

Power analysis via simulations

Readings for today

- Beaujean, A. A. (2014). Sample size determination for regression models using Monte Carlo methods in R. *Practical Assessment, Research, and Evaluation*, 19(1), 12.

Topics

1. Power

2. Monte Carlo methods

Power

Estimating data sufficiency

Q: Is my data sufficient to address my hypothesis?

Statistical Power: The ability to correctly reject a false H_0

n : sample size

α : Type-I error (false positive) rate

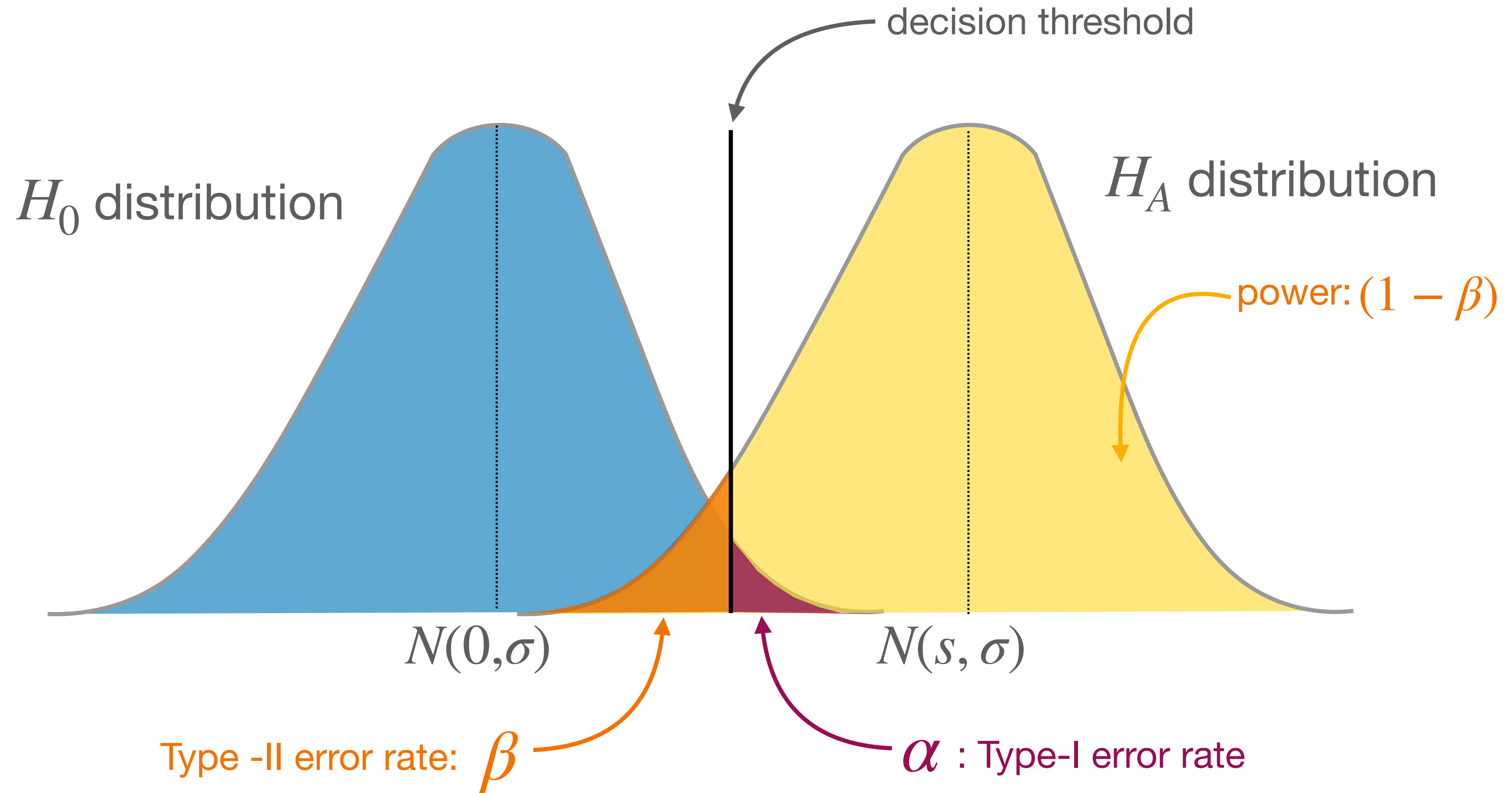
β : Type-II error (false negative) rate

s : effect size

p threshold

Power = $1 - \beta$

Statistical power



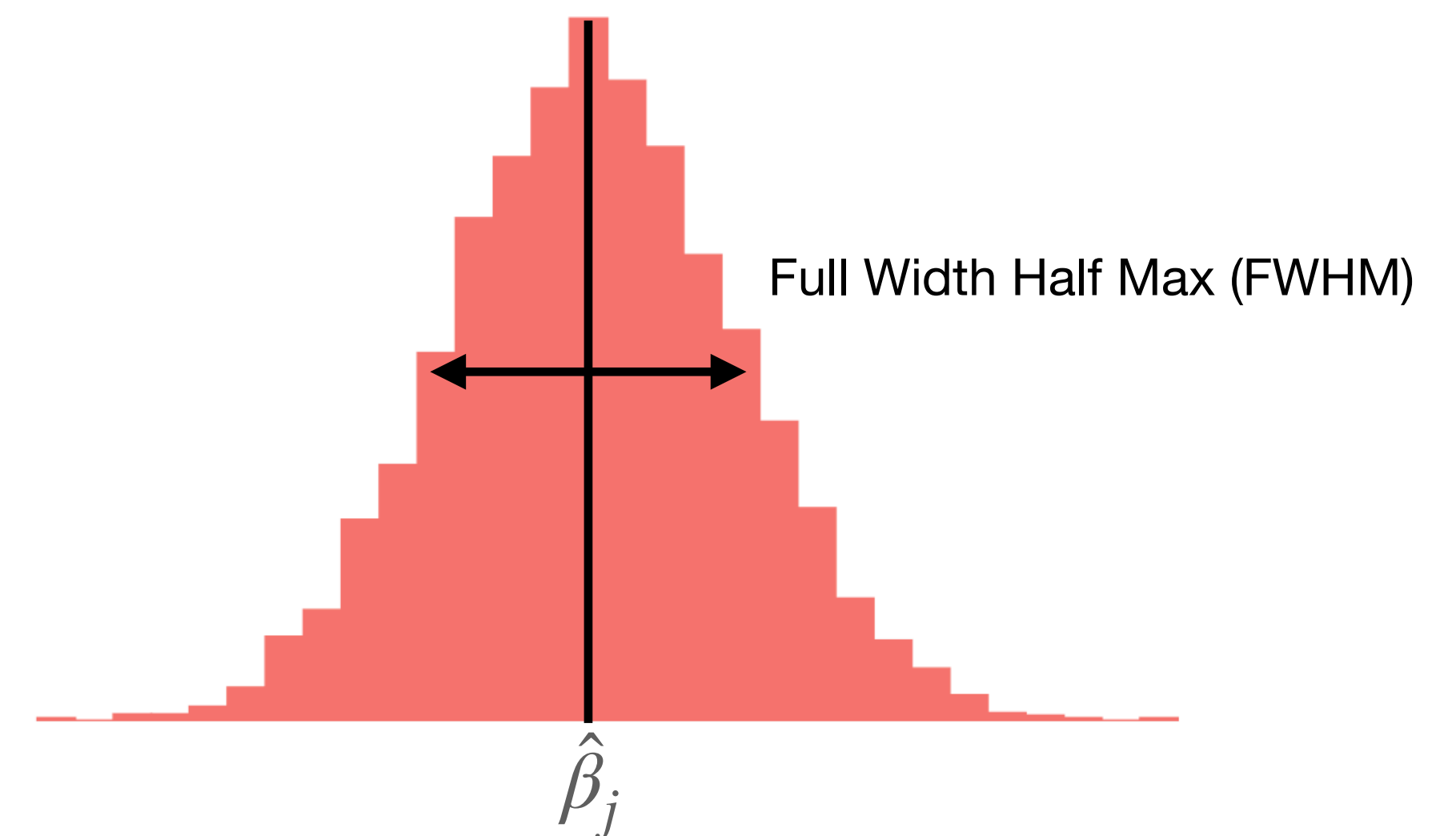
Goals of power analysis

Null Hypothesis Test Statistics (NHTS): Do you have sufficient power to detect whether a test statistic is “statistically significant”?

↪ $P(\hat{H}_0 \text{ rejected} \mid H_0 = \text{false})$

Accuracy in Parameter Estimation (AIPE) Do you have sufficient power to accurately recover the value of a statistical parameter (regardless of “statistical significance”)?

↪ $MSE(\hat{\beta}_j - \beta_j)$



Traditional approaches to estimating accuracy

Goal: Parameter estimation accuracy (\hat{s}) determination.

Step 1: Determine the distributions of your X and Y .

Step 2: Review the literature to find the best estimates of \hat{s} (i.e., $\hat{f}(X)$).

Step 3: Determine your best expected model $E[\hat{f}(X)]$.

Step 4: Set desired confidence interval range for a given Type-I error rate (α)

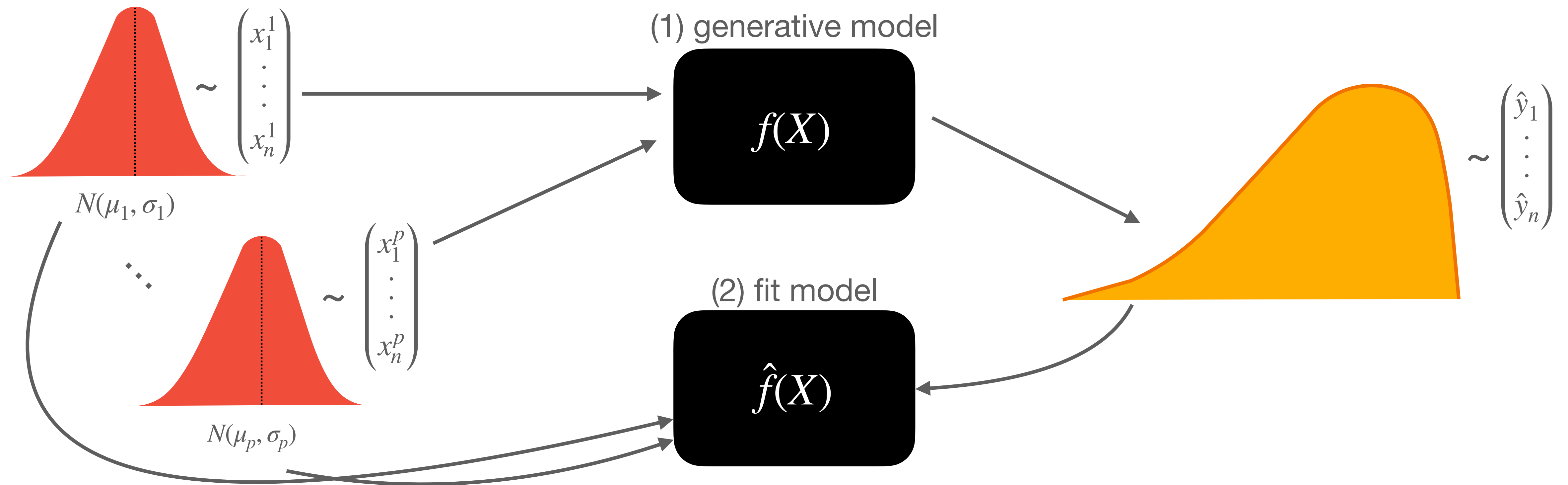
Step 5: Use *parametric* methods to determine the sample size needed to achieve your target $(1 - \alpha) \%$ confidence intervals.

Monte Carlo methods

Monte Carlo (MC) method

Goal: Generate *empirical* distributions of your expected effects, instead of estimated distributions, to determine sufficiency of your data set.

Monte Carlo method: Use random sampling to simple data for numerical analyses.



MC power analysis

Step 1: Determine model of interest (e.g., regression, t-test)

Step 2: Decide on values for all parameters.

- Sample distributions: $X_p \sim N(\mu_p, \sigma_p)$, $Y \sim N(\mu_y, \sigma_y)$
- Effect sizes: e.g., β_j

Step 3: Identify data “quirks” (e.g., missing value rate)

Step 4: Decide on aspects of MC simulations.

- $\alpha \rightarrow$ Type-I error rate
- $(1 - \beta) \rightarrow$ power
- $m \rightarrow$ number of iterations/simulations
- $n \rightarrow$ sample size
- $c \rightarrow$ number of different random seeds (at least 2)

MC power analysis

Step 5: Run m simulations using the parameters in Step 2.

Step 6: Evaluate performance of each simulation using quality metrics.

- If poor, $\uparrow n$ and repeat Step 5.

Step 7: Repeat Steps 5-6 using different random seeds. Check to see if results converge across seeds.

MC power analysis algorithm

For $i = 1$ to m

i) Sample: e.g., $\underbrace{X_1 \sim N(\mu_1, \sigma_1), \dots, X_j \sim B(n, p)}_{\text{normal}} \quad \underbrace{\hspace{10em}}_{\text{binomial}}$

ii) Sim: $Y = f(x, \hat{\theta}) \rightarrow \beta_0 + \sum_{j=1}^p \beta_j X_j + \epsilon$

iii) Evaluate: $T(y, x) \rightarrow \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$

iv) Track: $T = [T_1(Y_1, X_1), \dots, T_i(Y_i, X_i)]$

MC simulation quality metrics

Relative parameter
estimate bias:

$$\theta_{bias} = \frac{\hat{\theta} - \theta_H}{\theta_H}$$

- θ_H : pre-set (real) effect size
- $\hat{\theta}$: average fit parameter across simulations.

Relative standard
error bias:

$$\sigma_{bias} = \frac{\hat{\sigma}_\theta - \sigma_{\hat{\theta}}}{\sigma_{\hat{\theta}}}$$

- $\hat{\sigma}_\theta$: average of fit parameter SE across simulations.
- $\sigma_{\hat{\theta}}$: standard deviation of $\hat{\theta}$ across simulations.

Coverage:

$$c = \frac{1}{m} I(\theta \neq 0)$$

$$I(\theta \neq 0) = \begin{cases} 1, & \text{if } H_0 = \text{rejected} \\ 0, & \text{otherwise} \end{cases}$$

Example: Regression

Model:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

↪ $Y^* \sim N(\beta_0, \sigma_y)$

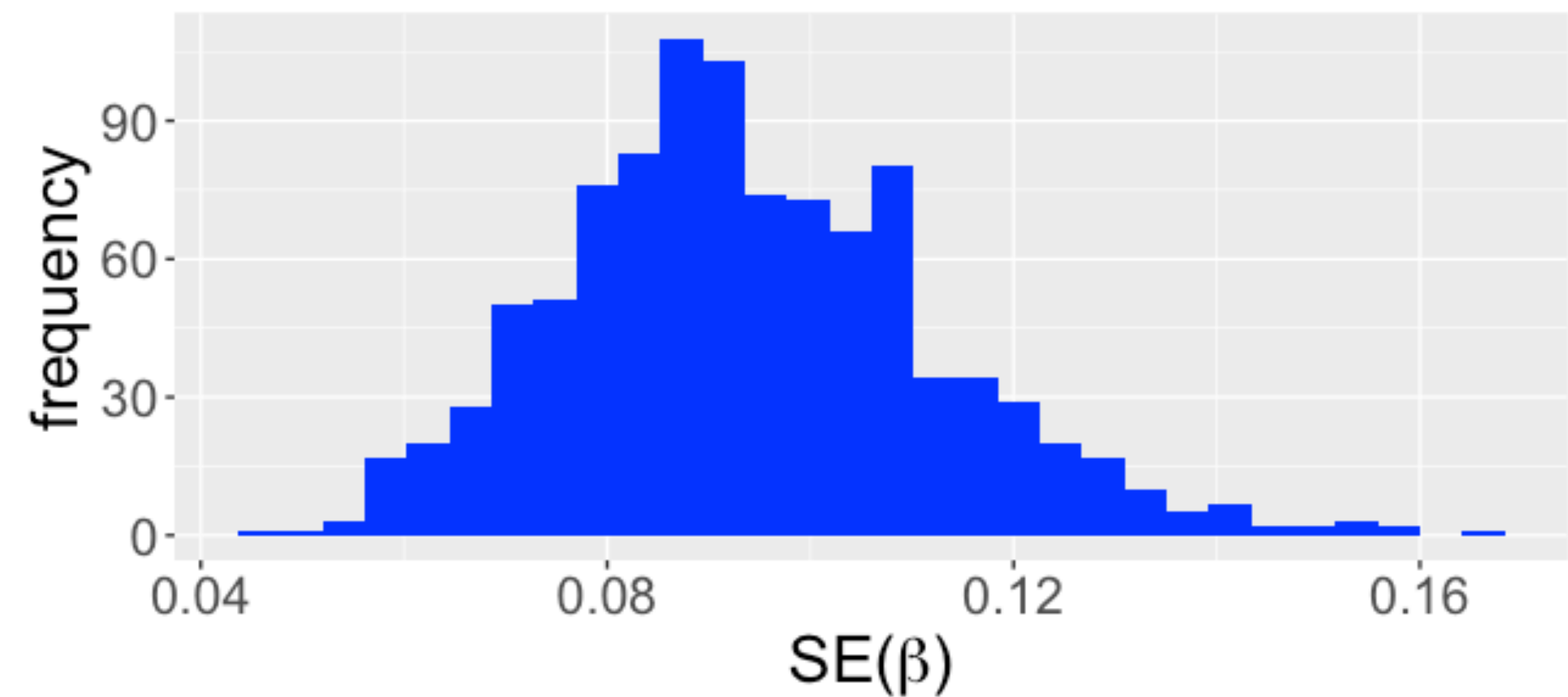
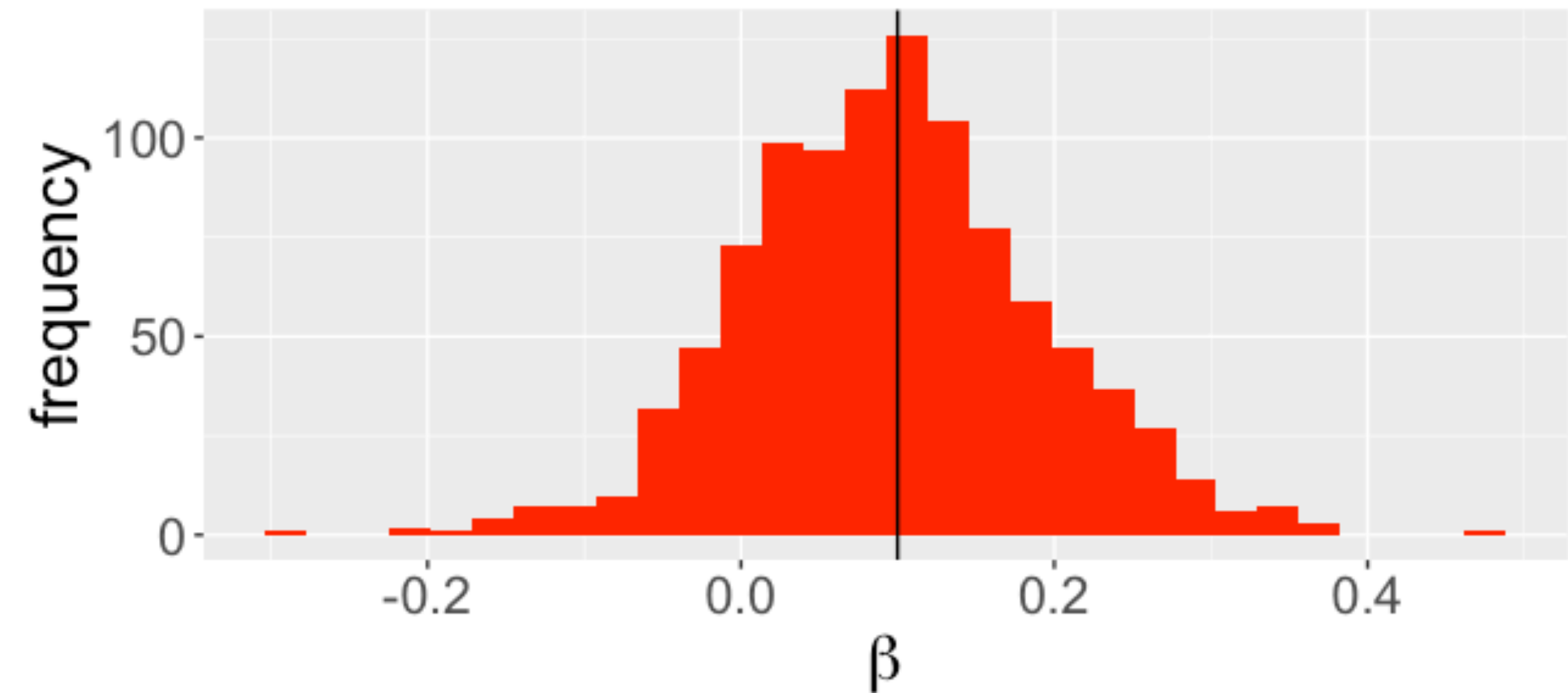
Parameters:

$$\alpha = 0.05 \quad m = 1000$$

$$X \sim N(3, 2) \quad n = 30$$

$$\beta_0 = 9 \quad \beta_1 = 0.1$$

$$\epsilon \sim N(0, 1) = N(0, \sigma_y)$$



Example: Regression

Relative parameter
estimate bias:

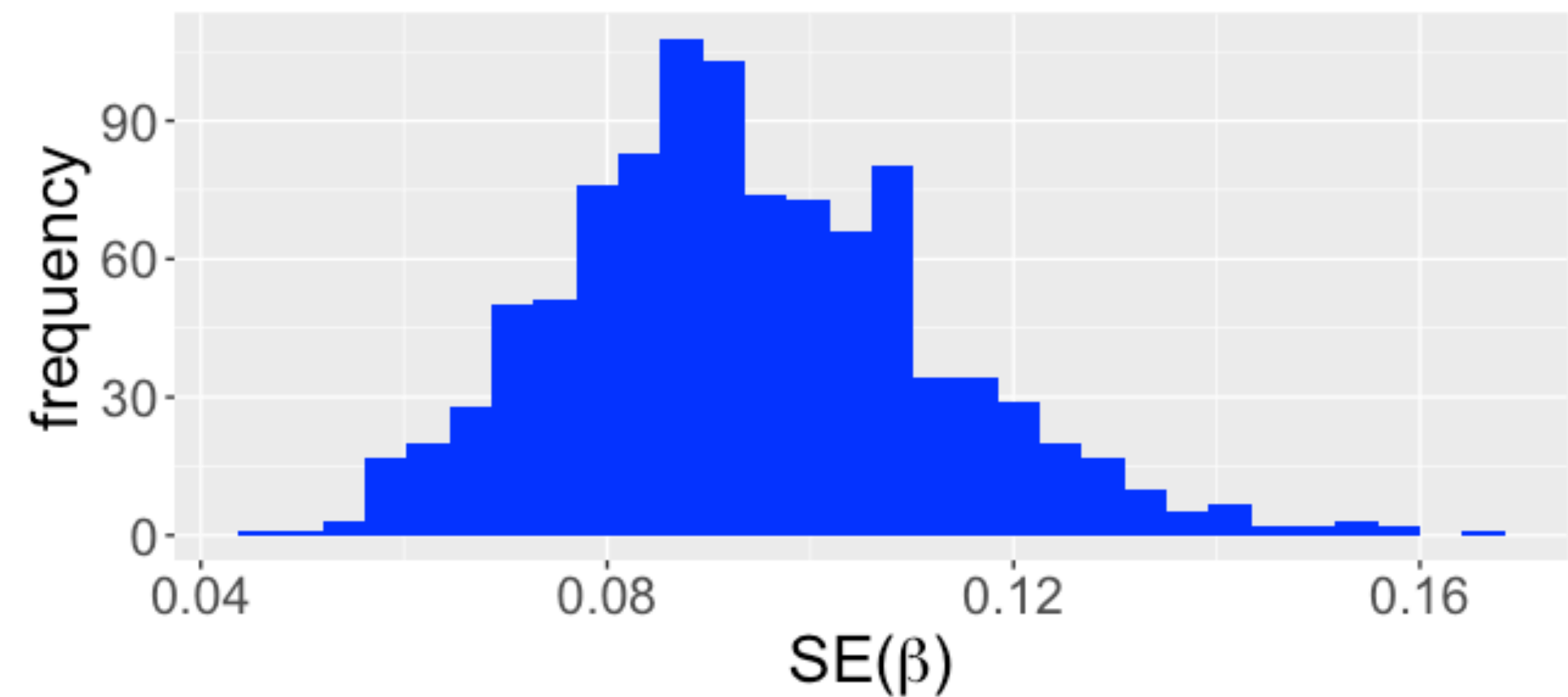
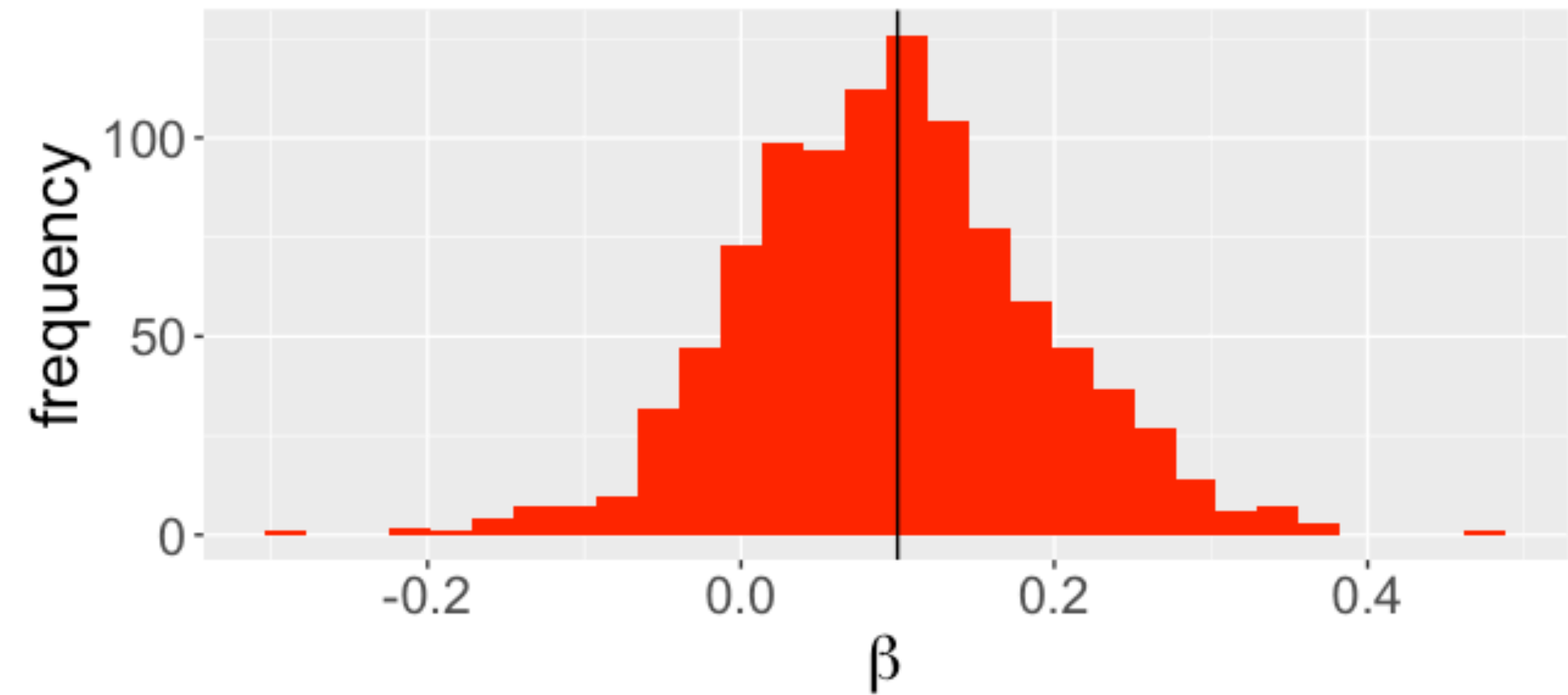
$$\theta_{bias} = \frac{\hat{\theta} - \theta_H}{\theta_H} = -0.045$$

Relative standard
error bias:

$$\sigma_{bias} = \frac{\hat{\sigma}_\theta - \sigma_{\hat{\theta}}}{\sigma_{\hat{\theta}}} = -0.010$$

Coverage:

$$c = \frac{1}{m} I(\theta \neq 0) = 0.173$$



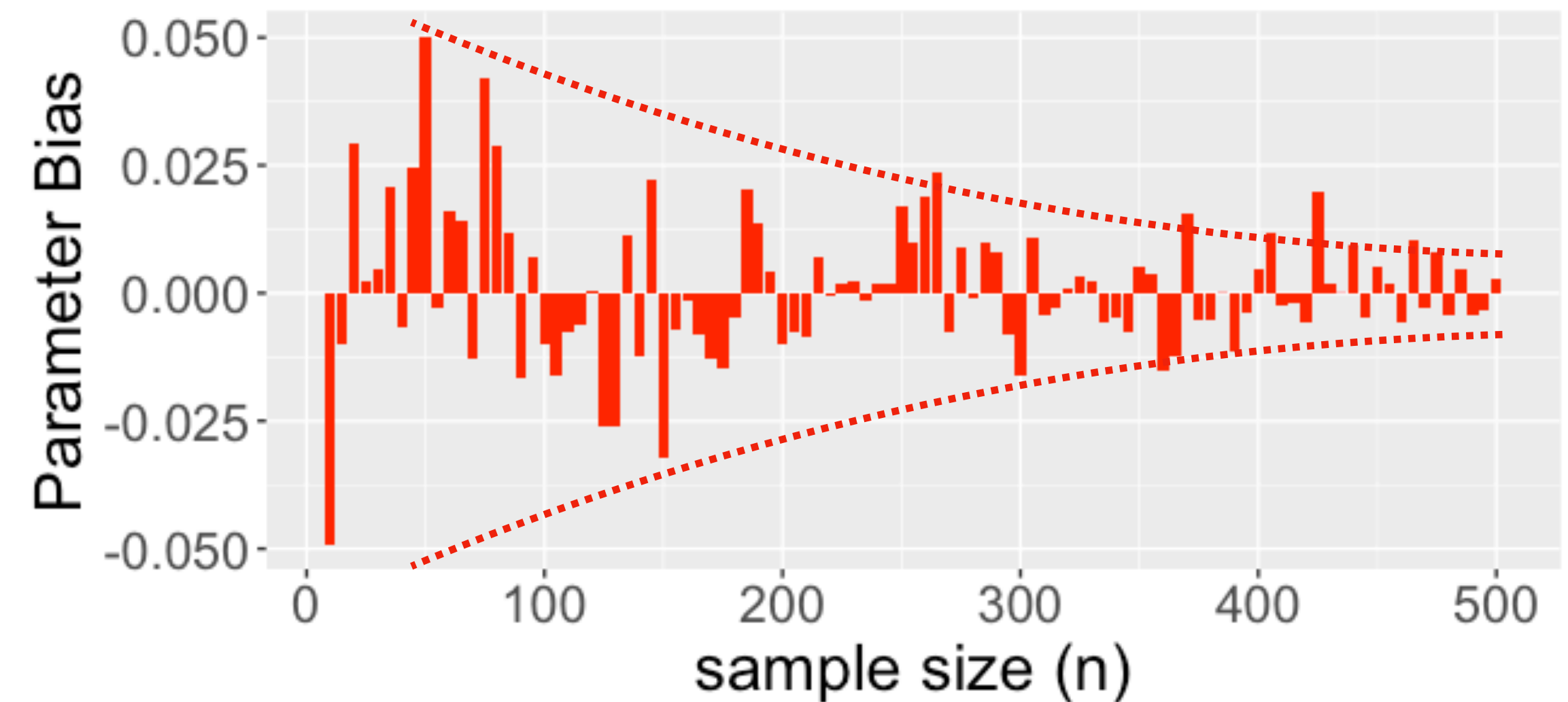
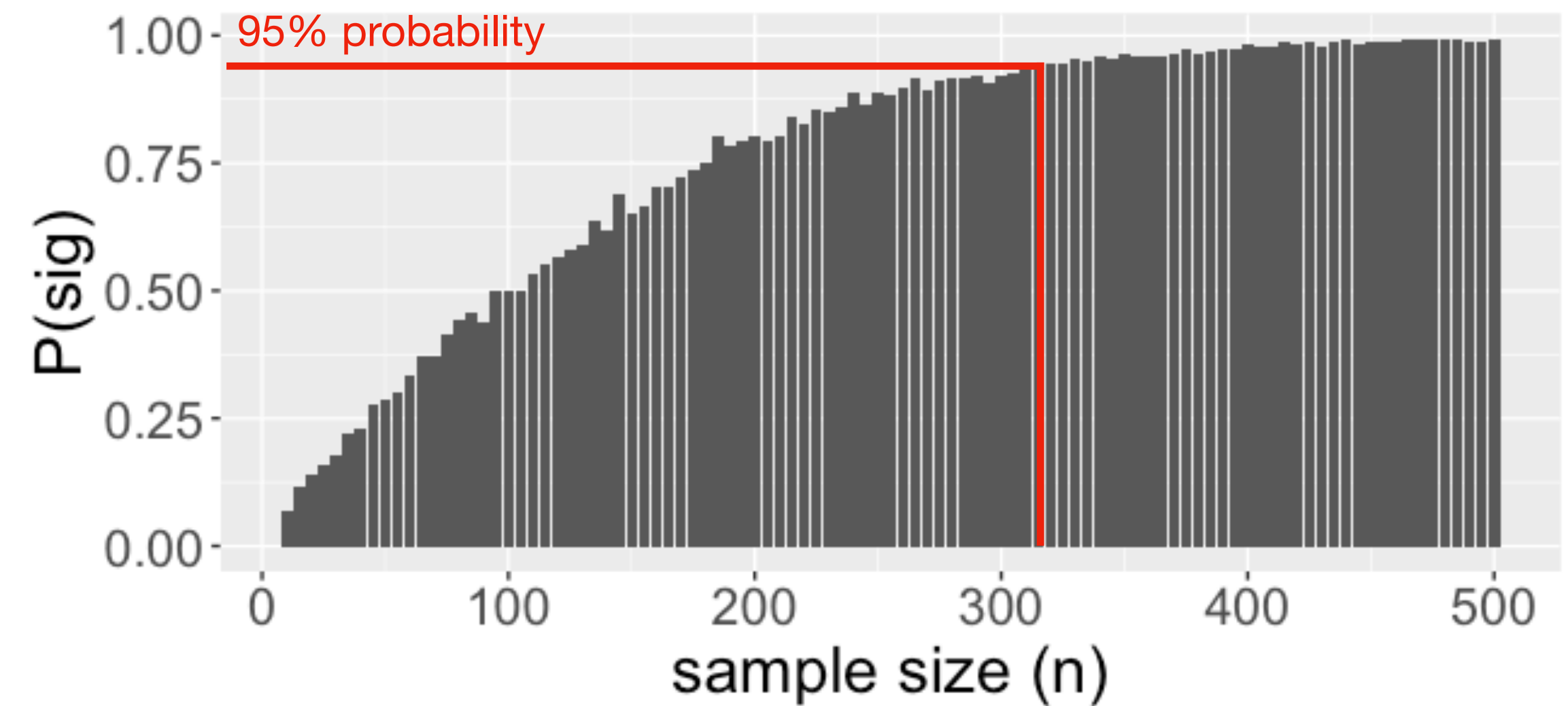
Example: power & accuracy

Goal: Determine the sample size (n) needed in order to:

1. Detect a significant effect of β_1 95% of the time.
2. Have a parameter bias $< |0.01|$

Approach: Repeat model sims from before, across a range of sample sizes (n).

$$n = \underbrace{10 - 500}_{\text{range}}$$



Take home message

- Because they rely on relatively few assumptions, Monte Carlo methods provide a robust and flexible way of estimating the statistical power of your model.