Learning Objectives

- **Describe** the four possible decisions in the *Null Hypothesis Statistical Testing (NHST)* framework
- **Identify** factors that increase the likelihood of rejecting the null hypothesis (H_0)
- **Justify** decisions to manage error
- **Visualize** statistical power

$(1 - \beta)$ = **Power**

- Research seeks at least 80% power
  - Not 40-60%!!

- Visualize in ***G*Power*** [*(or Laken's Shiny)*]
  - Most common tool for calculating power
  - Normal distribution (or *z*-distribution)
    - What is/are the *z*-score for: $\alpha_{2\text{-tail}} = .05$

Happiness Therapy

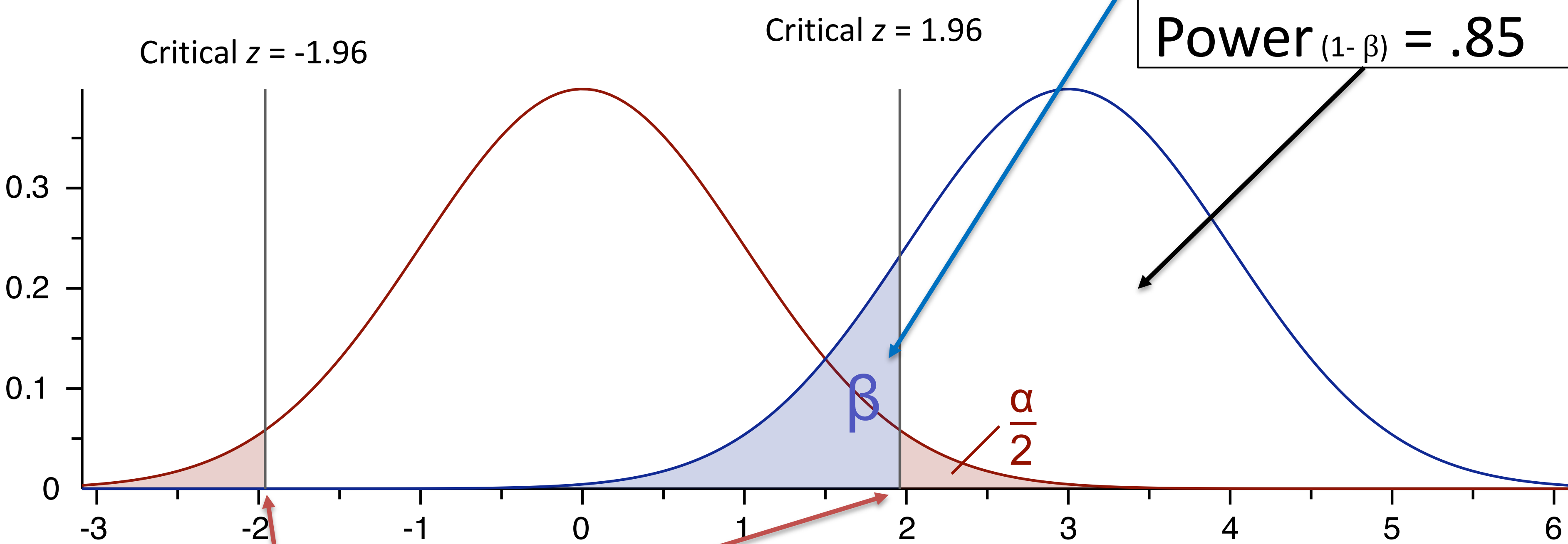
- *X* number of participants attend happiness therapy
  - Pre-test (on a scale of 1 to 10, how happy are you right now)
  - Post-test (on a scale of 1 to 10, how happy are you right now)
    - (Post-test – Pre-test) = positive numbers are consistent with the therapy working!
- H_0: Participants attending therapy will report similar levels of happiness following therapy
  - Differences in happiness will be due to chance
- H_1: Participants attending therapy will show different levels of happiness after therapy
  - Therapy changes happiness levels (in addition to chance)

Red = H_0 Distribution (If H_0 were true)

Blue = H_1 Distribution (If H_1 were true)

β = .15

Power $_{(1-\beta)}$ = .85



$\alpha_{2\text{-tail}} = .05$

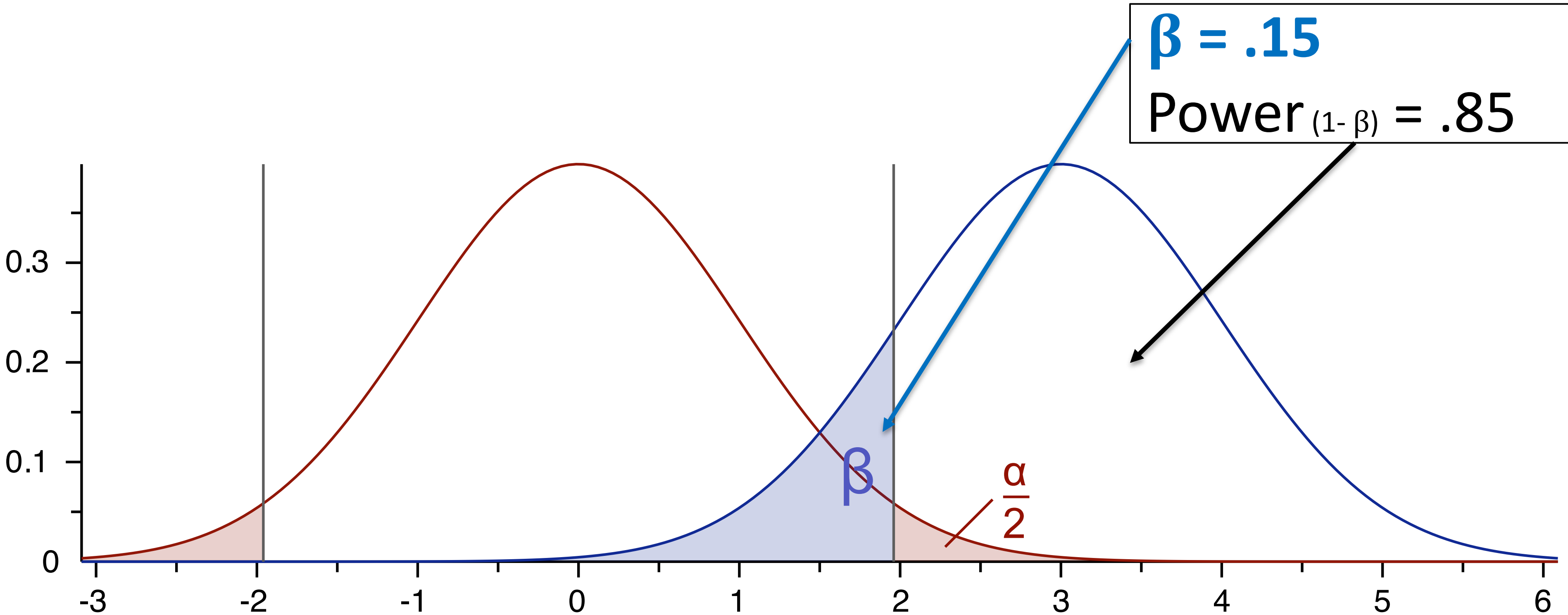
Critical $z = \pm 1.96$

$\bar{X}$-μ = 3 units

$\sigma_{\bar{X}}$ = 1 unit

Red = H_0 Distribution (If H_0 were true)

Blue = H_1 Distribution (If H_1 were true)

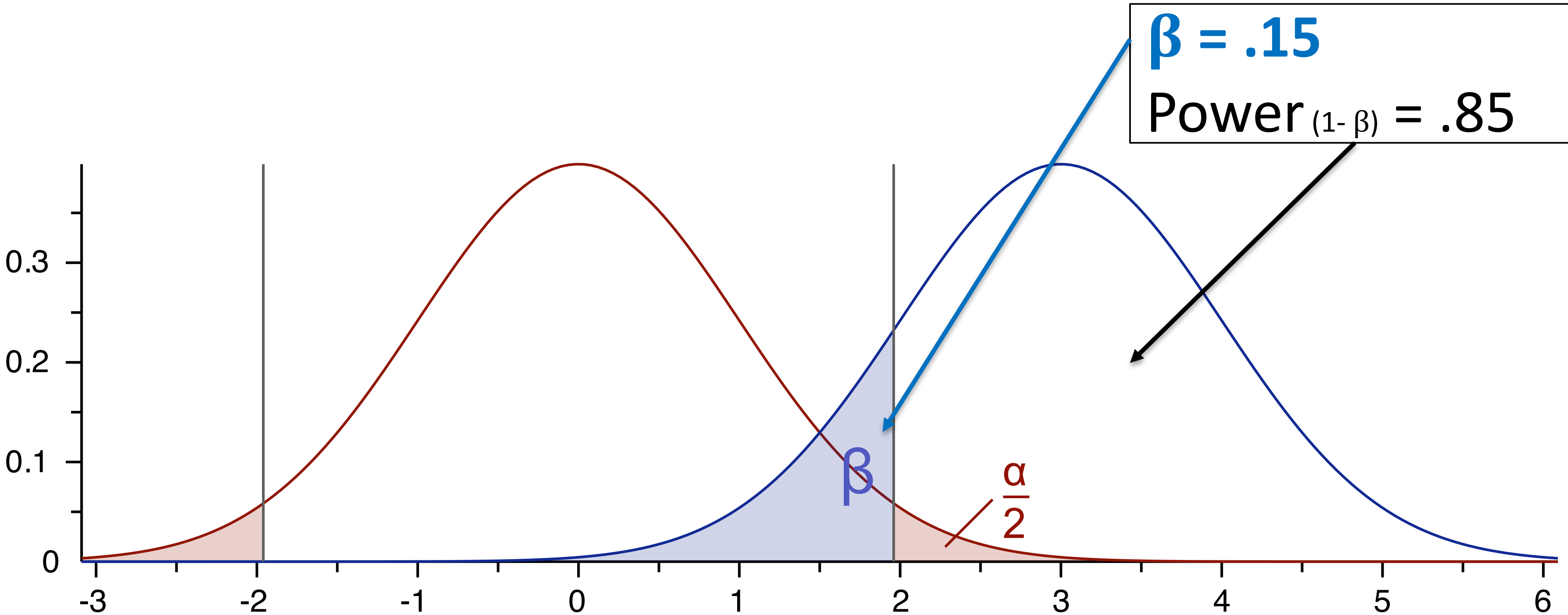


When H_0 is true (we're sampling from the **red** distribution), Type I errors occur 5% of the time.

When H_1 is true (sampling from **blue** distribution), Type II errors occur 15% of time.

Red = H_0 Distribution (If H_0 were true)

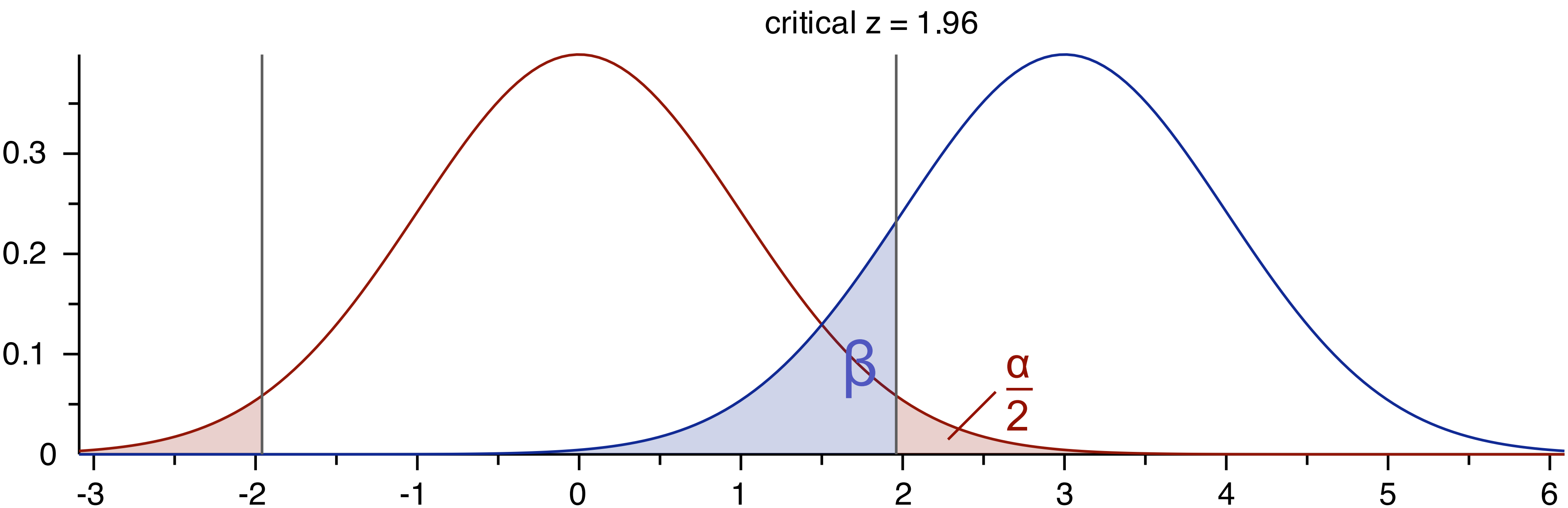
Blue = H_1 Distribution (If H_1 were true)



When H_0 is true, we correctly fail to reject H_0 95% of the time.
(95% = 1 – α)

When H_1 is true, we correctly reject H_0 85% of the time.
(85% = 1 – β)

Red = If H_0 is true
Blue = If H_1 is true



$\alpha_{2\text{-tail}}$ = .05

β = .15

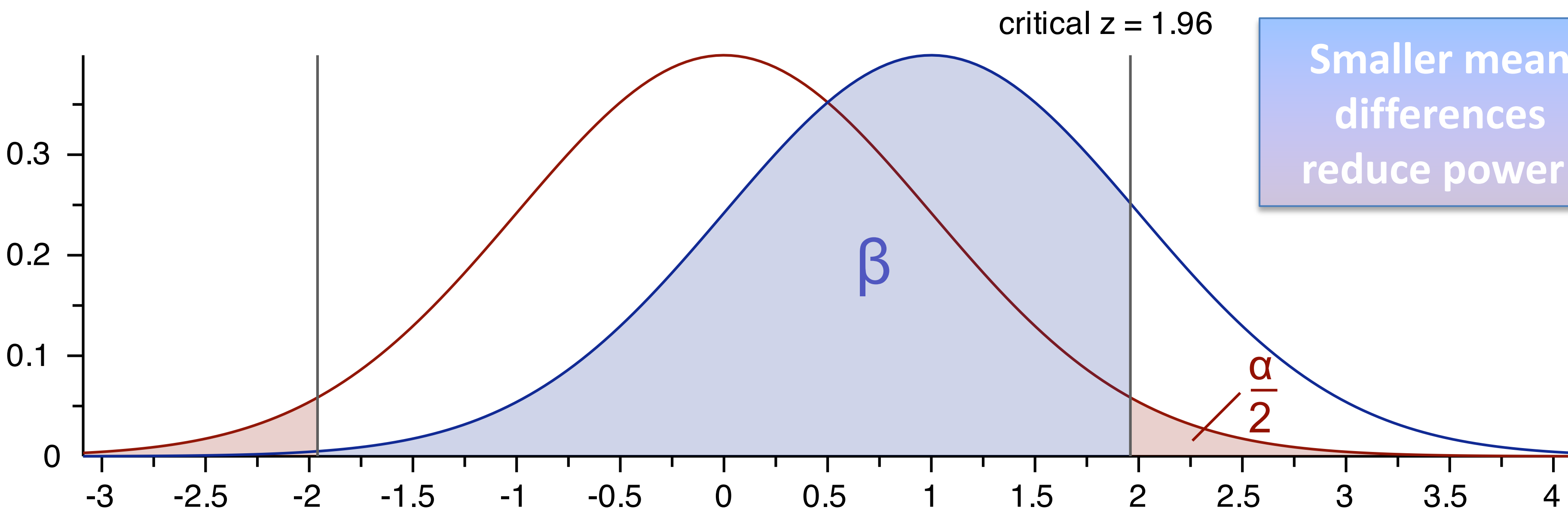
Power = .85

$\bar{X}$ - μ = 3 units

$\sigma_{\bar{X}}$ = 1 unit

Power when effect size is smaller!

Red = If H_0 is true
Blue = If H_1 is true



Smaller mean differences reduce power!

$\alpha_{2\text{-tail}} = .05$

β = .83

Power = .17

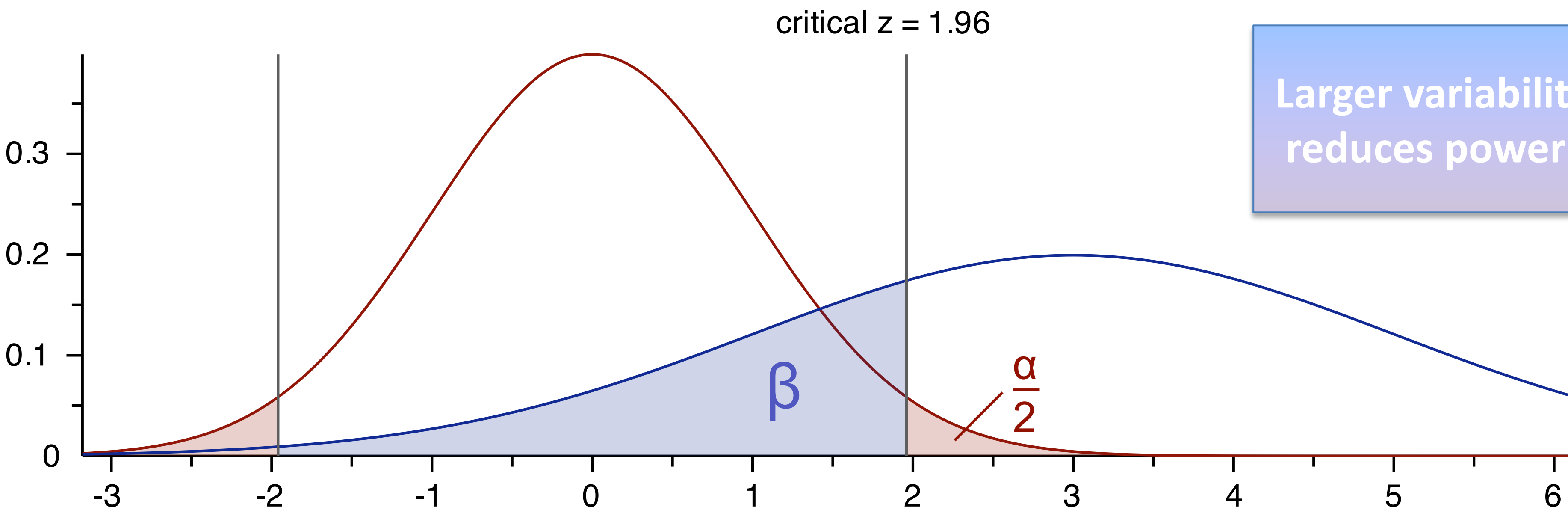
$\bar{X}$ - μ = **1 unit** **(was 3)**

$\sigma_{\bar{X}}$ = 1 unit

Power when variability is greater!

Red = If H_0 is true

Blue = If H_1 is true



Larger variability reduces power!

$\alpha_{\text{2-tail}} = .05$

β = .29

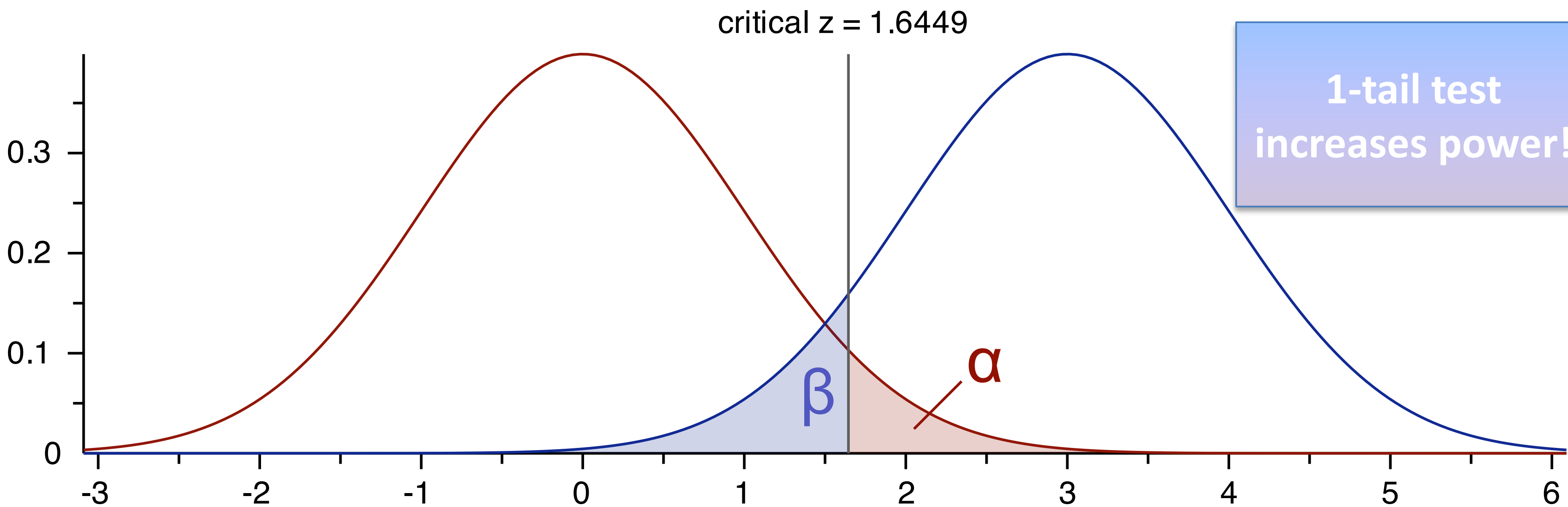
Power = .71

$\bar{X}$ - μ = 3 units

$\sigma_{\bar{X}}$ = **2 units** **(was 1)**

Power when 1-tail (vs. 2-tail) test!

Red = If H_0 is true
Blue = If H_1 is true



1-tail test increases power!

$\alpha_{\text{1-tail}}$ **= .05**

β = .09

Power = .91

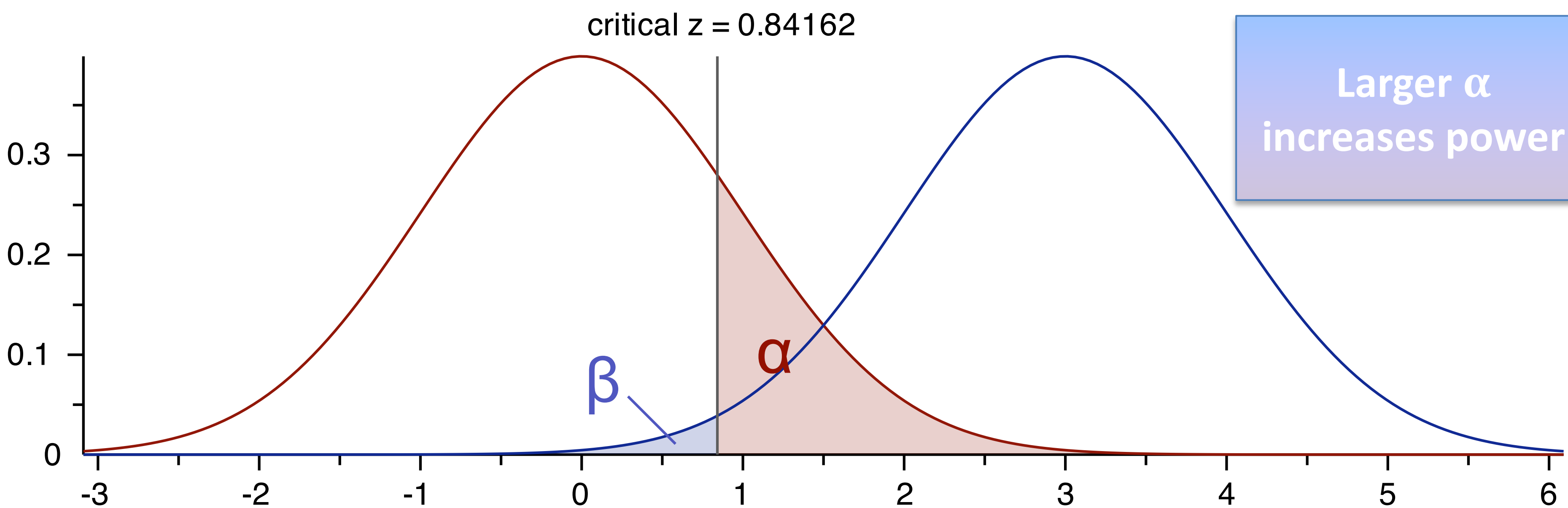
$\bar{X}$ - μ = 3 units

$\sigma_{\bar{X}}$ = 1 unit

Power when α is large!

Red = If H_0 is true

Blue = If H_1 is true



Larger α increases power

$\alpha_{1\text{-tail}}$ = **.20**

β = .01

Power = .99

$\bar{X}$ - μ = 3 units

$\sigma_{\bar{X}}$ = 1 unit

3 Common Types of Power

1. ***a priori* –** Before data, find *N* given:
   – α, $(1 - \beta)$, expected effect size

2. ***post hoc* –** After data, find power given:
   – α, *N*, **observed effect size**

3. ***Sensitivity* –** Before/after data, find detectable effect size given:
   – α, $(1 - \beta)$, *N*

Setting Power

- **Exp. 1**: "We sought to collect 80 participants… Sensitivity analysis indicated with power set at .80, we could detect an effect size as small as d_z = .317"

3. Sensitivity power b/c we don't know the size of the effect
    - α = .05
    - (1- β) = .80
    - *N* = 80

Setting Power

- **Exp. 2**: "Based on the observed effect size of d_z = .430 in Experiment 1, we sampled from 48 participants to set power at .80"

1. Use *a priori* power after getting an estimate of effect size in Exp. 1
    - α = .05
    - (1- β) = .80
    - Effect size = d_z = .430

Support the Guardian

Fund independent journalism with $13 per month

Support us →

The Guardian

News | Opinion | Sport | Culture | Lifestyle

World ▶ Europe **US** Americas Asia Australia Middle East Africa Inequality More

Hillary Clinton

This article is more than **8 years old**

Coin tosses used to determine county delegates in Clinton-Sanders race

Obscure party rule can be called upon to decide a tied result – and its use shows how close the Democratic race is

Were the coins ‘rigged’?

- Coin flips were used to decide the allocation of 6 delegates in Iowa (2016)
- If we’re testing whether these coin flips were somehow rigged, what are our hypotheses?
- How can we test these hypotheses?
  – What is our expectation given H_0?
  – One-tail test or two-tail test?
  – What level of Type I error should we accept?


The Binomial Distribution with n = 6 and p = 0.5
Probability
0.3
0.25
0.2
0.15
0.1
0.05
0
0
1
2
3
4
5
6
Number of Successes