# Decision Error

Grinnell College

April 1, 2024

Review

# Decision Making

For now, let's not worry about $p$-values (*we will revist), instead, let's go back to binary thinking since, in actuality, we must ultimately decide between one of two decisions:

1. There is sufficient evidence to reject $H_0$
2. There is *not* sufficient evidence to reject $H_0$

# Decision Making

Just as our confidence intervals were correct or incorrect, so to may be our decision regarding $H_0$. In this case, however, there are two distinct ways in which our decision can be incorrect:

1. $H_0$ is *TRUE* (i.e., there is no effect), yet we reject anyway
2. $H_0$ is *FALSE* (i.e., there is an effect), yet we fail to reject it

# Decision Making

These two types of errors are known as Type I and Type II errors, respectively:

1. $H_0$ is *TRUE* (i.e., there is no effect), yet we reject anyway
   - Type I error
   - False positive
   - Evidence leads to wrong conclusion
2. $H_0$ is *FALSE* (i.e., there is an effect), yet we fail to reject it
   - Type II error
   - False negative
   - Not enough evidence to conclude

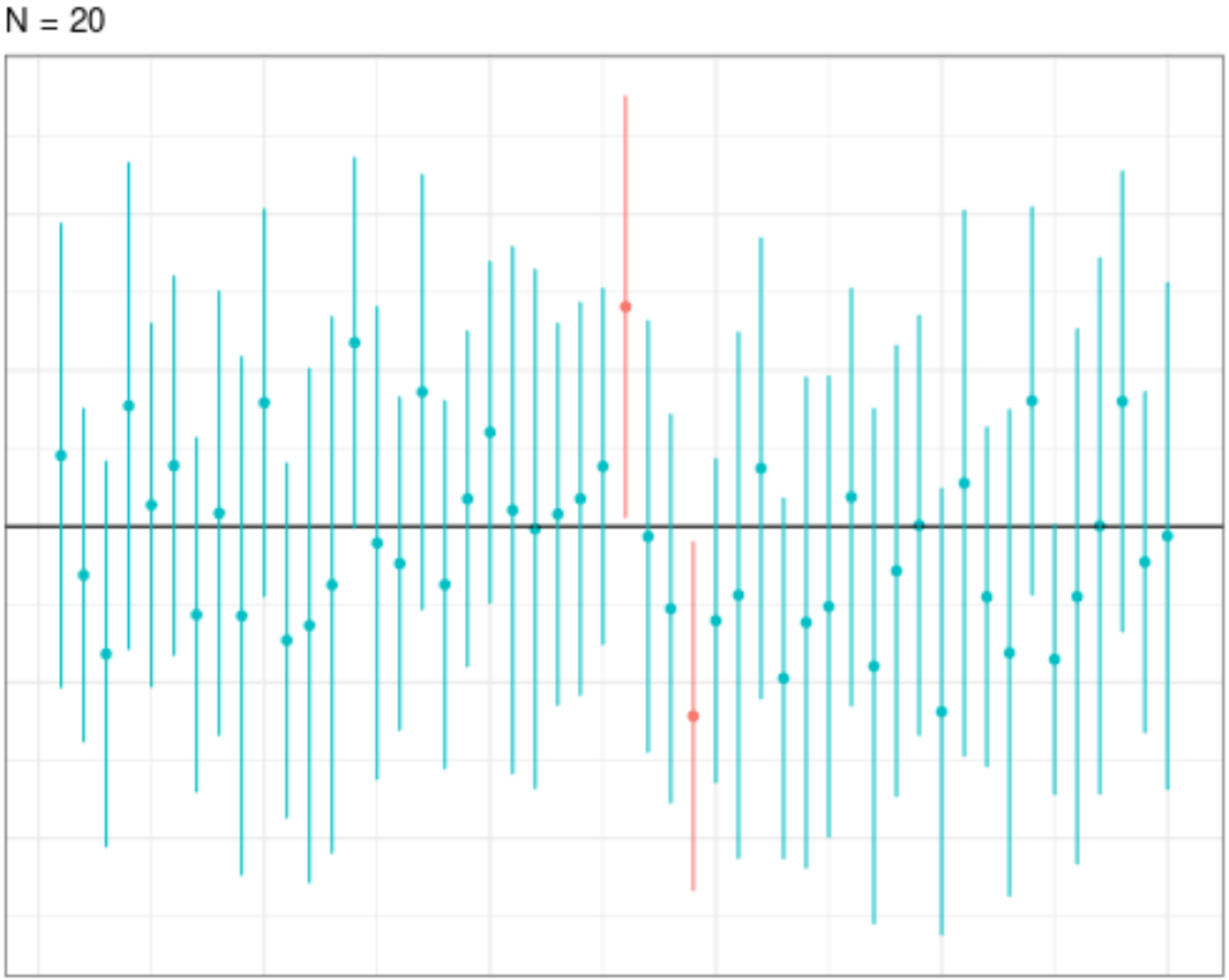
# Type I Errors

A Type I error describes a situation in which we incorrectly identify a null effect:

- Conclude that an intervention works when it does not
- Conclude that there is a relationship between two variables when there are not

A Type I error will occur, for example, when our constructed confidence does not contain $\mu_0$ when in actuality it should

# Type I Errors



# Type I Error Rate

We can control the rate at which we commit Type I errors with adjusting the *level of significance*, denoted $\alpha$.

This is also called the *Type I error rate*

The Type I error rate has a *one-to-one* correspondence with our confidence intervals – a 95% confidence interval will permit a Type I error 5% of the time, corresponding to $\alpha = 0.05$

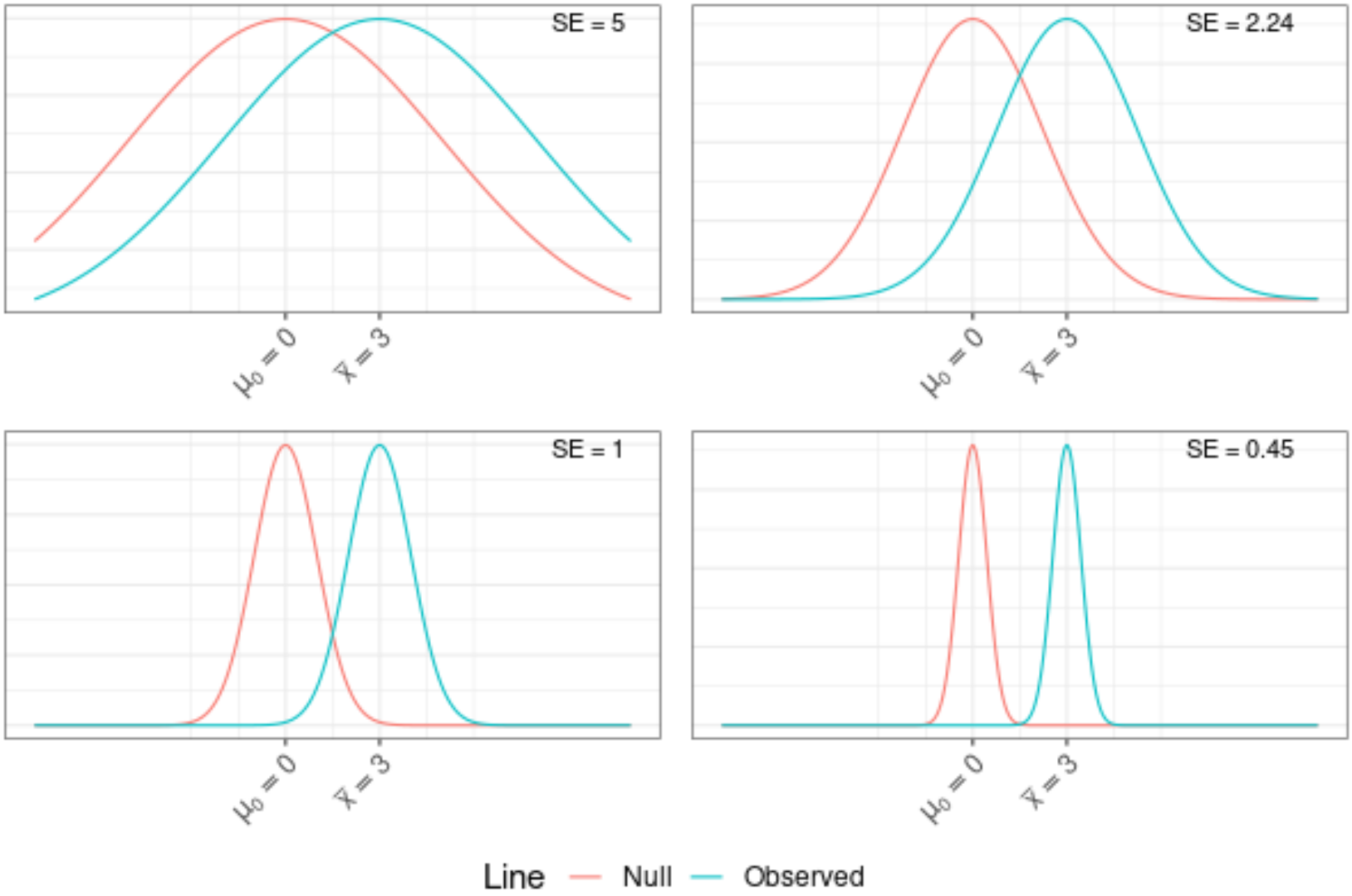
# Type II Errors

A Type II error describes a situation in which the null hypothesis is false, yet based on the evidence gathered we fail to reject it:

- An intervention has a clinical effect, but it is not detected
- An email is considered spam, but the filter does not detect it

Typically, a Type II error is the result of one or more factors:

- Too few observations in our sample
- The population has large variability
- The effect size is small


SE = 5
SE = 2.24
SE = 1
SE = 0.45
$\mu_0 = 0$
$\bar{x} = 3$
Line
Null
Observed

# Type II Error Rate

The Type II error rate is typically denoted $\beta$

More frequently, we consider the rate at which Type II errors do not occur $(1 - \beta)$, a term we refer to as *power*

A study that is unable to detect a true effect is said to be *underpowered*

# Drawing Conclusions

As we never truly know whether $H_0$ is correct or not, we must simultaneously be prepared to combat both types of error

| Test Result | True State of Nature | |
|---|---|---|
| | $H_0$ True | $H_0$ False |
| Fail to reject $H_0$ | Correct $(1-\alpha)$ | Incorrect Type II Error $(\beta)$ |
| Reject $H_0$ | Incorrect Type I Error $(\alpha)$ | Correct $(1-\beta)$ |

- Type I error $= P(\text{Reject } H_0 | H_0 \text{ true}) =$ false alarm
- Type II error $= P(\text{Fail to reject } H_0 | H_A \text{ true}) =$ missed opportunity

# Review

Based on the evidence observed, we will ultimately make one of two decisions:

1. Reject $H_0$
2. Fail to reject $H_0$

Depending on the true state of $H_0$, we can be incorrect in two ways:

1. Type I Error ($\alpha$): $H_0$ is true, yet we reject anyway
2. Type II Error ($\beta$): $H_0$ is false, yet we fail to reject it

We will discuss more how these errors are related on Wednesday