Hypothesis Testing withTwo Samples

Two Samples

- Up until now, we’ve been working with single samples
- But it is much more interesting to compare samples
  - Do men and women do the same amount of housework?
  - Do people who exercise live longer than people who don’t?

Two Samples

- Samples can be dependent, or independent
- Dependent samples means that whether someone is in one sample depends on who is in the other sample
- Comparing spouses, for example

Independent Samples

- Independent samples - who is in one sample has nothing to do with who is in a second sample

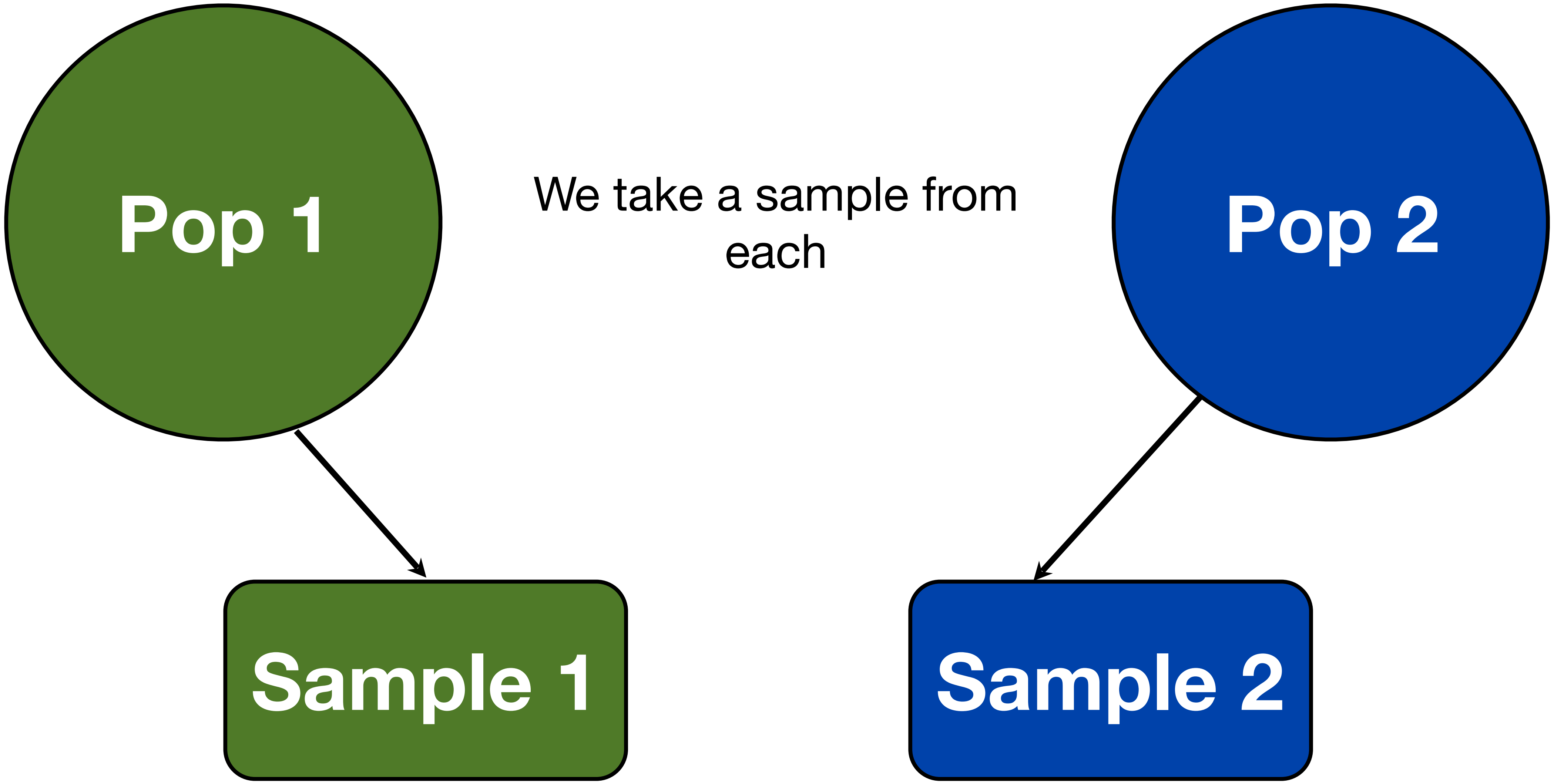
What can we do with two samples?

- Calculate the mean for each
- Calculate the difference in means
- Test whether means are statistically different

Notation

	Sample 1	Sample 2
Population mean	μ_1	μ_2
Sample Mean	$\bar{X}_1$	$\bar{X}_2$
Sample Size	N_1	N_2

Consider two populations



$\bar{X}_1 = 23.5$ Calculate sample means $\bar{X}_2 = 25.2$

$$\bar{X}_1 - \bar{X}_2 = 23.5 - 25.2 = -1.7$$

What are plausible values for the population mean difference?

Dependent Variable

- Dependent variable is the variable you think is influenced by other factors
- Value depends on values of independent variable(s)

Independent Variable

- Variable you think influences the outcome of the dependent variable
- In this case, defines what distinguishes the two samples

Confidence Intervals

- We can construct a confidence interval for the difference between two populations
- But to do so, we need to know what the sampling distribution is for the difference between two means

Sampling Distribution

- If we take multiple samples from each of the two populations and compute their means, we'll eventually create a distribution of possible differences
- This is exactly like the sampling distribution for a single mean

Assumptions

- Large Sample
- Independent random sample
- Interval level dependent variable
- Nominal level independent variable with two categories
- Normally distributed sampling distribution

Sampling Distribution for Difference

Mean: $$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$$

Standard Error: $$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_1^2 + \sigma_2^2}$$

$$= \sqrt{\frac{s_1^2}{N_1 - 1} + \frac{s_2^2}{N_2 - 1}}$$

Confidence Interval

$$(\bar{X}_1 - \bar{X}_2) \pm Z\sigma_{\bar{X}_1 - \bar{X}_2}$$

$$(\bar{X}_1 - \bar{X}_2) \pm Z\sqrt{\frac{s_1^2}{N_1 - 1} + \frac{s_2^2}{N_2 - 1}}$$

Practice!

You are interested in who are better students: anthropologists or sociologists. You collect some data and calculate the following statistics

	Anthro	Sociol
mean GPA	3.2	3.6
st dev	0.8	1.2
n	150	125

Construct a 95% confidence interval for the population mean difference between sociologists' and anthropologists' GPAs.

$$\bar{X}_1 = 3.6$$
$$\bar{X}_2 = 3.2$$
$$s_1 = 1.2$$
$$s_2 = 0.8$$
$$n_1 = 125$$
$$n_2 = 150$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{N_1 - 1} + \frac{s_2^2}{N_2 - 1}}$$
$$= \sqrt{\frac{1.2^2}{125 - 1} + \frac{0.8^2}{150 - 1}}$$
$$= \sqrt{\frac{1.44}{124} + \frac{0.64}{149}}$$
$$= \sqrt{0.0116 + 0.00430}$$
$$= 0.126$$

$$(\bar{X}_1 - \bar{X}_2) \pm Z\sigma_{\bar{X}_1 - \bar{X}_2}$$
$$(3.6 - 3.2) \pm 1.96(0.126)$$
$$0.4 \pm 0.247$$

We are 95% confident that the population mean difference between sociologists' and anthropologists' GPAs is between 0.153 and 0.647

Practice!

Someone suggests that people who drink coffee make more money than people who don't. You collect some data and calculate the following statistics:

	Coffee Drinkers	Coffee Abstainers
mean income	34,200	33,100
st dev	7000	6500
n	200	200

Construct a 95% confidence interval for the population mean difference in income between coffee drinkers and abstainers

$$\bar{X}_1 = 34,200$$
$$\bar{X}_2 = 33,100$$
$$s_1 = 7000$$
$$s_2 = 6500$$
$$n_1 = 200$$
$$n_2 = 200$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_1^2}{N_1-1} + \frac{s_2^2}{N_2-1}}$$

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{7000^2}{200-1} + \frac{6500^2}{200-1}}$$
$$= \sqrt{\frac{49000000}{199} + \frac{42250000}{199}}$$
$$= \sqrt{246231.16 + 212311.56}$$
$$= 677.16$$

$$(\bar{X}_1 - \bar{X}_2) \pm Z\sigma_{\bar{X}_1-\bar{X}_2}$$
$$(34200 - 33100) \pm 1.96(677.16)$$
$$1100 \pm 1327.23$$

We can be 95% confident that the population mean difference in income between coffee drinkers and coffee abstainers is between $-227.23 and $2427.23

Confidence Intervals

- Notice that in the previous example, 0 was a plausible value for the population difference in means.
- This implies that if we do a hypothesis test, we won't find statistical support for the claim that coffee drinkers make more than abstainers (zero difference)

Hypothesis Test

- We do a hypothesis test for the difference between means just as we did with only one sample
- This time, though, the test statistic we calculate is Z, and is representative of the difference in means

1. Assumptions

- Large sample (small samples use different equations)
- Independent random samples
- Normally distributed sampling distribution
- Interval level dependent variable, nominal level independent variable

2. Null Hypothesis

- When testing difference, the null is always no difference:

$$H_0 : \mu_1 = \mu_2$$

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

2. Alternative Hypothesis

- Two-tailed: $H_A : \mu_1 \neq \mu_2 \quad \mu_1 - \mu_2 \neq 0$

- One-tailed: $H_A : \mu_1 > \mu_2 \quad \mu_1 - \mu_2 > 0$

  $H_A : \mu_1 < \mu_2 \quad \mu_1 - \mu_2 < 0$

3. Test Statistic

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

Hypothesis Test

- From here, we could continue on as usual, comparing p-value to alpha and/or test statistic to the critical value

Practice!

Using the same data as before, test for whether coffee drinkers make more than coffee abstainers, using an alpha of 0.05

	Coffee Drinkers	Coffee Abstainers
mean income	34,200	33,100
st dev	7000	6500
n	200	200

1. Assumptions

- Large sample
- Independent random samples
- Normally distributed sampling distribution
- Interval level dependent variable, nominal level independent variable

2. Hypotheses

- Null: coffee drinkers and abstainers make the same income

$$H_0 : \mu_1 = \mu_2$$

- Alternative: coffee drinkers make more than coffee abstainers

$$H_A : \mu_1 > \mu_2$$

3. Test Statistic

$$\bar{X}_1 = 34200$$
$$\bar{X}_2 = 33100$$
$$\sigma_{\bar{X}_1 - \bar{X}_2} = 677.16$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

$$Z = \frac{(34200 - 33100) - 0}{677.16}$$
$$= \frac{1100}{677.16}$$
$$Z = 1.624$$

4. P-Value

Second decimal place in z										
0.09	*0.08*	*0.07*	*0.06*	*0.05*	*0.04*	*0.03*	*0.02*	*0.01*	*0.00*	*z*
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	*−1.9*
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	*−1.8*
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	*−1.7*
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	*−1.6*
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	*−1.5*

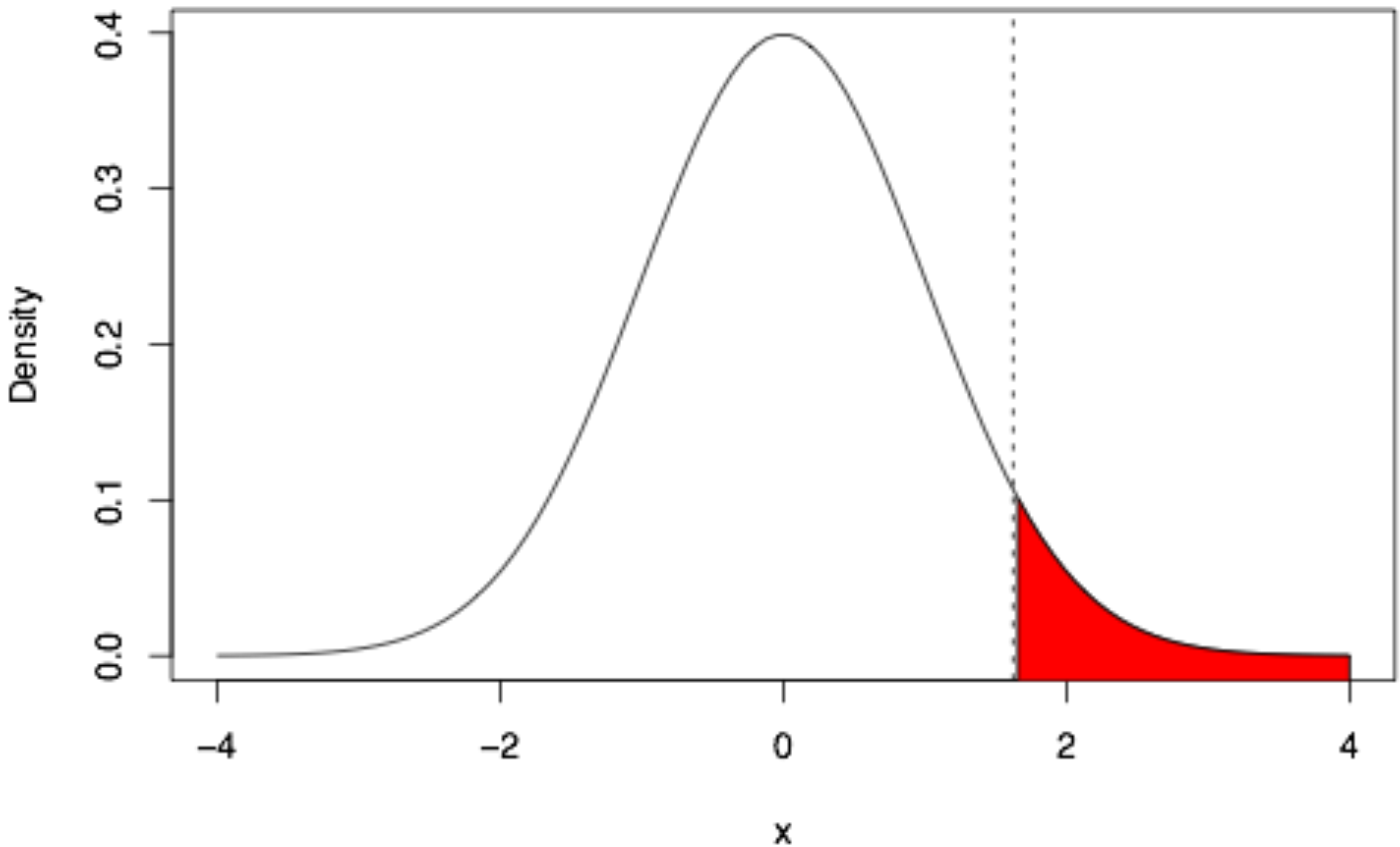
So our p-value is 0.0526. Since this is a one tailed test, we do not need to multiply by two. Our p-value is greater than our alpha (0.05)

4. Critical Region

- For a one tailed test, with a large sample and an alpha of 0.05, our critical value is 1.65

- Our test statistic is 1.624, which is smaller than the critical value

- Our test statistic does not fall in the critical region



Density
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
x

5. Conclusion

- With a p-value of 0.0526 and a test statistic of 1.624, we fail to reject the null hypothesis.
- There is not enough evidence to claim that coffee drinkers make more money than coffee abstainers

Break

Two Proportions

Two Proportions

- When both variables are nominal
- Ex: College graduate and opinion on the death penalty

Notation

	Sample 1	Sample 2
Population Proportion	π_1	π_2
Sample Proportion	$\hat{\pi}_1$	$\hat{\pi}_2$
Sample Size	N_1	N_2

Sampling Distribution

- Normally distributed as long as you have large samples
- Often people say at least 5 observations in each category

Sampling Distribution

- Mean

$$\mu_{\hat{\pi}_1 - \hat{\pi}_2} = (\hat{\pi}_1 - \hat{\pi}_2)$$

- Standard Error

$$\sigma_{\hat{\pi}_1 - \hat{\pi}_2} = \sqrt{\frac{\pi(1-\pi)}{N_1} + \frac{\pi(1-\pi)}{N_2}}$$

Standard Error

$$\sigma_{\hat{\pi}_1 - \hat{\pi}_2} = \sqrt{\frac{\pi(1-\pi)}{N_1} + \frac{\pi(1-\pi)}{N_2}}$$

- But what if we don't know π (probability of success in the population)?

- Pooled estimate:

$$\hat{\pi} = \frac{N_1\hat{\pi}_1 + N_2\hat{\pi}_2}{N_1 + N_2}$$

Standard Error

- Remember: we assume the null hypothesis is true until proven otherwise

- The null hypothesis in this case is:

$$\pi_1 = \pi_2$$

- For that to be true, given the observed successes and cases, then the computed, pooled population estimate is the best estimate for the population proportion

Practice!

The following data come from the 2010 GSS, in which it was asked: is gay sex wrong?:

	Wrong	Not Wrong	Total
Male	349	210	559
Female	355	309	664

Test the hypothesis that women are more supportive of gay sex than men, using an alpha of 0.05

1. Assumptions

- Independent random samples
- Normally distributed sampling distribution
- Nominal dependent and independent variables

2. Hypotheses

- Null: $H_0\colon\ \pi_M = \pi_F$
- Alternative: $H_A\colon\ \pi_M < \pi_F$

3. Test Statistic

First, we need to calculate some proportions:

$$\hat{\pi}_M = \frac{210}{559} = 0.376 \qquad \hat{\pi}_F = \frac{309}{664} = 0.465$$

$$\hat{\pi} = \frac{N_M \hat{\pi}_M + N_F \hat{\pi}_F}{N_M + N_F}$$

$$\begin{aligned} \hat{\pi} &= \frac{559 * 0.376 + 664 * 0.465}{559 + 664} \\ &= \frac{519}{1223} \\ &= 0.424 \end{aligned}$$

$$\sigma_{\hat{\pi}_1-\hat{\pi}_2} = \sqrt{\frac{\pi(1-\pi)}{N_1} + \frac{\pi(1-\pi)}{N_2}}$$

$$\begin{aligned}\sigma_{\hat{\pi}_1-\hat{\pi}_2} &= \sqrt{\frac{0.424(1-0.424)}{559} + \frac{0.424(1-0.424)}{664}} \\ &= \sqrt{\frac{0.2442}{559} + \frac{0.2442}{664}} \\ &= 0.02837\end{aligned}$$

$$\begin{aligned}Z &= \frac{(\hat{\pi}_F - \hat{\pi}_M) - 0}{\sigma_{\hat{\pi}_F-\hat{\pi}_M}} \\ &= \frac{(0.465 - 0.376) - 0}{0.02837} \\ Z &= 3.137\end{aligned}$$

4. P-Value

- The p-value associated with a test statistic of 3.137 is essentially 0, which is less than our alpha of 0.05

4. Critical Region

- For a one tailed test with alpha of 0.05, the critical value is 1.65
- Our test statistic is 3.137, which is greater than the critical value
- Therefore, our test statistic lies within the critical region

5. Conclusion

- With a p-value of essentially 0 and a test statistic of 3.137, we reject the null hypothesis in favor of the alternative

- Women are more likely to be supportive of homosexuality than men.

Worksheet & Break