Additional topics in testing

Stat 120

April 26 2023

Question 1

A new blood thinning drug is being tested against the current drug in a double-blind experiment. Is there evidence that the mean blood thinness rating is higher for the new drug? Using n for the new drug and o for the old drug, which of the following are the null and alternative hypotheses?

A. $\mathrm{H_0} : \mu_{\mathrm{n}} > \mu_0 \quad \text{Vs} \quad \mathrm{H_a} : \mu_{\mathrm{n}} = \mu_{\mathrm{o}}$

B. $H_0 : \mu_n = \mu_0 \quad \text{Vs} \quad H_2 : \mu_n \neq \mu_0$

C. $\mathrm{H_0} : \mu_{\mathrm{n}} = \mu_0 \quad \text{Vs} \quad \mathrm{H_a} : \mu_{\mathrm{n}} > \mu_0$

D. $\mathrm{H_0} : \bar{x}_n = \bar{x}_o \quad \text{Vs} \quad \mathrm{H_a} : \bar{x}_n \neq \bar{x}_o$

E. $\mathrm{H_0} : \bar{x}_n = \bar{x}_o \quad \text{Vs} \quad \mathrm{H_a} : \bar{x}_n > \bar{x}_o$

▶ Click for answer

Question 2

A new blood thinning drug is being tested against the current drug in a double-blind experiment, and the hypotheses are:

$$\mathrm{H_0} : \mu_{\mathrm{n}} = \mu_{\mathrm{o}} \quad \mathrm{vS} \quad \mathrm{H_a} : \mu_{\mathrm{n}} > \mu_{\mathrm{o}}$$

What does a Type I Error mean in this situation?

A. We reject $\mathbf{H_0}$

B. We do not reject $\mathbf{H_0}$

C. We find evidence the new drug is better when it is really not better.

D. We are not able to conclude that the new drug is better even though it really is.

E. We are able to conclude that the new drug is better.

▶ Click for answer

Question 3

A new blood thinning drug is being tested against the current drug in a double-blind experiment, and the hypotheses are:

$$\mathrm{H_0}: \mu_{\mathrm{n}} = \mu_{\mathrm{o}} \quad \text{vS} \quad \mathrm{H_a}: \mu_{\mathrm{n}} > \mu_{\mathrm{o}}$$

What does a Type II Error mean in this situation?

A. We reject $\mathbf{H_0}$

B. We do not reject $\mathbf{H_0}$

C. We find evidence the new drug is better when it is really not better.

D. We are not able to conclude that the new drug is better even though it really is.

E. We are able to conclude that the new drug is better.

▶ Click for answer

Question 4

A new blood thinning drug is being tested against the current drug in a double-blind experiment, and the hypotheses are:

$$\mathrm{H}_0 : \mu_\mathrm{n} = \mu_\mathrm{o} \quad \text{vS} \quad \mathrm{H}_\mathrm{a} : \mu_\mathrm{n} > \mu_\mathrm{o}$$

If the new drug has potentially serious side effects, we should pick a significance level that is:

A. Relatively small (such as 1%)

B. Middle of the road (5%)

C. Relatively large (such as 10%)

▶ Click for answer

Question 5

A new blood thinning drug is being tested against the current drug in a double-blind experiment, and the hypotheses are:

$$\mathrm{H}_0 : \mu_{\mathrm{n}} = \mu_{\mathrm{o}} \quad \mathrm{vS} \quad \mathrm{H}_{\mathrm{a}} : \mu_{\mathrm{n}} > \mu_{\mathrm{o}}$$

If the new drug has no real side effects and costs less than the old drug, we might pick a significance level that is:

A. Relatively small (such as 1%)

B. Middle of the road (5%)

C. Relatively large (such as 10%)

▶ Click for answer

Probability of Type II Error

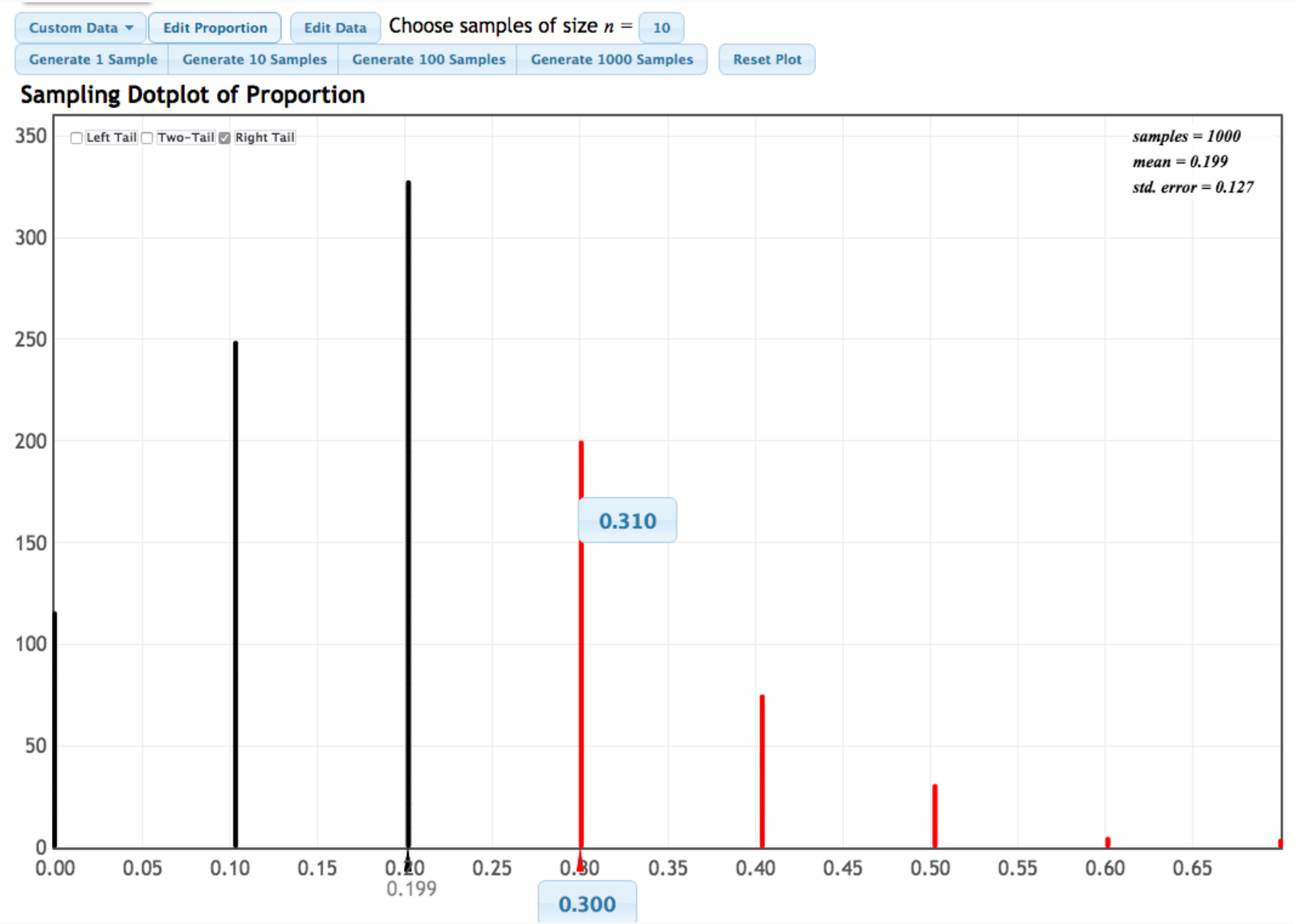
How can we reduce the probability of making a Type II Error (not rejecting a false null)?

a) Decrease the sample size

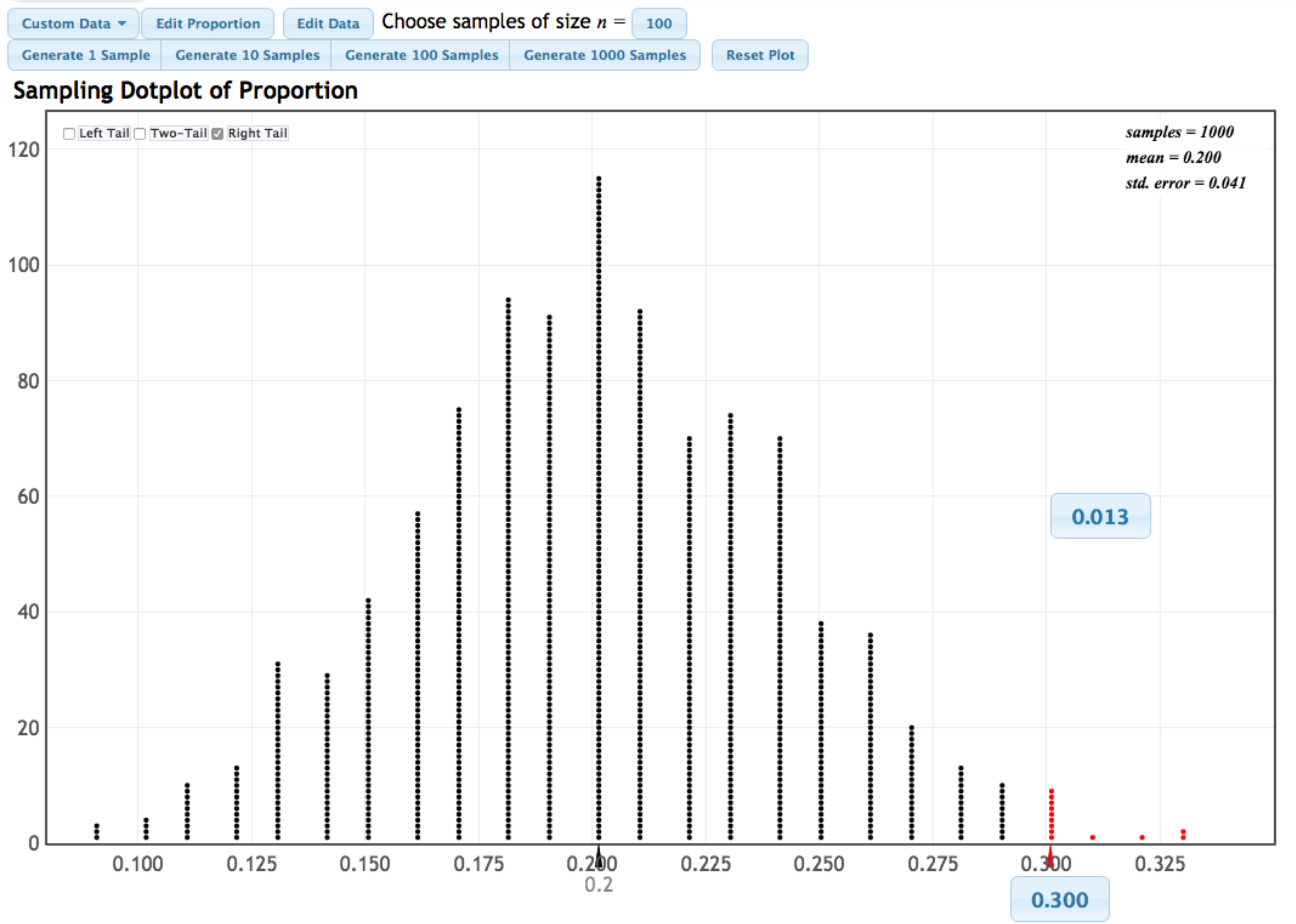
b) Increase the sample size

▶ Click for detailed answer

Randomization distribution with $n = 10$



Randomization distribution with $n = 100$



Bootstrap Distribution

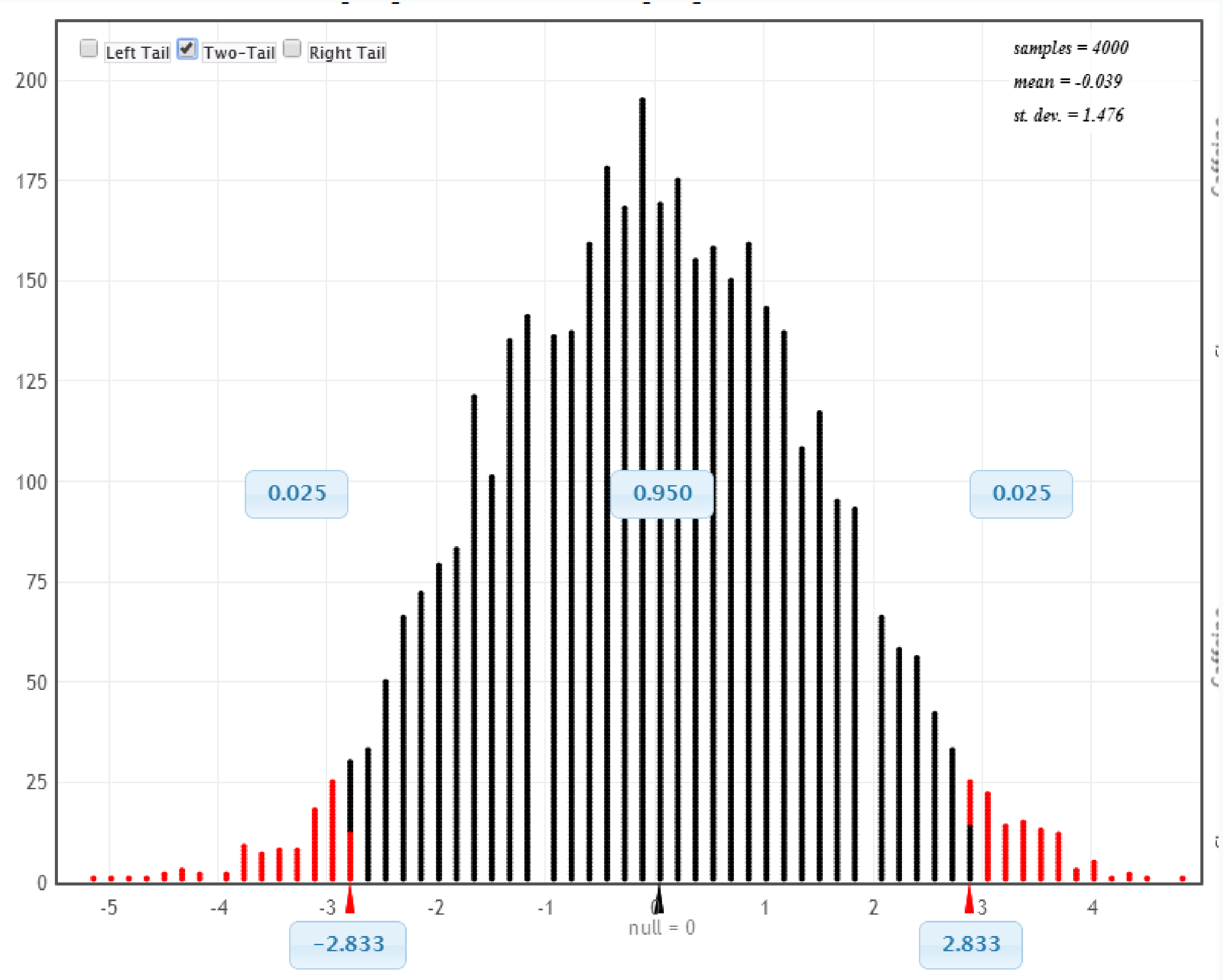
- Our best guess at the distribution of sample statistics
- Centered around the observed sample statistic
- Simulate sampling from the population by resampling from the original sample

Randomization Distribution

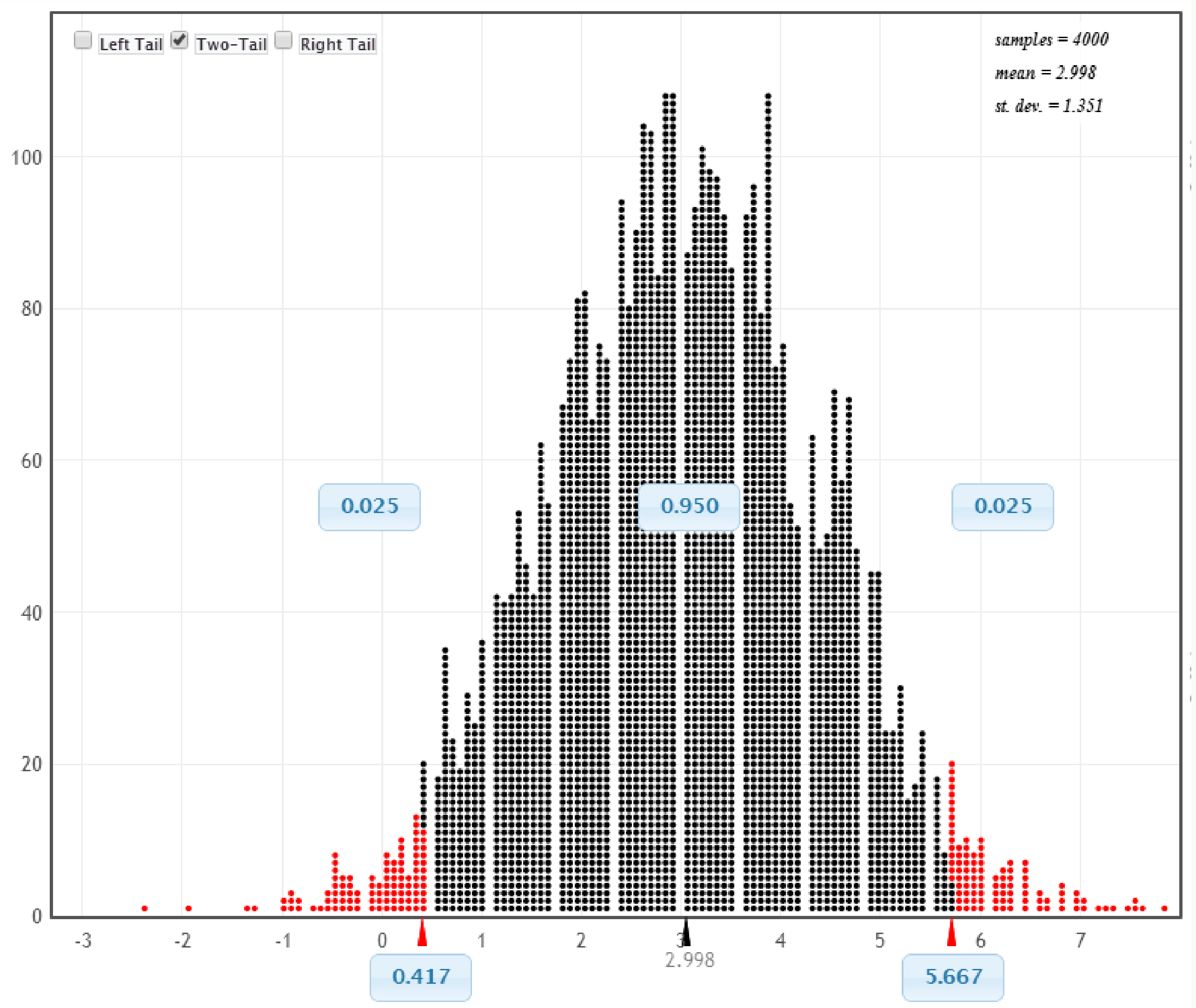
- Our best guess at the distribution of sample statistics, if H_0 were true
- Centered around the null hypothesized value
- Simulate samples assuming H_0 were true

Big difference: a randomization distribution assumes H_0 is true, while a bootstrap distribution does not

Example: bootstrap or randomization?



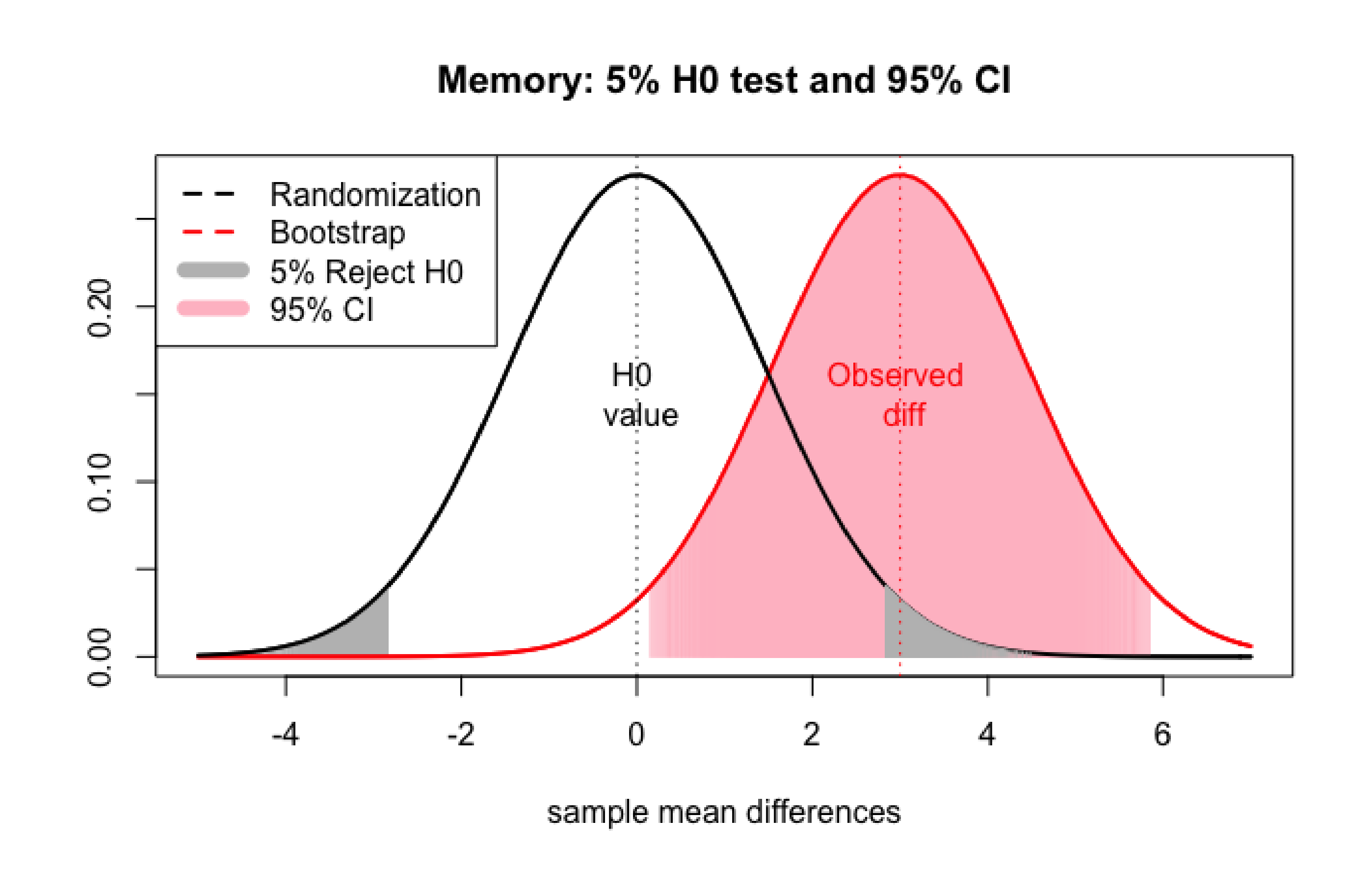
Example: bootstrap or randomization?



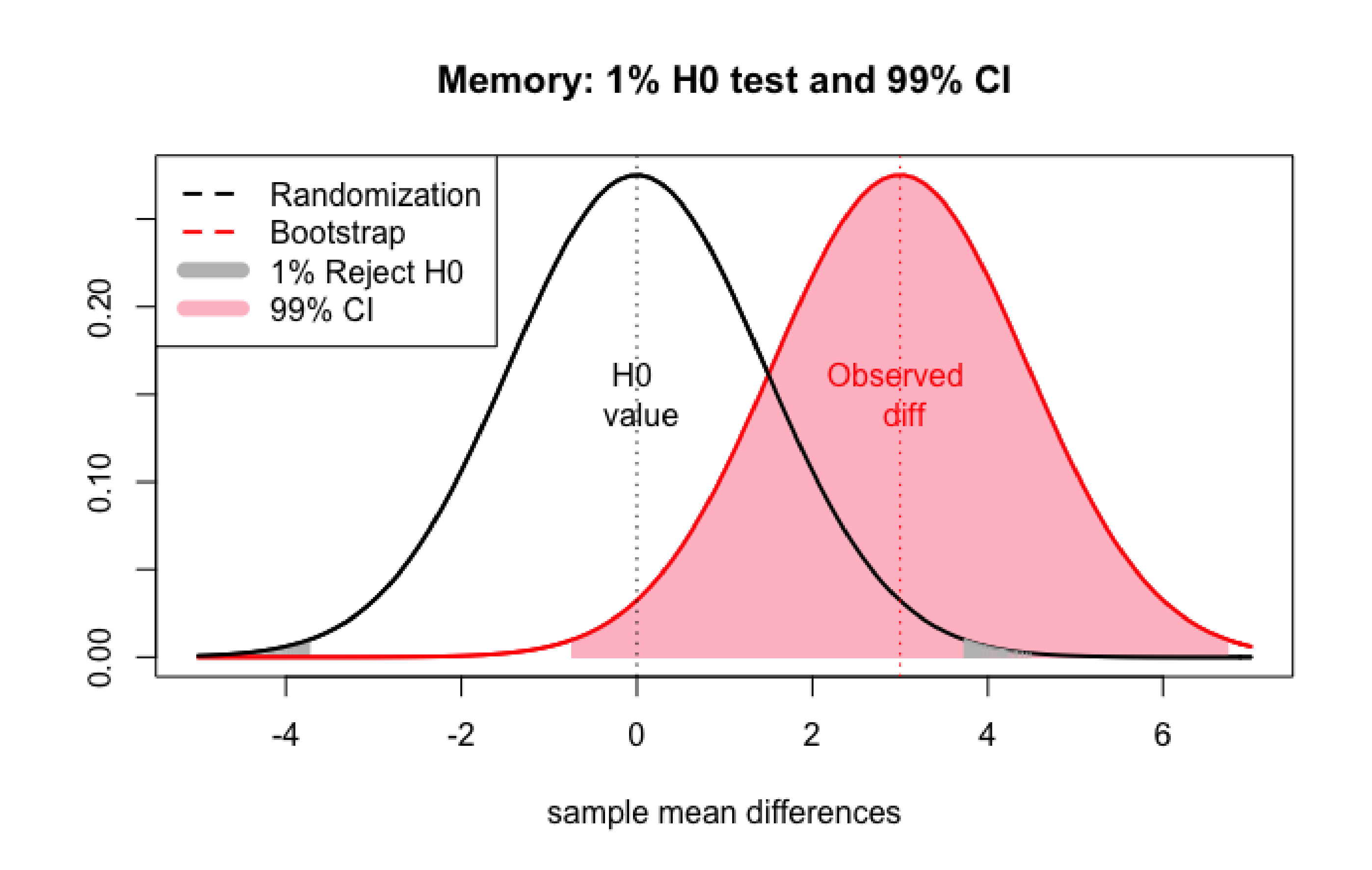
Using confidence intervals for hypothesis testing

If a 95% CI contains the parameter in H_0, then a two-tailed test should not reject H_0 at a 5% significance level.

If a 95% CI misses the parameter in H_0, then a two-tailed test should reject H_0 at a 5% significance level.



The 95% confidence interval misses null difference of 0. Reject the null at 5% level



The 99% confidence interval contains null difference of 0. Do not reject the null at 1% level

Multiple Testing/Comparison

When multiple hypothesis tests are conducted, the chance that at least one test incorrectly rejects a true null hypothesis increases with the number of tests.

- ***If the null hypotheses are all true, α of the tests will yield statistically significant results just by random chance.***

Diet and Sex of Baby

- *Are certain foods in your diet associated with whether or not you conceive a boy or a girl?*
- *To study this, researchers asked women about their eating habits, including asking whether or not they ate 133 different foods regularly*

http://www.newscientist.com/article/dn13754-breakfast-cereals-boost-chances-of-conceiving-boys.html

Diet and Sex of Baby

For each of the 133 foods studied, a hypothesis test was conducted for a difference between mothers who conceived boys and girls in the proportion who consume each food

What are the null and alternative hypotheses?

Compare two populations: mothers who have boys vs. mothers who have girls

p_b ***: proportion of mothers who have boys that consume the food regularly***

p_g ***: proportion of mothers who have girls that consume the food regularly***

$$\mathrm{H}_0 : p_b = p_g$$
$$\mathrm{H}_a : p_b \neq p_g$$

Diet and Sex of Baby

A significant difference was found for breakfast cereal (mothers of boys eat more), prompting the headline

"Breakfast Cereal Boosts Chances of Conceiving Boys"

How might you explain this?

Random chance: several tests (about 6 or 7) are going to be significant, even if no differences exist

Diet and Sex of Baby

If there are NO differences (all 133 null hypotheses are true), about how many significant differences would be found using α = 0.05?

$$133 \pm 0.05 = 6.65$$

Expect about 6-7 statistically significant foods even if the rate of food consumption is equal for women who have boys and women who have girls

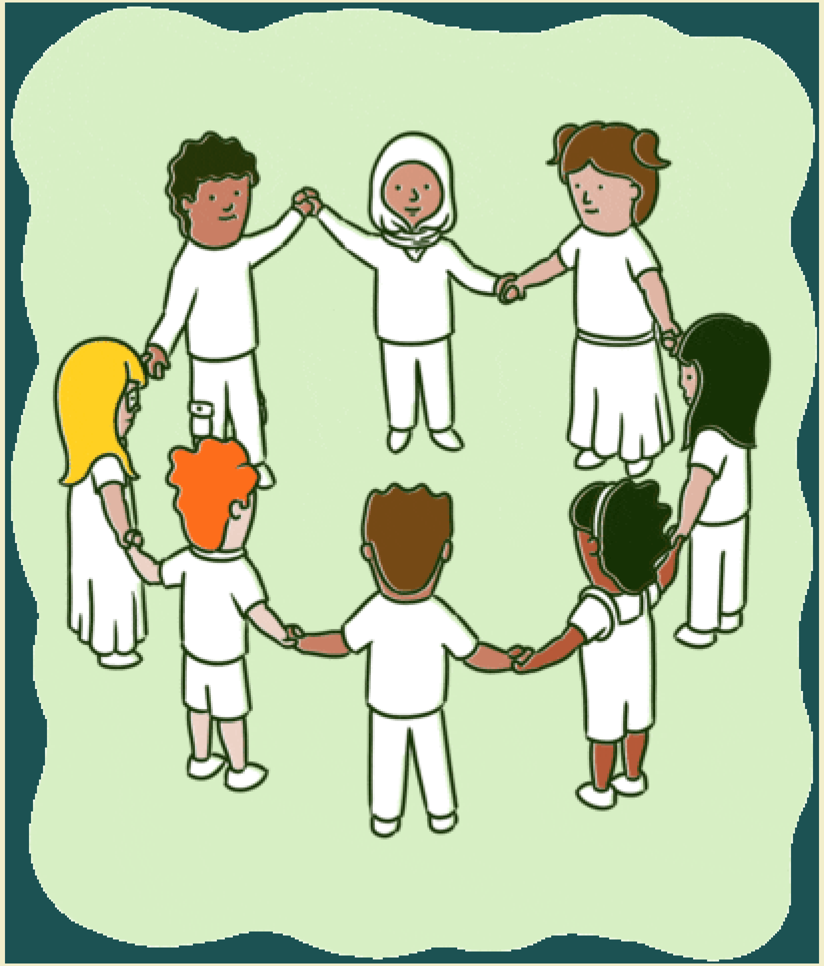
Multiple Comparisons

The most important thing is to be aware of this issue, and not to trust claims that are obviously one of many tests (unless they specifically mention an adjustment for multiple testing)

There are ways to account for this (e.g. Bonferroni's Correction), but these are beyond the scope of this class

YOUR TURN 1

10:00



Please go over the class activity for today and let me know if you have any questions.