ONE-FACTOR ANALYSIS: PART 2

Chapter 3-5, 8

LEARNING OBJECTIVES

- Define a sampling distribution for a statistic
- Write null and alternative hypotheses for two-sample tests and simple linear regression
- Understand relationship between null and alternative distributions
- Explain how to calculate Type 1 error and power

BREAD YEAST AND TIME TO RISE

- Yeast is added to bread to make it "rise"
  - Contacts warm water and feeds on sugars in the flour
  - Releases carbon dioxide at some rate
- Bread dough rests for a period of time to allow this process to happen
- Compare two different rise times to see if there is a significant difference in dough height
  - 35 and 45 minutes (quantitative or categorical?)
- Large batch of dough partitioned into *N* loaf pans

TWO SAMPLE HYPOTHESIS TESTING

- Have r_1 and r_2 loaf pans for 35 and 45 minutes
  - Conclude there are differences in rise time if the mean dough heights are different
- Write effects model for this data (as in R)

- Hypothesis test:
  - Assume opposite of what you want to show (null hypothesis)

$$H_0: \mu_1 = \mu_2 \Leftrightarrow \mu_1 - \mu_2 = 0$$

  - Choose test statistic whose distribution changes depending on whether the null or or alternative hypothesis is true
  - Statistic distribution is called sampling distribution
  - If the observed test statistic is unlikely to come from null distribution, we reject the null hypothesis

COMMENTS ABOUT HYPOTHESIS TESTING

- Prefer to write null hypothesis in terms of estimable functions when we can
  - Can only test hypotheses that can be written this way
- Need to know the null and alternative distributions
- Always chance we falsely reject null (Type I error) but we can control this error rate with α
- If we fail to reject the null we cannot conclude the null hypothesis is true

TWO SAMPLE HYPOTHESIS TESTING

- Unbiased estimator for $\mu_1 - \mu_2$ is $\bar{Y}_1. - \bar{Y}_2.$

- Sampling distribution:

$$N\left(\mu_1 - \mu_2, \sigma^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right)\right)$$

- Basic idea: if $\bar{Y}_1. - \bar{Y}_2.$ "far" from 0 then we reject null,

- What does it mean to be far?
  - Take into account expected variability
  - Cutoff for unlikely values controls Type 1 error

TWO-SAMPLE Z-STATISTIC

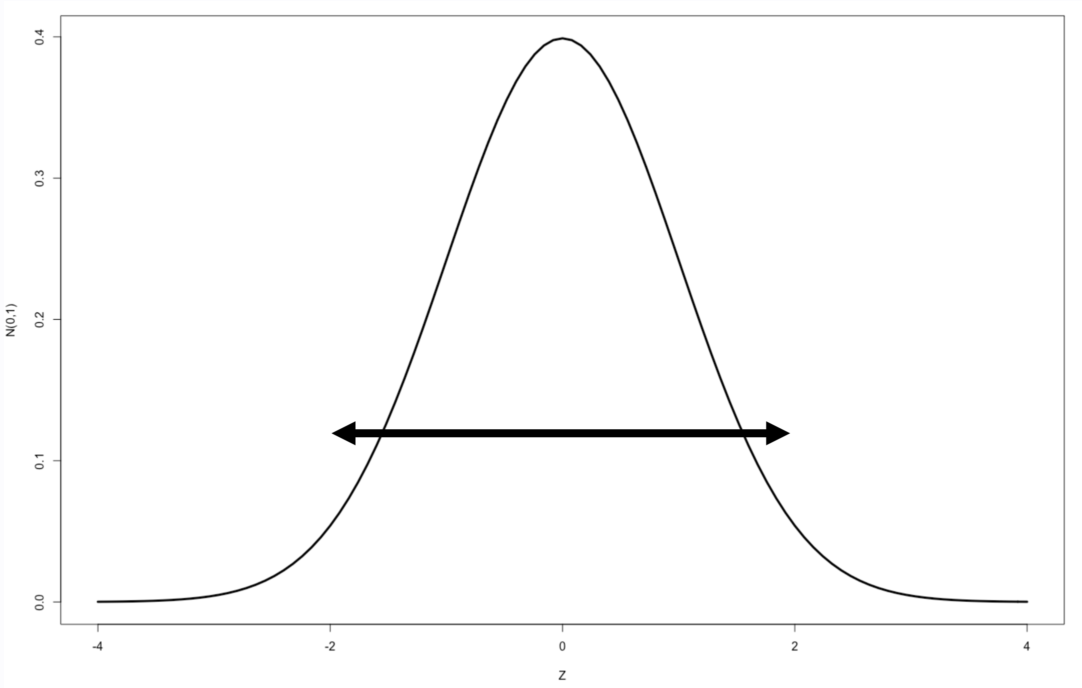
- A related test statistic we could use is

$$z = \frac{(\bar{y}_{1\cdot} - \bar{y}_{2\cdot})}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}}$$

- If $\mu_1 - \mu_2 = 0$ and we knew σ then *Z* would be N(0, 1)
  - Call this null distribution
  - Unlikely for |*z*| > 2 (absolute value of z)
- If $\mu_1 - \mu_2 \neq 0$, *Z* does not follow N(0,1) distribution
  - Call this alternative distribution
  - More likely for |*z*| > 2

TWO-SAMPLE Z-STATISTIC NULL DISTRIBUTION

- Plot of null distribution

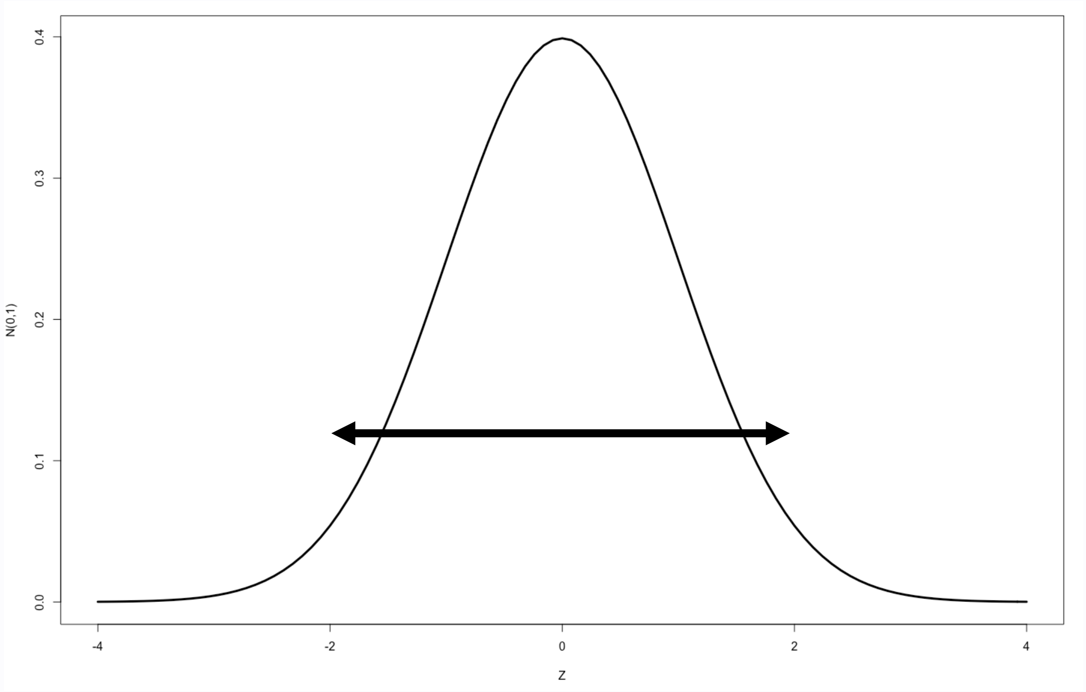


If null true, then 95% chance observed Z is between -1.96 and 1.96

If observed Z is outside this range, makes sense to reject null

TWO-SAMPLE Z-STATISTIC CRITICAL VALUE REJECTION

- Plot of null distribution

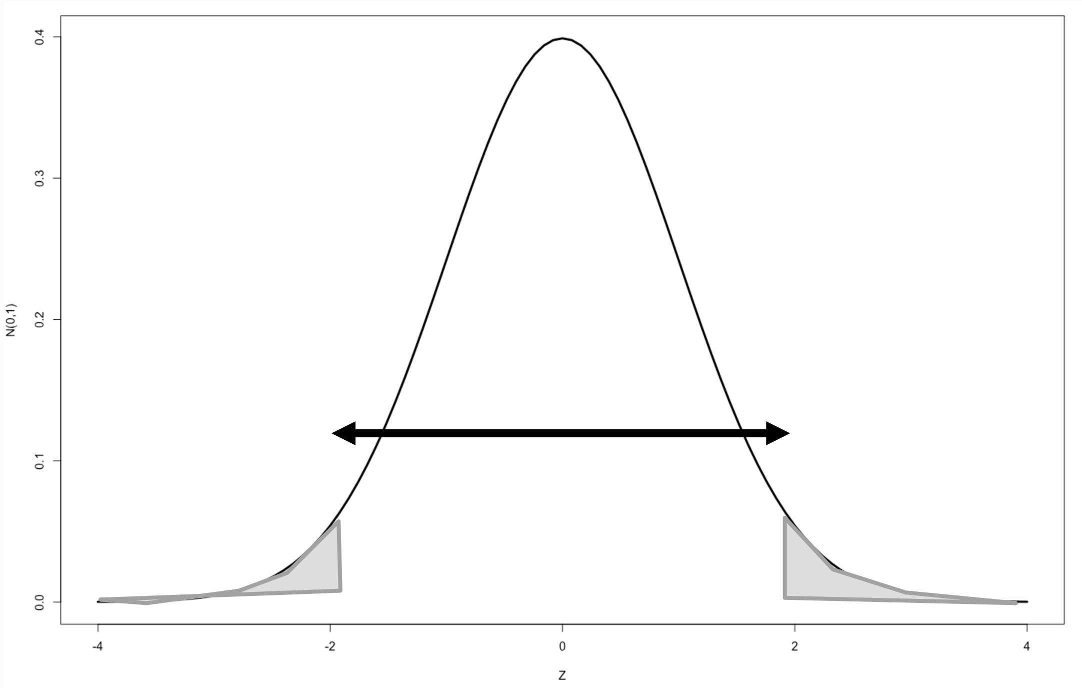


Rejection rule: if $|Z_{obs}| > 1.96$ then reject null hypothesis

Call 1.96 the critical value

TYPE 1 ERROR

- **Type 1 error:** reject when null is true (false reject)

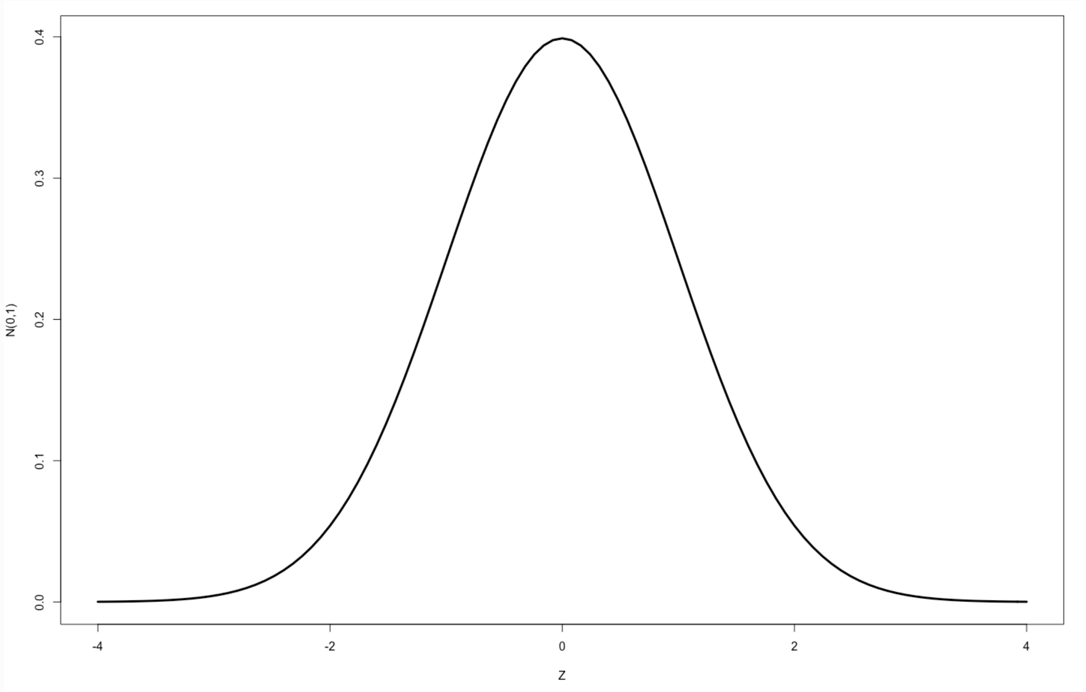


$$P(Type\ 1\ error) = P_{H_0}(|Z| > 1.96) = 0.05 = \alpha$$

Assuming H_0 is true

TWO-SAMPLE Z-STATISTIC P-VALUE REJECTION

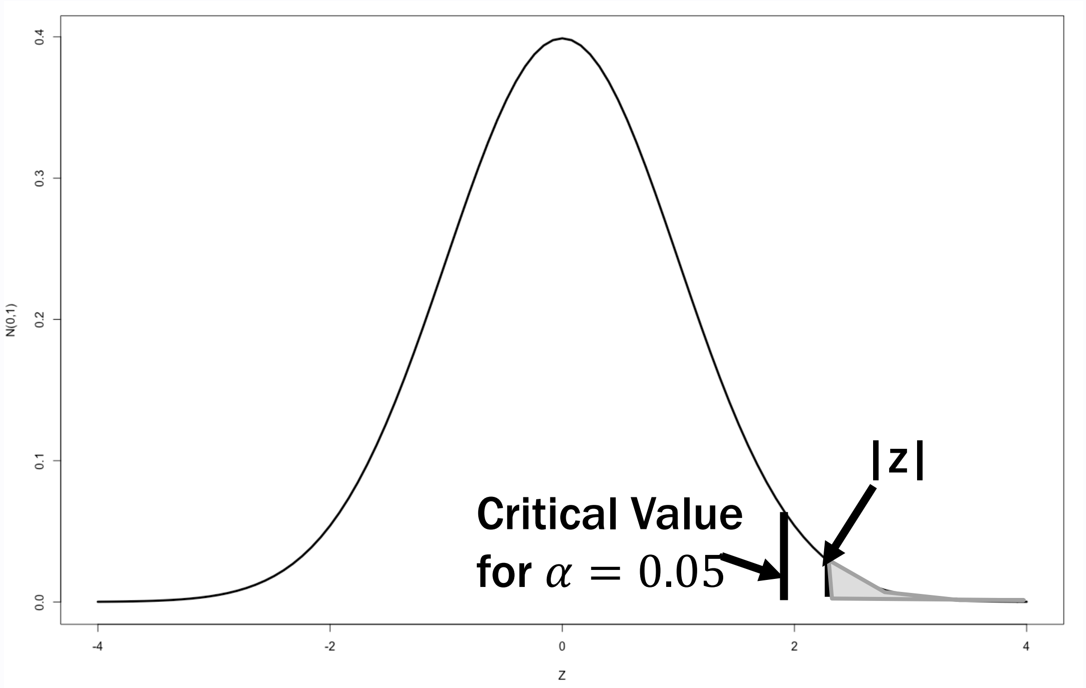
- An alternative rejection rule



Calculate probability of observed statistic (z) or more extreme value assuming the null is true

TWO-SAMPLE Z-STATISTIC P-VALUE REJECTION

- An alternative rejection rule



$$2 \times P_{H_0}(Z > |z|) = p - value$$

Reject null if $p - value < \alpha$

POWER

- Power: probability of rejecting assuming fixed alternative hypothesis is true

- For given data set, power depends on:
  - Alternative hypothesis (specific $|\mu_1 - \mu_2| > 0$)
  - Critical level for desired significance level
  - Sample sizes
  - σ

- Use power to determine "optimal" relative sample sizes after fixing alternative, critical value, and σ

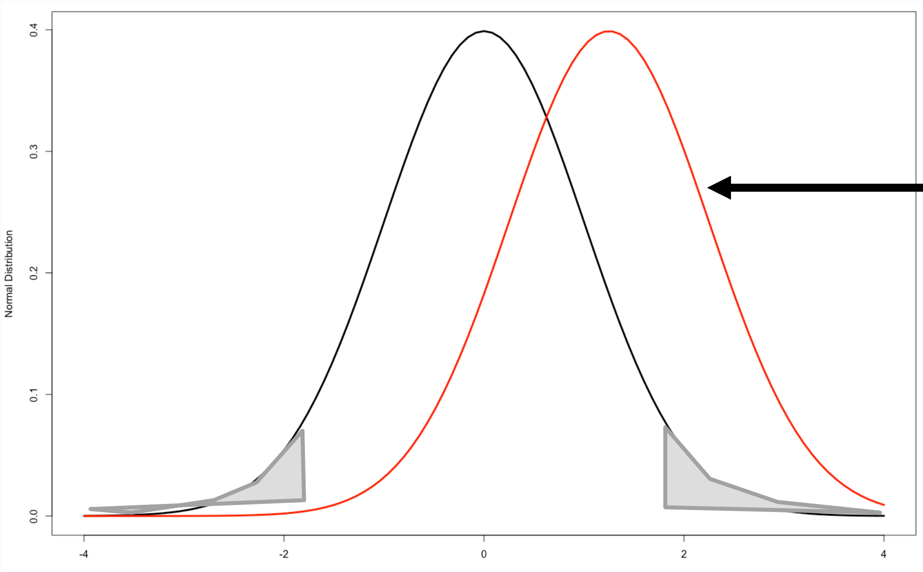
TWO-SAMPLE Z-STATISTIC POWER

- If $\mu_1 - \mu_2 \neq 0$ then alternative distribution of Z is

$$Z \sim N\left(\boxed{\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}}}, 1\right)$$

Noncentrality parameter

- The larger this mean is, the more likely we observe "extreme" values



$$P_{H_0}(|Z| > 1.96) = 0.05 = \alpha$$

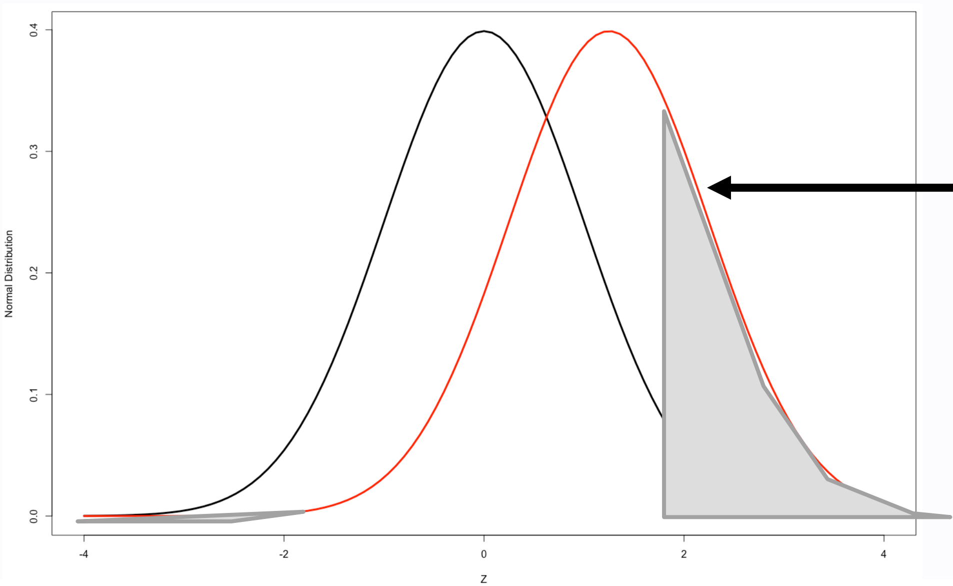
$\sigma = 1$ $\quad\mu_1 - \mu_2 = 0.25$

$r_1 = 10$ $\quad r_2 = 10$

$$\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}} = 1.25$$

TWO-SAMPLE Z-STATISTIC POWER

- If $\mu_1 - \mu_2 \neq 0$ then true distribution of Z is
$$Z \sim N\left(\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}}, 1\right)$$
- The larger this mean is, the more likely we observe "extreme" values



$P_{H_A}(|Z| > 1.96) = 0.2395 > 0.05$

Assuming H_A is true

$\sigma = 1 \quad \mu_1 - \mu_2 = 0.25$

$r_1 = 10 \quad r_2 = 10$

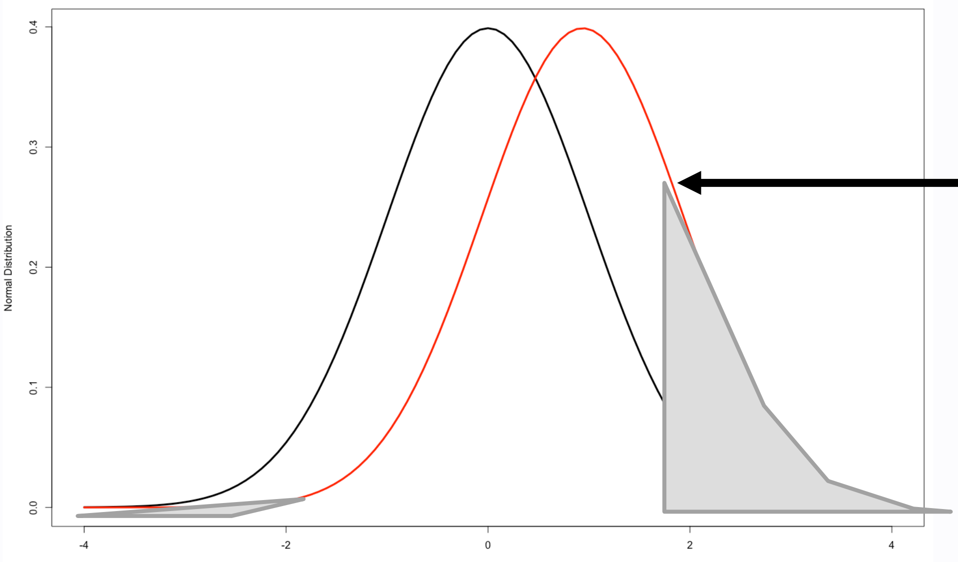
$$\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}} = 1.25$$

TWO-SAMPLE Z-STATISTIC POWER

- If $\mu_1 - \mu_2 \neq 0$ then true distribution of *Z* is

$$Z \sim N\left(\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}}, 1\right)$$

- The larger this mean is, the more likely we observe "extreme" values



$P_{H_A}(|Z| > 1.96) = 0.1552 > 0.05$

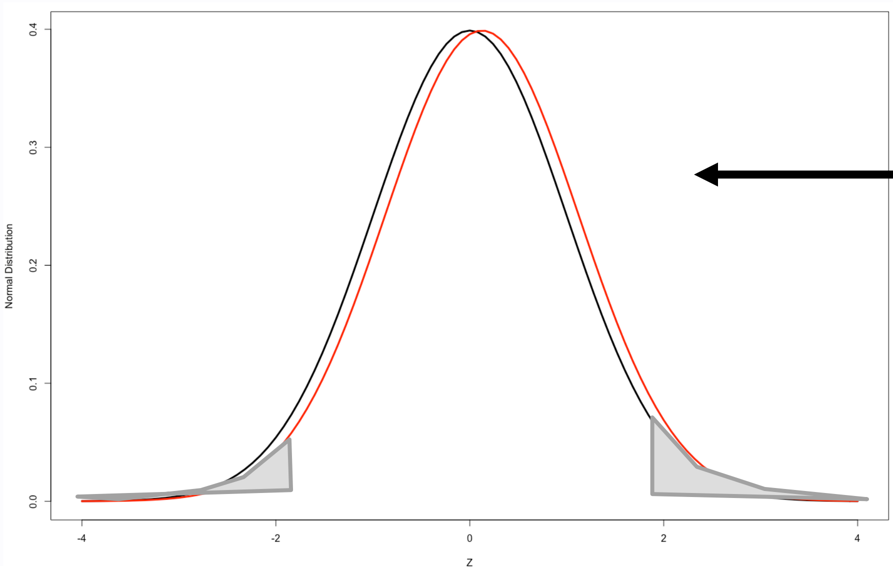
$\sigma = 1 \qquad \mu_1 - \mu_2 = 0.25$

$r_1 = 5 \qquad r_2 = 15$

$$\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}} = 0.9375$$

TWO-SAMPLE Z-STATISTIC POWER

- If $\mu_1 - \mu_2 \neq 0$ then true distribution of Z is
$$Z \sim N\left(\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}}, 1\right)$$
- The larger this mean is, the more likely we observe "extreme" values



$P_{H_A}(|Z| > 1.96) = 0.0518 > 0.05$

$\sigma = 10$ $\mu_1 - \mu_2 = 0.25$

$r_1 = 10$ $r_2 = 10$

$$\frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}} = 0.125$$

LESSONS FROM POWER

- Achieving high power is a delicate balance between
  - Alternative hypothesis (specific $|\mu_1 - \mu_2| > 0$)
  - Critical level for desired significance level
  - Sample sizes
  - σ
- Equal sample sizes maximize power when the two populations have equal variance
- Minimizing σ maximizes power (local error control)
- Power can only be calculated for guesses of the true parameters
- If the alternative is true, may have low power which is why we never accept the null, only fail to reject

TWO-SAMPLE CONFIDENCE INTERVALS

- If we reject the null we should provide estimate of what the difference truly is
- Use a $100(1-\alpha)\%$ confidence interval centered at estimate

$$\underbrace{(\bar{Y}_{1\cdot} - \bar{Y}_{2\cdot})}_{\text{Estimator}} \pm \underbrace{t_{n-2,\alpha/2}}_{\text{Critical value from T-distribution with df}} \underbrace{\hat{\sigma}\sqrt{\frac{1}{r_1}+\frac{1}{r_2}}}_{\text{Standard error of estimate}}$$

- Like “inverting” hypothesis test

LEARNING OBJECTIVES REVIEW

- Define a sampling distribution for a statistic
- Write null and alternative hypotheses for two-sample tests and simple linear regression
- Understand relationship between null and alternative distributions
- Explain how to calculate Type 1 error and power