HYPOTHESIS TESTING

Frances Chumney, PhD

CONTENT OUTLINE

- Logic of Hypothesis Testing
- Error & Alpha
- Hypothesis Tests
- Effect Size
- Statistical Power

how we conceptualize hypotheses

LOGIC OF HYPOTHESIS TESTING

HYPOTHESIS TESTING LOGIC

- **Hypothesis Test**

  statistical method that uses sample data to evaluate a hypothesis about a population

- The Logic
  - State a hypothesis about a population, usually concerning a population parameter
  - Predict characteristics of a sample
  - Obtain a random sample from the population
  - Compare obtained data to prediction to see if they are consistent

STEPS IN HYPOTHESIS TESTING

- Step 1: State the Hypotheses
  - **Null Hypothesis (H_0)**

    in the general population there is no change, no difference, or no relationship; the independent variable will have no effect on the dependent variable

    - Example
      - All dogs have four legs.
      - There is no difference in the number of legs dogs have.
  - **Alternative Hypothesis (H_1)**

    in the general population there is a change, a difference, or a relationship; the independent variable will have an effect on the dependent variable

    - Example
      - 20% of dogs have only three legs.

STEP 1: STATE THE HYPOTHESES (EXAMPLE)

- Example

How to Ace a

Statistics Exam

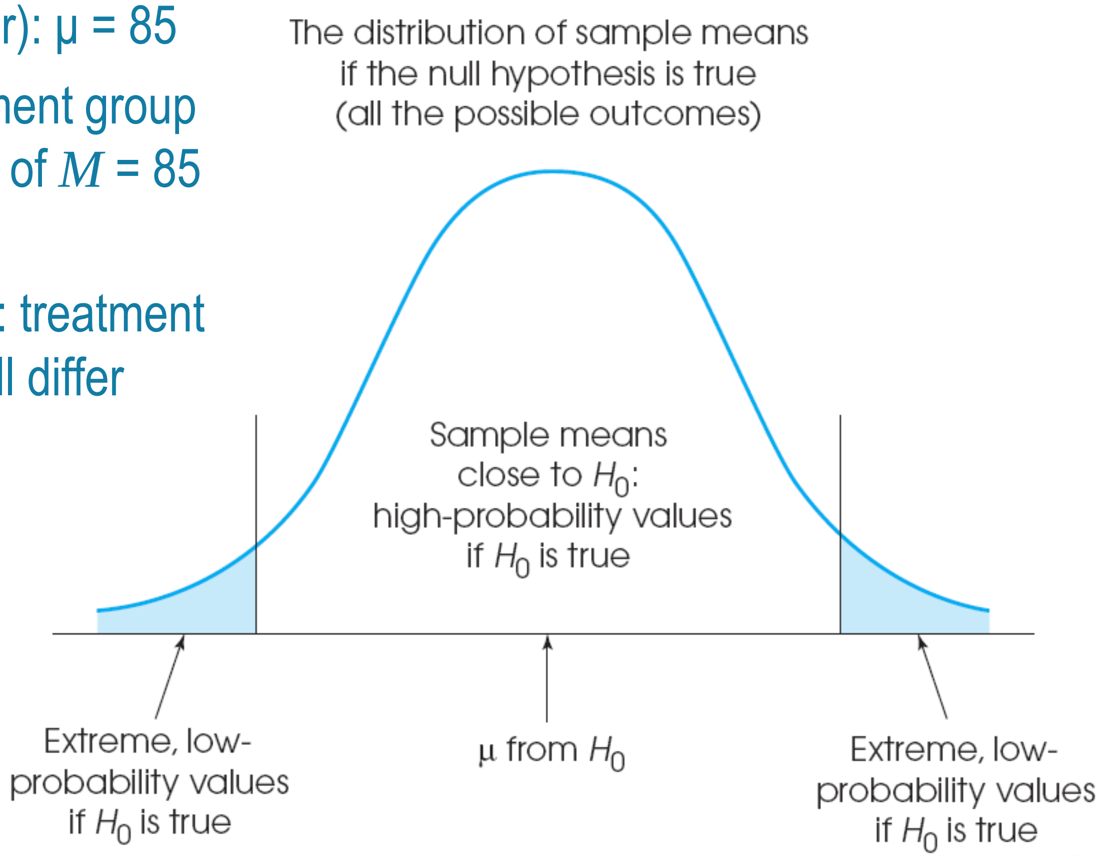
little known facts about
the positive impact of
alcohol on memory
during "cram" sessions

STEP 1: STATE THE HYPOTHESES (EXAMPLE)

- Dependent Variable
  - Amount of alcohol consumed the night before a statistics exam
- Independent/Treatment Variable
  - Intervention: Pamphlet (treatment group) or No Pamphlet (control group)
- Null Hypothesis (H_0)
  - No difference in alcohol consumption between the two groups the night before a statistics exam.
- Alternative Hypothesis (H_1)
  - The treatment group will consume more alcohol than the control group.

STEP 2: SET CRITERIA FOR DECISION

- Example
  - Exam 1 (Previous Semester): μ = 85
  - Null Hypothesis (H_0): treatment group will have mean exam score of M = 85 (σ = 8)
  - Alternative Hypothesis (H_1): treatment group mean exam score will differ from M = 85



STEP 2: SET CRITERIA FOR DECISION

- **Alpha Level/Level of Significance**

  probability value used to define the (unlikely) sample outcomes if the null hypothesis is true; e.g., α = .05, α = .01, α = .001

- **Critical Region**

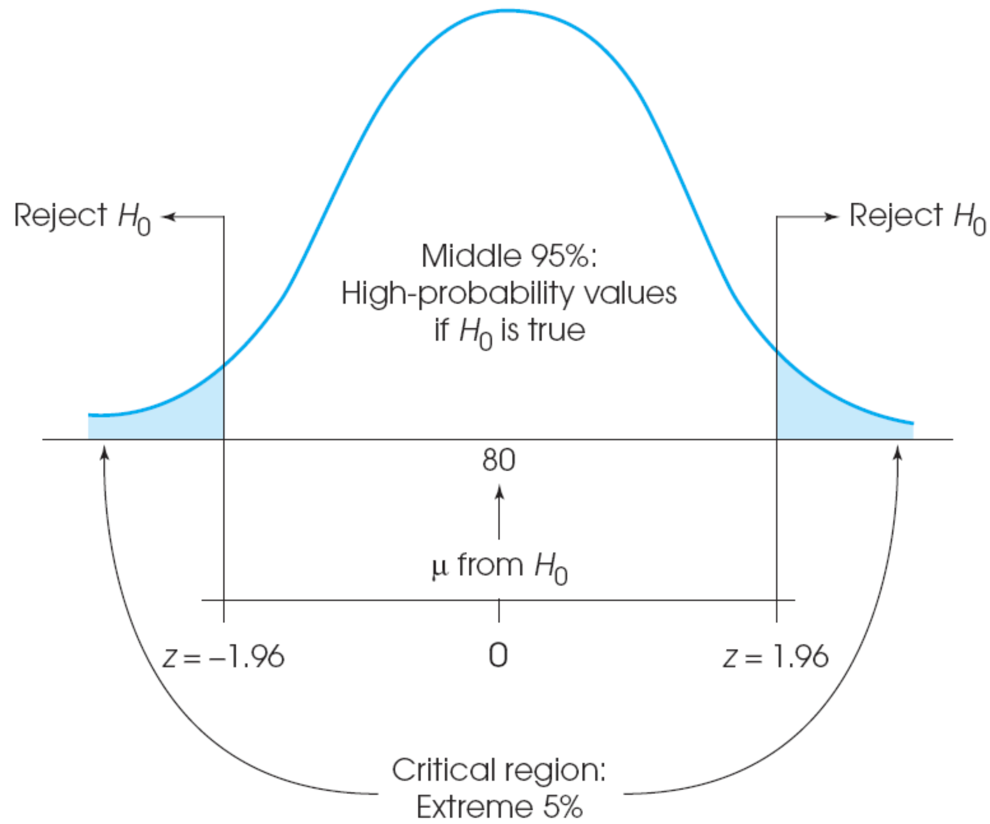
  extreme sample values that are very unlikely to be obtained if the null hypothesis is true

  - Boundaries determined by alpha level
  - If sample data falls within this region (the shaded tails), reject the null hypothesis

STEP 2: SET CRITERIA FOR DECISION

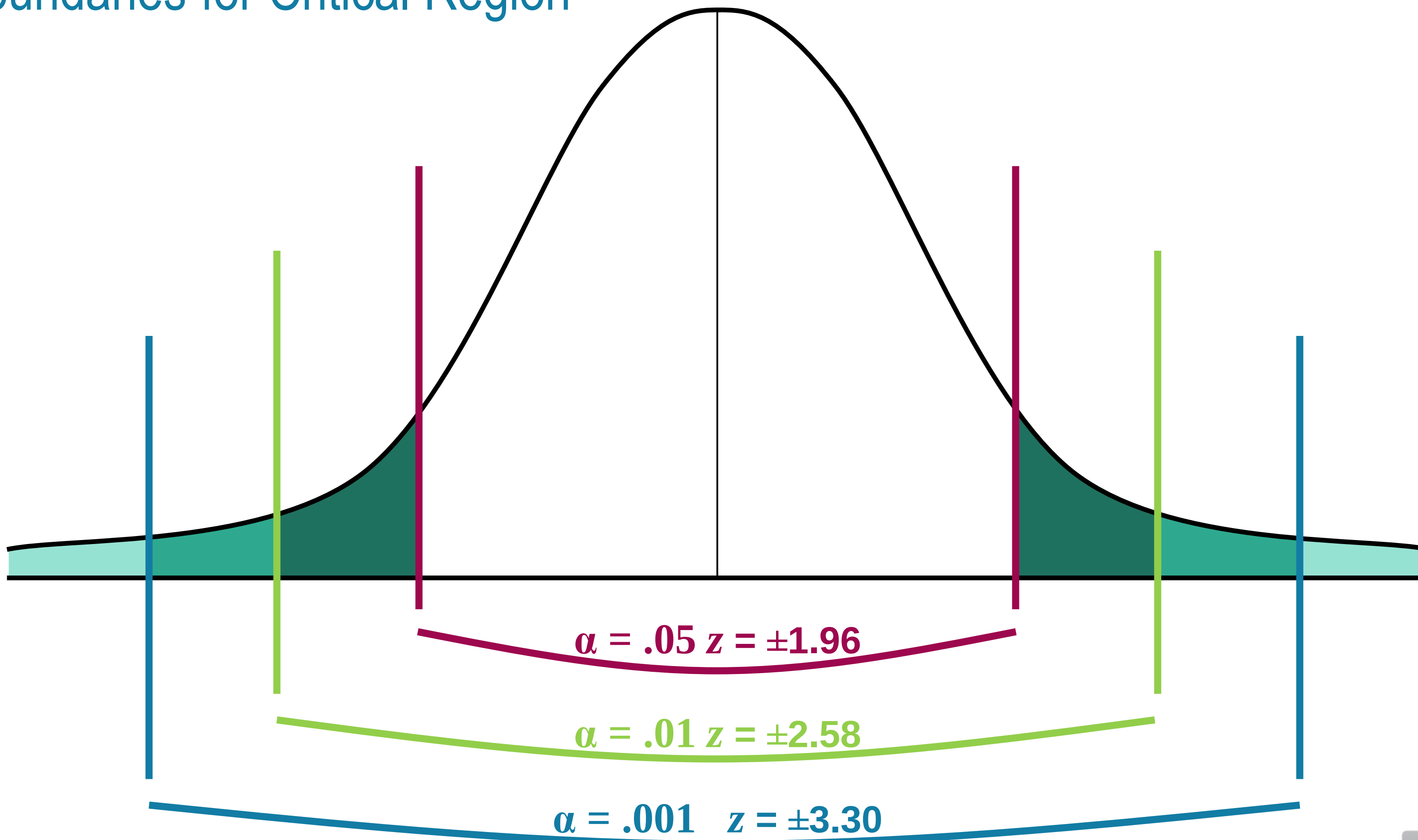
➢ Critical Region Boundaries

- ❖ Assume normal distribution
- ❖ Alpha Level + Unit Normal Table
- ❖ Example: if $\alpha = .05$, boundaries extreme 5%
    - ○ 2.5% in each tail (2-tailed)



STEP 2: SET CRITERIA FOR DECISION

- Boundaries for Critical Region



STEP 3: COLLECT, COMPUTE

- Collect data
- Compute sample mean
- Transform sample mean M to z-score

$$z = \frac{M - \mu}{\sigma_M}$$

- Example #2

$$z = \frac{95 - 85}{1.13} = \frac{10}{1.13} = 8.85$$

STEP 4: MAKE A DECISION

- Compare z-score with boundary of critical region for selected level of significance

- If…
  - z-score falls in the tails, our mean is significantly different from H_0
    - Reject H_0
  - z-score falls between the tails, our mean is not significantly different from H_0
    - Fail to reject H_0

HYPOTHESIS TESTING: AN EXAMPLE (2-TAIL)

- How to Ace a Statistics Exam…
  - Population: $\mu = 85$, $\sigma = 8$
  - Hypotheses
    - H_0: Sample mean will not differ from $M = 85$
    - H_1: Sample mean will differ from $M = 85$
  - Set Criteria (Significance Level/Alpha Level)
    - $\alpha = .05$

EXAMPLE (2-TAIL)

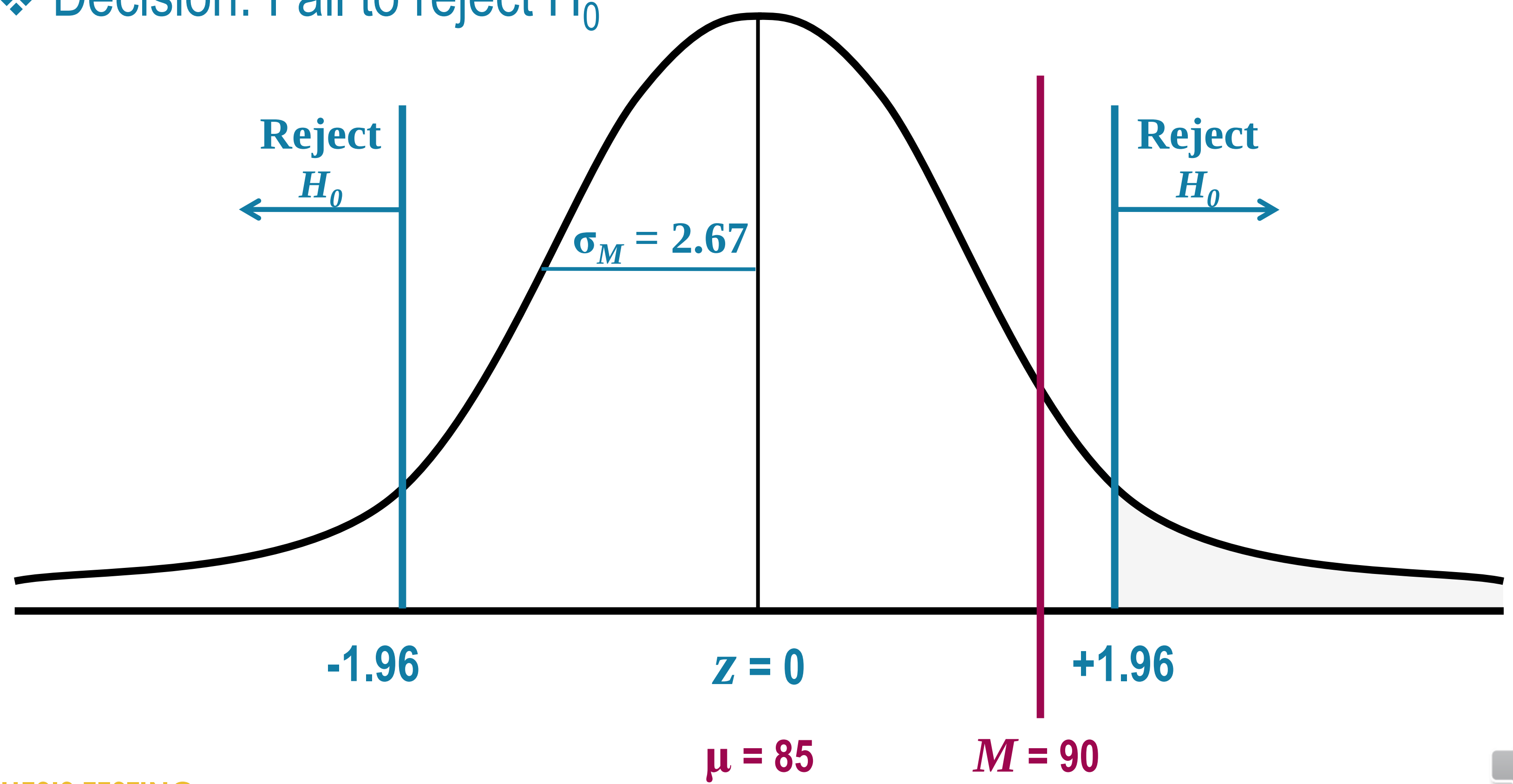
- How to Ace a Statistics Exam…
  - Collect Data & Compute Statistics
    - Intervention to 9 students
    - Mean exam score, M = 90

$$\sigma_M = \frac{\sigma}{\sqrt{9}} = \frac{8}{3} = 2.67$$

$$z = \frac{90 - 85}{2.67} = \frac{5}{2.67} = 1.87$$

EXAMPLE (2-TAIL)

- How to Ace a Statistics Exam…
  - Decision: Fail to reject H_0



REVISITING *Z*-SCORE STATISTICS

- A Test Statistic
  - Single, specific statistic
  - Calculated from the sample data
  - Used to test H_0
- Rule of Thumb…
  - Large values of z
    - Sample data pry DID NOT occur by chance – result of IV
  - Small values of z
    - Sample data pry DID occur by chance – not result of IV

uncertainty leads to error

ERROR & ALPHA

UNCERTAINTY & ERROR

- Hypothesis Testing = Inferential Process
  - LOTS of room for error

- Types of Error
  - Type I Error
  - Type II Error

TYPE 1 ERRORS

error that occurs when the null hypothesis is rejected even though it is really true; the researcher identifies a treatment effect that does not really exist (a false positive)

- Common Cause & Biggest Problem
  - Sample data are misleading due to sampling error
  - Significant difference reported in literature even though it isn't real
- Type I Errors & Alpha Level
  - Alpha level = probability of committing a Type I Error
  - Lower alphas = less chances of Type I Error

TYPE II ERRORS

error that occurs when the null hypothesis is not rejected even it is really false; the researcher does not identify a treatment effect that really exists (a false negative)

- Common Cause & Biggest Problem
  - Sample mean in not in critical region even though there is a treatment effect
  - Overlook effectiveness of interventions

- Type II Errors & Probability
  - β = probability of committing a Type II Error

TYPE I & TYPE II ERRORS

- Experimenter's Decision

	Actual Situation	
	No Effect, H_0 True	Effect Exists, H_0 False
Reject H_0	Type I Error	☺
Retain H_0	☺	Type II Error

SELECTING AN ALPHA LEVEL

- Functions of Alpha Level
  - Critical region boundaries
  - Probability of a Type I error
- Primary Concern in Alpha Selection
  - Minimize risk of Type I Error without maximizing risk of Type II Error
- Common Alpha Levels
  - $\alpha = .05$, $\alpha = .01$, $\alpha = .001$

testing null hypotheses

HYPOTHESIS TESTS

HYPOTHESIS TESTS: INFLUENTIAL FACTORS

- Magnitude of difference between sample mean and population mean (in *z*-score formula, larger difference ⇨ larger numerator)

$$z = \frac{M - \mu}{\sigma_M} \qquad \sigma_M = \frac{\sigma}{\sqrt{n}}$$

- Variability of scores (influences σ_M; more variability ⇨ larger σ_M)

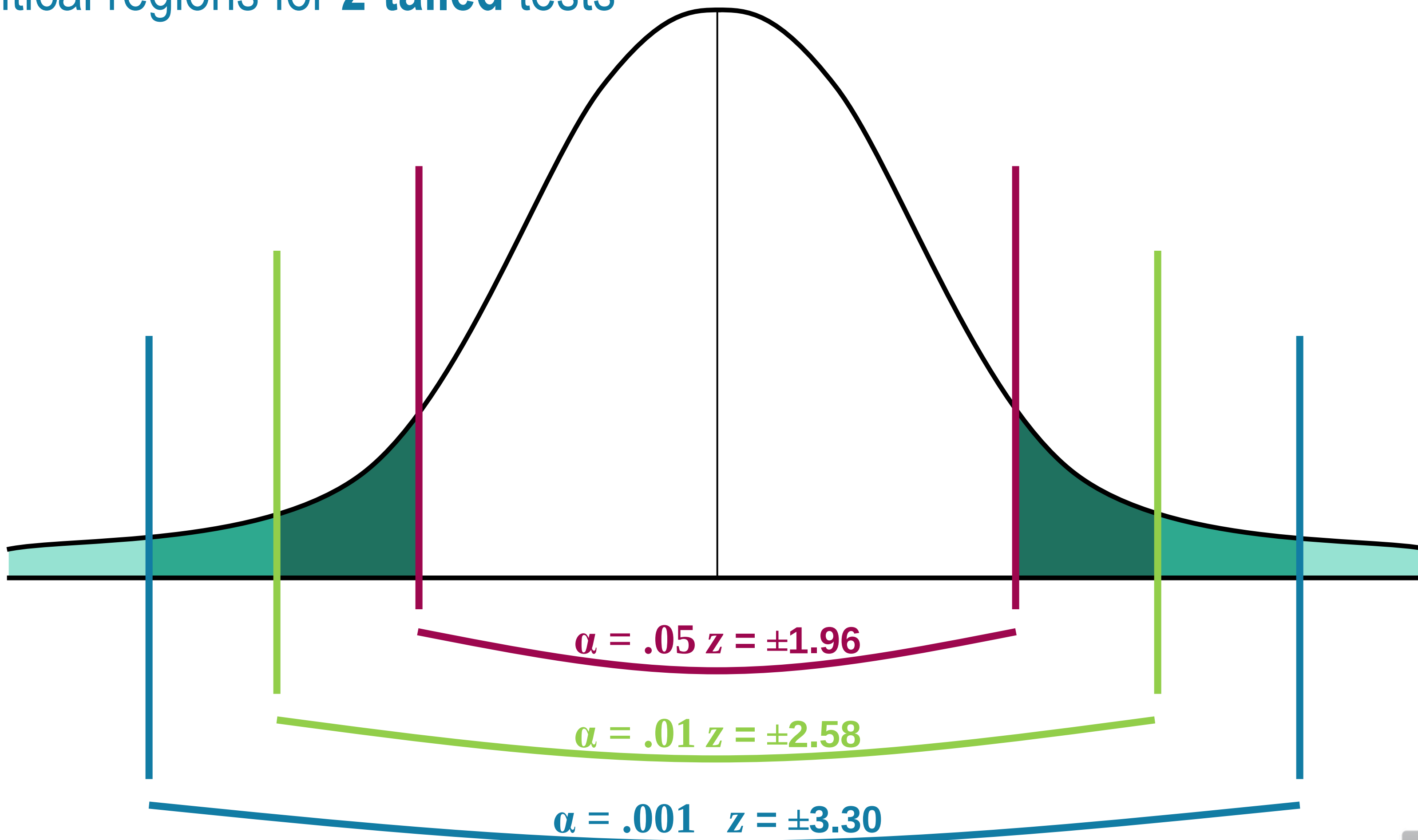
- Sample size (influences σ_M; larger sample size ⇨ smaller σ_M)

HYPOTHESIS TESTS: ASSUMPTIONS

- Random Sampling
- Independent Observations
- Value of σ is Constant
  - Despite treatment
- Normal sampling distribution

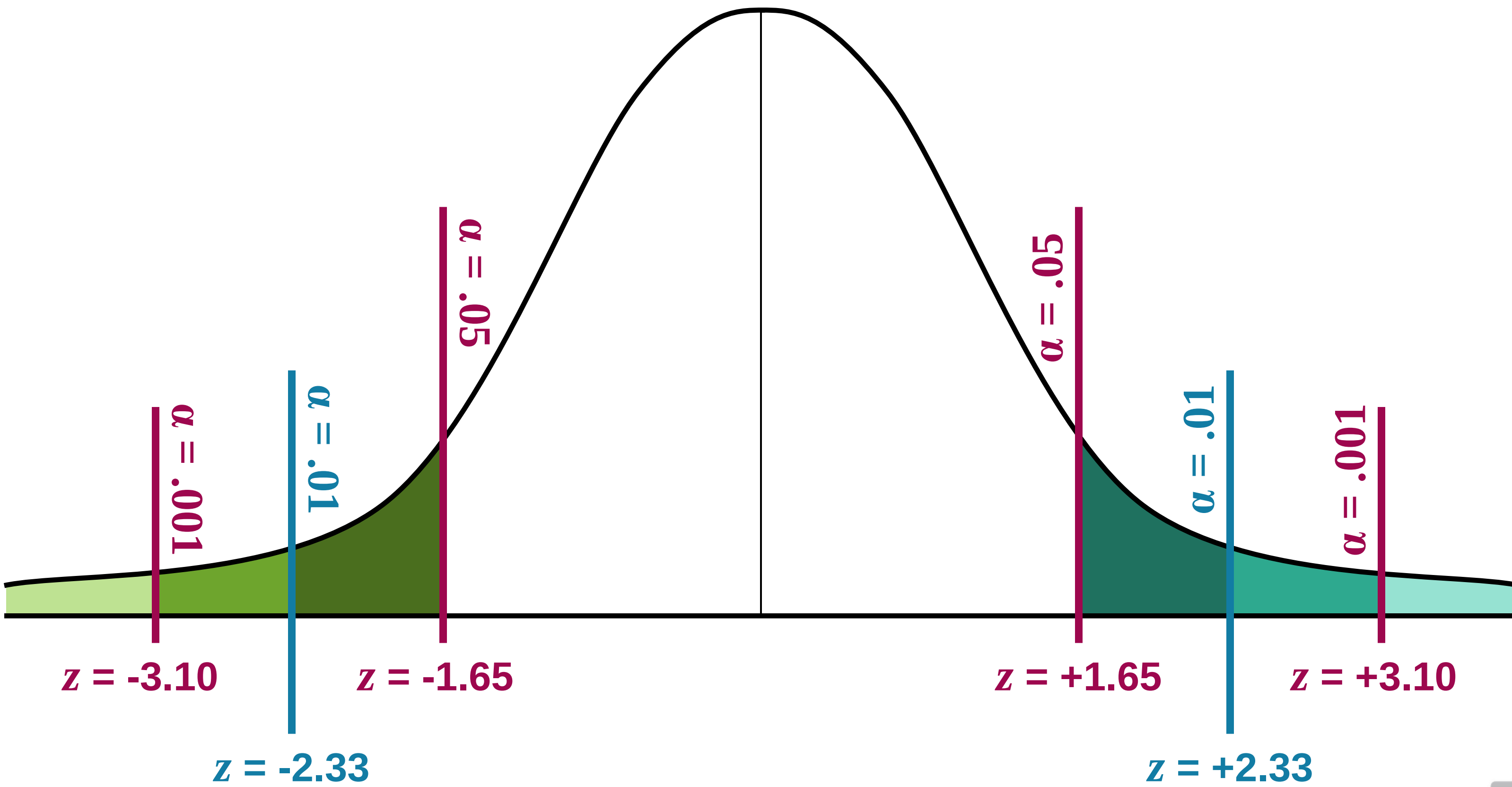
NON-DIRECTIONAL HYPOTHESIS TESTS

- Critical regions for **2-tailed** tests



DIRECTIONAL HYPOTHESIS TESTS

- Critical regions for **1-tailed** tests
  - Blue **or** Green tail of distribution – **NOT BOTH**



ALTERNATIVE HYPOTHESES

- Alternative Hypotheses for **2-tailed** tests
  - Do not specify direction of difference
  - Do not hypothesize whether sample mean should be lower or higher than population mean

- Alternative Hypotheses for **1-tailed** tests
  - Specify a difference
  - Hypothesis specifies whether sample mean should be lower or higher than population mean

NULL HYPOTHESES

- Null Hypotheses for **2-tailed** tests
  - Specify no difference between sample & population

- Null Hypotheses for **1-tailed** tests
  - Specify the opposite of the alternative hypothesis
  - Example #2
    - H_0: $\mu \leq 85$ (There is no increase in test scores.)
    - H_1: $\mu > 85$ (There is an increase in test scores.)

HYPOTHESIS TESTS: AN EXAMPLE (1-TAIL)

- How to Ace a Statistics Exam…
  - Population: μ = 85, σ = 8
  - Hypotheses
    - H_0: Sample mean will be less than or equal to M = 85
    - H_1: Sample mean be greater than M = 85
  - Set Criteria (Significance Level/Alpha Level)
    - α = .05

HYPOTHESIS TESTS: AN EXAMPLE (1-TAIL)

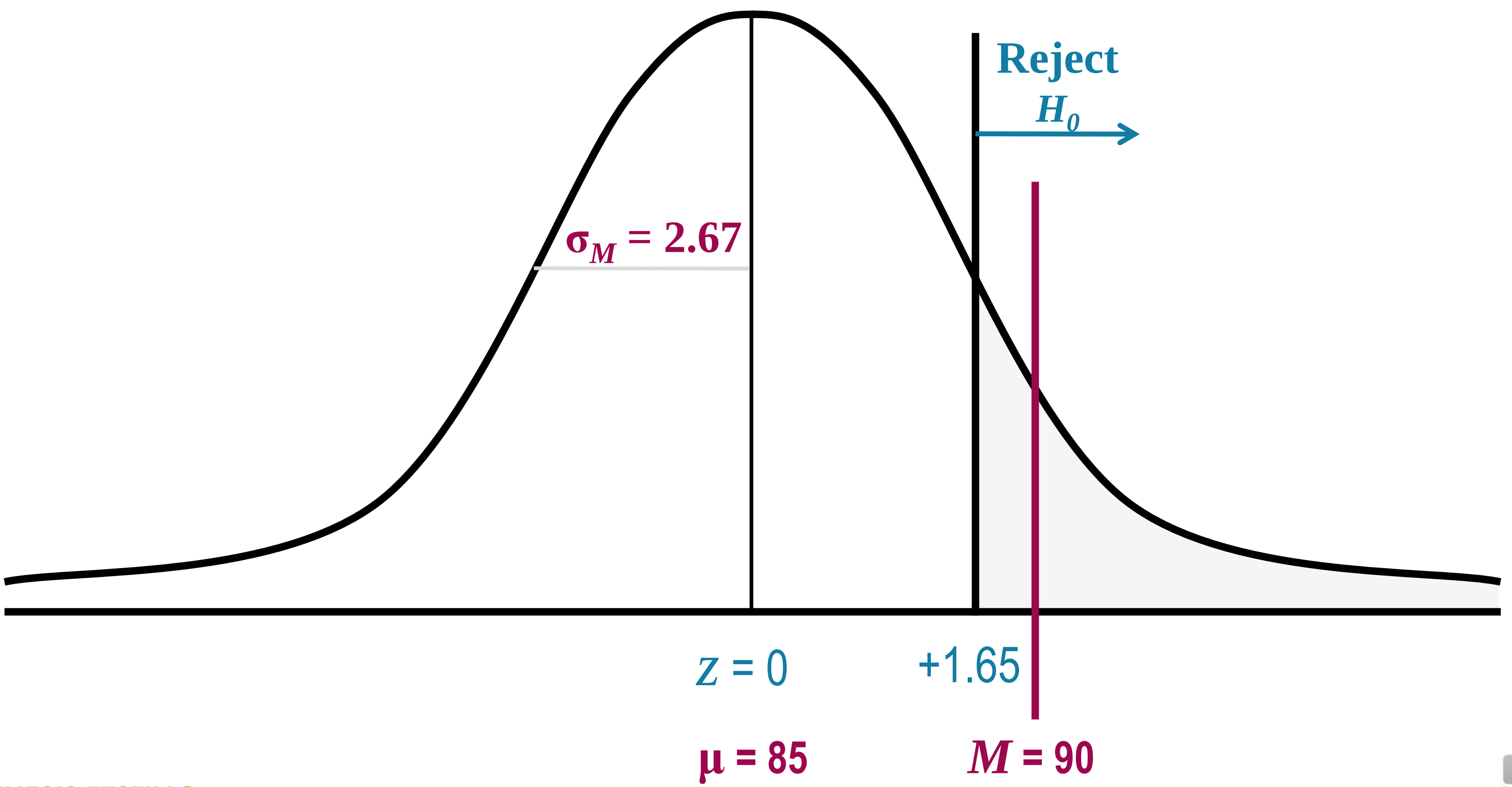
- How to Ace a Statistics Exam…
  - Collect Data & Compute Statistics
    - Intervention to 9 students
    - Mean exam score, M = 90

$$\sigma_M = \frac{\sigma}{\sqrt{9}} = \frac{8}{3} = 2.67$$

$$z = \frac{M - \mu}{\sigma_M} = \frac{90 - 85}{2.67} = \frac{5}{2.67} = 1.87$$

HYPOTHESIS TESTS: AN EXAMPLE (1-TAIL)

- How to Ace a Statistics Exam…
  - Decision: Reject H_0



estimating the magnitude of an effect

EFFECT SIZE

EFFECT SIZE

- Problem with hypothesis testing
  - Significance ≠ Meaningful/Important/Big Effect
    - Significance is relative comparison: treatment effect compared to standard error

- Effect Size

  statistic that describes the magnitude of an effect

- Measures size of treatment effect in terms of (population) standard deviation

EFFECT SIZE: COHEN'S *D*

- **Not** influenced by sample size

$$\text{Cohen's } d = \frac{\text{mean difference}}{\text{standard deviation}}$$

- Evaluating Cohen's d
  - $d = 0.2$ – Small Effect (mean difference ≈ 0.2 standard deviation)
  - $d = 0.5$ – Medium Effect (mean difference ≈ 0.5 standard deviation)
  - $d = 0.8$ – Large Effect (mean difference ≈ 0.8 standard deviation)

- Calculated the same for 1-tailed and 2-tailed tests *probability of correctly rejecting a false null hypothesis*

STATISTICAL POWER

STATISTICAL POWER

the probability of correctly rejecting a null hypothesis when it is not true; the probability that a hypothesis test will identify a treatment effect when if one really exists

- A priori
  - Calculate power before collecting data
  - Determine probability of finding treatment effect
- Power is influenced by…
  - Sample size
  - Expected effect size
  - Significance level for hypothesis test (α)