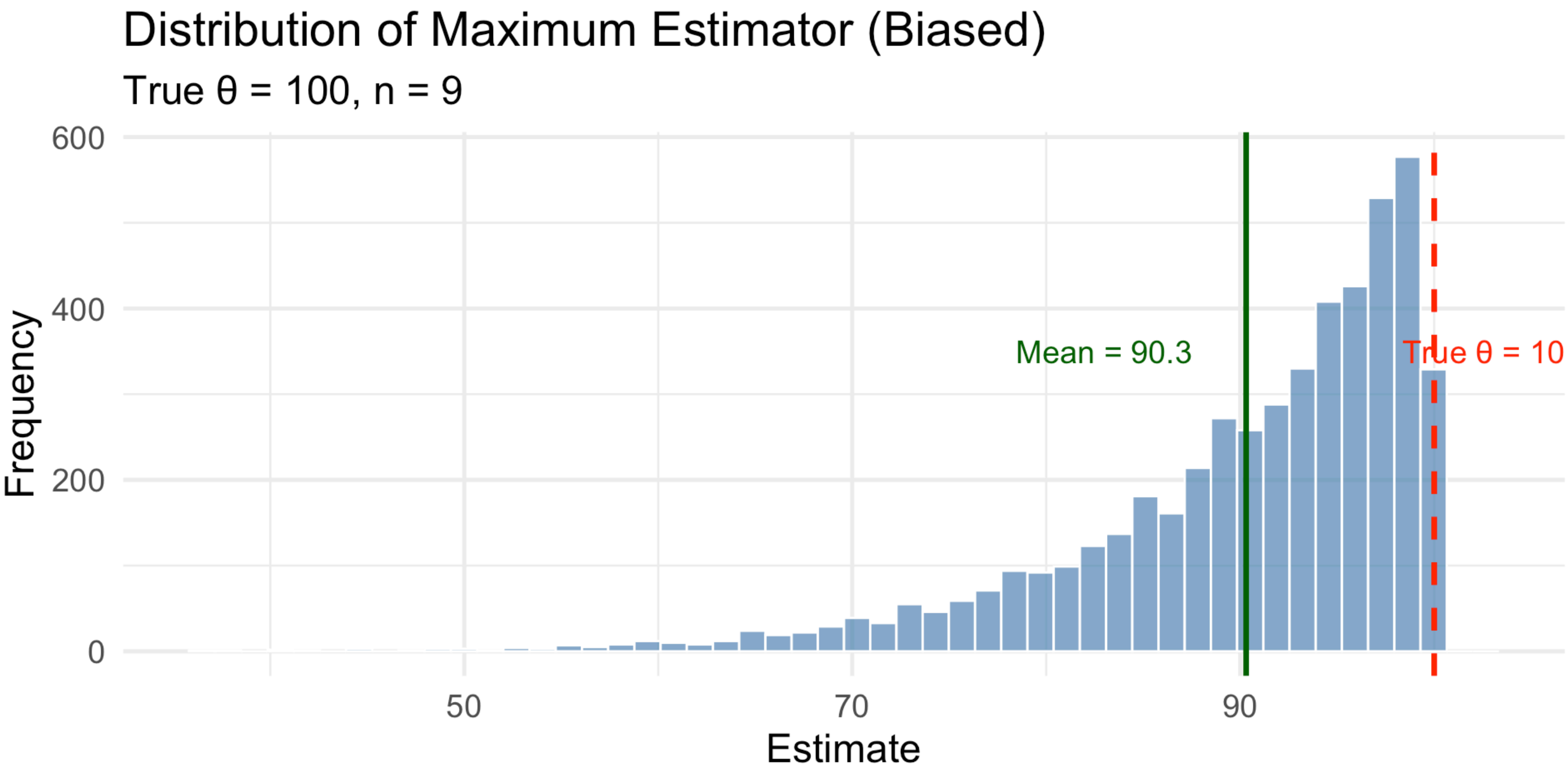
Simulation: Visualizing the Biased Estimator

```r
true_theta <- 100
n <- 9
n_sims <- 5000

max_simulation <- tibble(sim = 1:n_sims) |>
  mutate(
    max_estimate = map_dbl(sim, \(s) max(runif(n, 0, true_theta)))
  )

# Calculate empirical bias
max_simulation |>
  summarize(
    theoretical_E = n / (n + 1) * true_theta,
    empirical_mean = mean(max_estimate),
    theoretical_bias = -true_theta / (n + 1),
    empirical_bias = mean(max_estimate) - true_theta
  )
```

Simulation: Visualizing the Biased Estimator

```
# A tibble: 1 × 4
  theoretical_E empirical_mean theoretical_bias empirical_bias
          <dbl>          <dbl>            <dbl>          <dbl>
1            90           90.3              -10          -9.70
```

Visualizing the Biased Estimator



Correcting the Bias

Idea: Multiply by a correction factor to "un-bias" the estimator.

Since $E(\max) = \frac{n}{n+1}\theta$, we can define:

$$\hat{\theta}_{\text{unbiased}} = \frac{n+1}{n} \cdot \max(X_1, \ldots, X_n)$$

Your Turn: Apply the Correction

Exercise: Using the lab instrument example with $n = 9$ and $\theta = 100$:

1. If you observe $\max = 92$, what is the biased estimate?
2. What is the unbiased estimate?

Standard Error: Measuring Precision

Standard Error: Definition

Definition

The **standard error** of an estimator is its standard deviation:

$$SE(\hat{\theta}) = \sqrt{\mathrm{Var}(\hat{\theta})}$$

Worked Example: Standard Error of Sample Mean

Problem: In a blood pressure study, the population SD is $\sigma = 15$ mmHg. Calculate the standard error of $\bar{X}$ for sample sizes $n = 25$ and $n = 100$.

Your Turn: Calculate Standard Error

Exercise: A survey measures patient satisfaction on a 0-100 scale. The population standard deviation is $\sigma = 20$.

1. What is the SE of $\bar{X}$ for $n = 16$ patients?
2. What sample size is needed to achieve $SE = 2$?

The Problem with Unknown Parameters

Issue: Standard errors often involve unknown parameters!

- SE of $\bar{X}$ requires knowing σ
- SE of $\hat{p}$ requires knowing p

Example: Estimated Standard Error

```r
# Blood pressure data from 25 patients
bp_reductions <- c(12, 8, 15, 10, 7, 14, 11, 9, 13, 16,
                   8, 12, 10, 14, 11, 9, 15, 13, 7, 12,
                   10, 8, 14, 11, 13)

n <- length(bp_reductions)
x_bar <- mean(bp_reductions)
s <- sd(bp_reductions)

# Estimated standard error
se_estimated <- s / sqrt(n)

tibble(
  Statistic = c("Sample Mean", "Sample SD", "Sample Size", "Estimated SE"),
  Value = c(x_bar, round(s, 2), n, round(se_estimated, 2))
)
```

```
# A tibble: 4 × 2
  Statistic    Value
  <chr>        <dbl>
1 Sample Mean  11.3
2 Sample SD     2.64
3 Sample Size  25
4 Estimated SE  0.53
```

Mean Squared Error

Mean Squared Error: Combining Bias and Variance

What if we have to choose between a biased estimator with low variance and an unbiased estimator with high variance?

Proving the MSE Formula

$$\mathrm{MSE}(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

Worked Example: Computing MSE

Setup: Estimating θ from Uniform[0, θ] with $n = 9$ and $\theta = 100$.

Biased estimator: $\hat{\theta}_b = \max(X_i)$

From theory:

- $E(\hat{\theta}_b) = \frac{9}{10}(100) = 90$
- $\mathrm{Var}(\hat{\theta}_b) = \frac{n\theta^2}{(n+1)^2(n+2)} = \frac{9 \times 100^2}{100 \times 11} = 81.82$

Your Turn: Calculate MSE

Exercise: For the **unbiased** estimator $\hat{\theta}_u = \frac{n+1}{n} \max(X_i)$ with $n = 9$ and $\theta = 100$:

1. What is the bias?
2. If $\mathrm{Var}(\hat{\theta}_u) = 101.01$, what is the MSE?
3. Which estimator has lower MSE: biased or unbiased?

The Bias-Variance Tradeoff

The Bias-Variance Tradeoff

Sometimes a **biased** estimator has **lower MSE** than an unbiased one!

Comparing Proportion Estimators: Theory

For $\hat{p}_1 = X/n$ (standard):

- Bias = 0
- Variance = $\frac{p(1-p)}{n}$
- MSE = $\frac{p(1-p)}{n}$

Your Turn: Calculate Bias of Add-Two Estimator

Exercise: For $n = 20$ and $p = 0.3$:

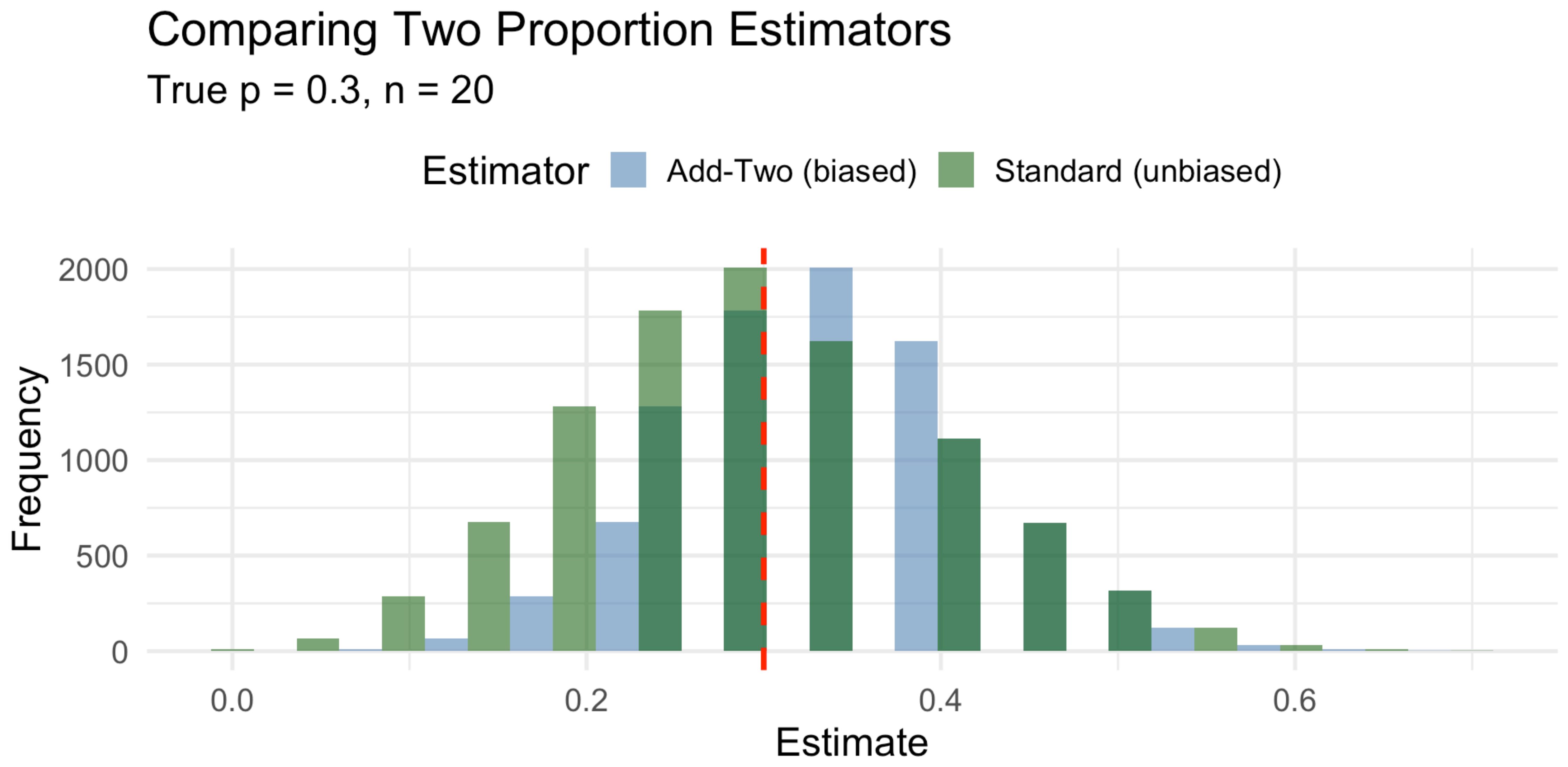
1. Calculate the bias of $\hat{p}_2 = \frac{X+2}{n+4}$

Hint: $E(X) = np$ for binomial, so $E(\hat{p}_2) = \frac{np+2}{n+4}$

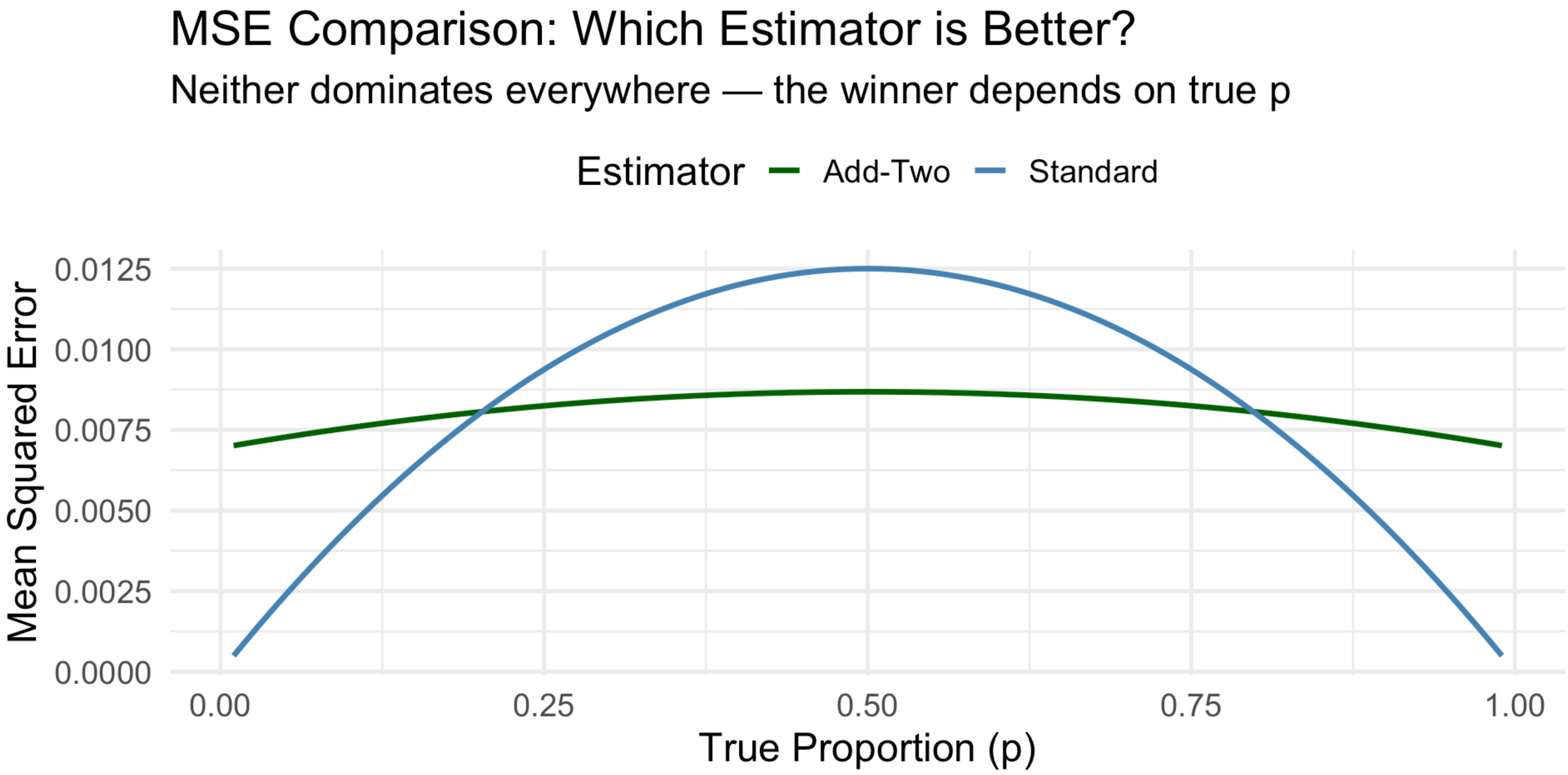
Simulation: Comparing the Estimators

```r
true_p <- 0.3
n <- 20
n_sims <- 10000

# Simulate both estimators
comparison_sim <- tibble(sim = 1:n_sims) |>
  mutate(
    x = rbinom(n_sims, size = n, prob = true_p),
    p_hat_standard = x / n,
    p_hat_addtwo = (x + 2) / (n + 4)
  )

# Compare MSE
comparison_sim |>
  summarize(
    `Standard Bias` = mean(p_hat_standard) - true_p,
    `Add-Two Bias` = mean(p_hat_addtwo) - true_p,
    `Standard Variance` = var(p_hat_standard),
    `Add-Two Variance` = var(p_hat_addtwo),
    `Standard MSE` = mean((p_hat_standard - true_p)^2),
    `Add-Two MSE` = mean((p_hat_addtwo - true_p)^2)
  ) |>
  pivot_longer(everything(), names_to = "Metric", values_to = "Value") |>
```

Visualizing the Tradeoff



MSE Comparison Across Different True Values



Key Insight: The “best” estimator depends on the true parameter value!

Medical Application: Disease Prevalence

Scenario: Estimating prevalence of a rare disease ($p \approx 0.05$) vs. a common condition ($p \approx 0.5$).

```r
# Compare MSE at different prevalence levels
n <- 50

mse_at_p <- function(p, n) {
  mse_std <- p * (1-p) / n
  bias_add2 <- (2 - 4*p) / (n + 4)
  var_add2 <- n * p * (1-p) / (n + 4)^2
  mse_add2 <- var_add2 + bias_add2^2

  tibble(p = p, MSE_Standard = mse_std, MSE_AddTwo = mse_add2,
         Better = ifelse(mse_std < mse_add2, "Standard", "Add-Two"))
}

bind_rows(
  mse_at_p(0.05, n),
  mse_at_p(0.50, n)
) |>
  mutate(across(where(is.numeric), \(x) round(x, 5)))
```

```
# A tibble: 2 × 4
      p MSE_Standard MSE_AddTwo Better
  <dbl>        <dbl>      <dbl> <chr>
```

Sample Variance: Why n-1?

Two formulas for sample variance:

$$S^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} \quad \text{vs.} \quad \tilde{S}^2 = \frac{\sum(X_i - \bar{X})^2}{n}$$

Simulation: Comparing Variance Estimators

```r
true_variance <- 100  # σ² = 100
n <- 10
n_sims <- 10000

variance_sim <- tibble(sim = 1:n_sims) |>
  mutate(
    sample_data = map(sim, \(s) rnorm(n, 0, sqrt(true_variance))),
    s2_n_minus_1 = map_dbl(sample_data, var),
    s2_n = map_dbl(sample_data, \(x) sum((x - mean(x))^2) / n)
  )

variance_sim |>
  summarize(
    `True σ²` = true_variance,
    `E[S² with n-1]` = mean(s2_n_minus_1),
    `E[S² with n]` = mean(s2_n),
    `Bias (n-1)` = mean(s2_n_minus_1) - true_variance,
    `Bias (n)` = mean(s2_n) - true_variance
  ) |>
  mutate(across(where(is.numeric), \(x) round(x, 2)))
```

```
# A tibble: 1 × 5
  `True σ²` `E[S² with n-1]` `E[S² with n]` `Bias (n-1)` `Bias (n)`
```

Your Turn: Comprehensive Example

Exercise: A clinical trial measures cholesterol reduction. Based on $n = 36$ patients:

- Sample mean: $\bar{x} = 25$ mg/dL
- Sample SD: $s = 12$ mg/dL

Calculate:

1. The estimated standard error of $\bar{X}$
2. If the true mean reduction is $\mu = 24$, and we repeated this trial many times, what would be the expected MSE of $\bar{X}$?

Summary and Looking Ahead

Putting It All Together: Estimator Summary

Property	Formula	Interpretation
Bias	$E(\hat{\theta}) - \theta$	Systematic error
Variance	$E[(\hat{\theta} - E(\hat{\theta}))^2]$	Random variability
Std Error	$\sqrt{\mathrm{Var}(\hat{\theta})}$	Typical deviation
MSE	$\mathrm{Var} + \mathrm{Bias}^2$	Total error

Lesson 2 Summary

Key Concepts:

1. **Point Estimators:** Rules for calculating estimates from data
   - Estimator = random variable; estimate = specific value
2. **Bias:** $\mathrm{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$
   - Unbiased if $E(\hat{\theta}) = \theta$
   - Can sometimes correct biased estimators
3. **Standard Error:** $SE(\hat{\theta}) = \sqrt{\mathrm{Var}(\hat{\theta})}$
   - Measures precision of the estimator
4. **Mean Squared Error:** $\mathrm{MSE} = \mathrm{Var} + \mathrm{Bias}^2$
   - Captures total estimation error
   - Bias-variance tradeoff: sometimes biased is better!

Lesson 2 Practice Problems

1. For a Uniform[0, θ] distribution with $n = 20$ observations and $\theta = 50$, calculate:
   - The expected value of the maximum
   - The bias of using the maximum as an estimator
   - The corrected unbiased estimator
2. Using simulation, compare the MSE of the standard proportion estimator vs. the add-two estimator for $n = 10$ and $p = 0.1, 0.3, 0.5$.
3. For a sample of size n from Exponential(λ), the MLE is $\hat{\lambda} = 1/\bar{X}$. It can be shown that $E(\hat{\lambda}) = \frac{n}{n-1}\lambda$. Calculate the bias and propose an unbiased estimator.

Next Week Preview

Week 2: Minimum Variance Unbiased Estimators

- Among all unbiased estimators, which has smallest variance?
- The Cramér-Rao lower bound
- Efficiency of estimators
- Introduction to Maximum Likelihood Estimation

References

- Devore, Berk, and Carlton. *Modern Mathematical Statistics with Applications* (Springer). Chapter 7.1
- Chihara and Hesterberg. *Mathematical Statistics with Resampling and R* (Wiley). Chapter 6.

Questions?

Thank you!

```
decktape docs/lessons/02_Point_Estimation/02_Point_Estimation.html lessons/02_Point_Estimat
```