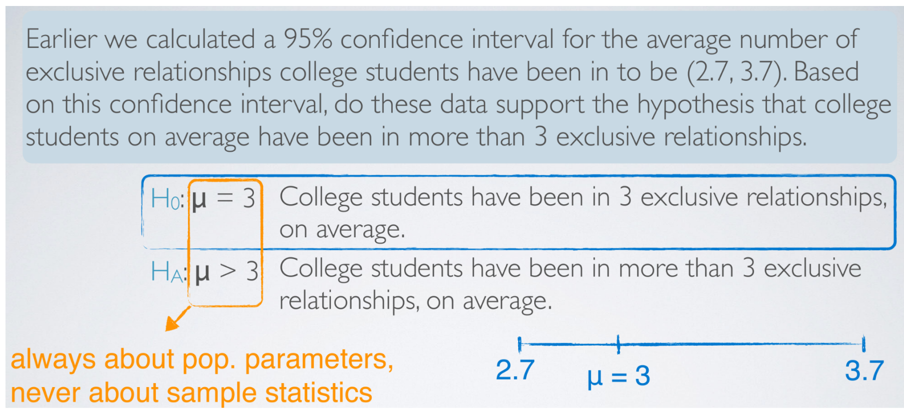
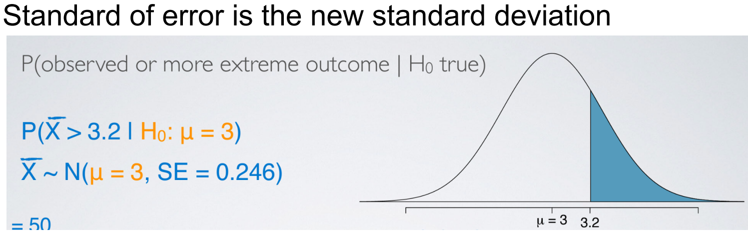
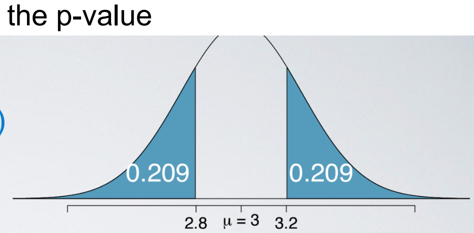
Hypothesis Testing	
Hypotheses	Null Hypothesis (H0) - claim to be tested Alternative Hypothesis (HA) - alternative claim, range of values that would reject/fail to reject null hypothesis Always about population parameters, never about sample statistics
P-value calculation	Probability of the observed value to be true given that null hypothesis is true.  Let's say we want to know the probability of more than 3.2 relationships. - x-bar>3.2 Step 1: Find the distribution of sample statistics (x-bar) using Central Limit Theorem - Population mean is the center - Standard of error is the new standard deviation -  Step 2: Identify what you're calculating for, the percent (area under the curve) that is above the test statistic of 3.2 Step 3: Find the z-score of the test statistics - z-score shows how many standard deviation from the mean

	- $Z = \frac{3.2 - 3}{0.246} = 0.81$ - This indicates that 3.2 is 0.81 standard deviation from the mean. - Looking at the z-table, it says that area under the curve from the left ~79% Step 3: Find x-bar>3.2 - Since we are looking for probability that the test statistic is greater than 3.2 (aka z > 0.81), then the area under the curve that's to the right is 21% - $p\text{-value} = P(Z > 0.81) = 0.209$ Step 4: Contextualize the result - High probability that a random of sample of 50 college students would yield a sample mean of 3.2 or higher simply by chance (or sampling variability) Therefore the probability that the alternative hypothesis that people have greater than 3.2 partners has a 20.9% chance of coming true given than the null hypothesis is true. Since both can occur, then we failed to reject the null hypothesis
P-value interpretation	P-value is low (<5% significance level) - reject H0 - Probability of observing the data if H0 is true is very unlikely - Something other than random chance is at play P-value is high (>5%) - Fail to reject H0 - Likely to observe the data even if the null hypothesis is true. - Observed data aligns with what you might expect from random chance alone.
Two-sided Test	Might be interested in looking at divergence in any direction and not just one direction. Using example earlier, just double the p-value $P(\bar{X} > 3.2 \text{ OR } \bar{X} < 2.8 \mid H_0: \mu = 3)$  p-value = $= P(Z > 0.81) + P(Z < -0.81)$ $= 2 \times 0.209$ $= 0.418$

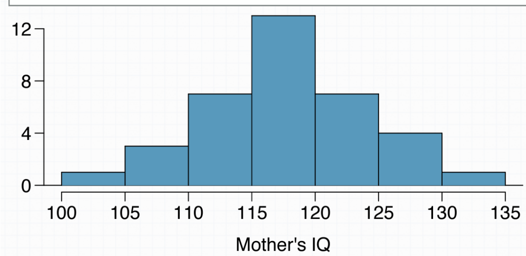
Summarize steps of Hypothesis Testing

Hypothesis testing for a single mean:

1. Set the hypotheses: $H_0 : \mu = null\ value$
   $H_A : \mu\ <\ or\ >\ or\ \neq\ null\ value$
2. Calculate the point estimate: $\bar{x}$
3. Check conditions:
   1. ***Independence:*** Sampled observations must be independent (random sample/assignment & if sampling without replacement, n < 10% of population)
   2. ***Sample size/skew:*** n ≥ 30, larger if the population distribution is very skewed.
4. Draw sampling distribution, shade p-value, calculate test statistic $Z = \frac{\bar{x} - \mu}{SE}, \quad SE = \frac{s}{\sqrt{n}}$
5. Make a decision, and interpret it in context of the research question:
   - If p-value < **α**, reject H_0; the data provide convincing evidence for H_A.
   - If p-value > **α**, fail to reject H_0 the data *do not* provide convincing evidence for H_A.

Hypothesis Testing Example 1:

Researchers investigating characteristics of gifted children collected data from schools in a large city on a random sample of thirty-six children who were identified as gifted children soon after they reached the age of four. In this study, along with variables on the children, the researchers also collected data on their mothers' IQ scores. The histogram shows the distribution of these data, and also provided are some sample statistics.



n	36
min	101
mean	118.2
sd	6.5
max	131

Perform a hypothesis test to evaluate if these data provide convincing evidence of a difference between the average IQ score of mothers of gifted children and the average IQ score for the population at large, which is 100. Use a significance level of 0.01.

Step 1: Set the Hypothesis
- How does the mothers' IQ score compare with the population at large? If the mother's IQ is same as the average population's, then the children are not outside the average IQ.
- u= average IQ score of mothers of gifted children

$H_0: \mu = 100 \qquad H_A: \mu \neq 100$

-
- If mother IQ is greater than 100, children are gifted
- If mother IQ is less than 100, then children are gifted

Step 2. Check the test statistic we are estimating for
- Mother IQ Mean = 118.2

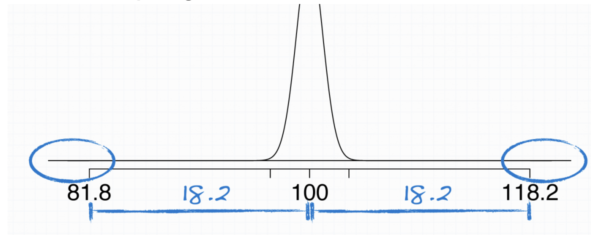
Step 3: Check conditions

1. random & 36 < 10% of all gifted children -> independence
2. n > 30 & sample not skewed -> nearly normal sampling distribution

-

Step 4: Find z-score

- Draw sampling distribution



- Find test statistic distribution using Central Limit Theorem

$$\bar{X} \sim N(\mu = 100, SE = \frac{s}{\sqrt{n}} = \frac{6.5}{\sqrt{36}} \approx 1.083)$$

- Find z-score

$$Z = \frac{118.2 - 100}{1.083} = 16.8$$

- Translate z-score via z-table or qnorm(16.8) to find the percent-> close to 0

Step 5: Find p-value

- P-value = P(z>118.2) and P(z<81.8) = 2* close to 0
- $p\text{-value} \approx 0$

Step 6: Interpret the p-value

p-value is very low -> strong evidence against the null

We reject the null hypothesis and conclude that the data provide convincing evidence of a difference between the average IQ score of mothers of gifted children and the average IQ score for the population at large.

Significance	
Point Estimation Notation	sample mean $\bar{x}$ difference between sample means $\bar{x}_1 - \bar{x}_2$ sample proportion $\hat{p}$ difference between sample proportions $\hat{p}_1 - \hat{p}_2$ Point estimate = difference between sample means. Doing the hypothesis testing but with difference between two averages. Point estimates are unbiased - aka a good estimate that does not naturally over/under estimate the parameter
Point Estimation Calculation	Confidence Interval: $point\ estimate \pm z^{\star} \times SE$ Z-score for Point estimates $Z = \dfrac{point\ estimate - null\ value}{SE}$
Point Estimation Example:	A 2010 Pew Research foundation poll indicates that among 1,099 college graduates, 33% watch The Daily Show (an American late-night TV show). The standard error of this estimate is 0.014. Estimate the 95% confidence interval for the proportion of college graduates who watch The Daily Show. P-hat = point estimate = 0.33 Standard error (given) = 0.014 Step 1: Confidence Interval for Point Estimation formula $\hat{p} \pm z^{*}\ SE$ Step 2: Find critical point aka z* of 95% confidence interval - Confidence level is 95%, so 5% margin of error - Each tail is 2.5%, so area to the left is 0.975 - The z-score that corresponds to 0.975 probability using the z-table is 1.96

<table>
<tr>
<td></td>
<td>
Step 3: Find the confidence interval
$0.33 \pm 1.96 \times 0.014$
0.33 ± 0.027
$(0.303, 0.357)$
</td>
</tr>
<tr>
<td>Point Estimate Example:</td>
<td>
The 3rd NHANES collected body fat percentage (BF%) and gender data from 13,601 subjects ages 20 to 80. The average BF% for the 6,580 men in the sample was 23.9, and this value was 35.0 for the 7,021 women. The standard error for the difference between the average male and female BF%s was 0.114. Do these data provide convincing evidence that men and women have different average BF%s. You may assume that the distribution of the point estimate is nearly normal.

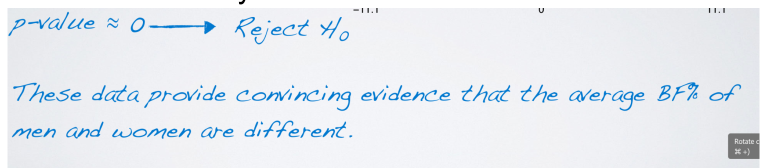
Step 1: Set the hypothesis / write down info

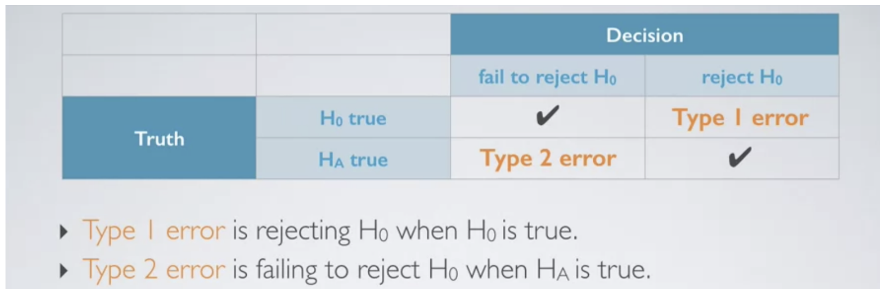
$H_0: \mu_{men} = \mu_{women}$ $H_A: \mu_{men} \neq \mu_{women}$
- SE = 0.114
- X-men = 23.9
- X-women = 35
Step 2: Find point estimate (difference between sample means)
$\bar{X}_{men} - \bar{X}_{women} = 23.9 - 35 = -11.1$

Step 3: Check Conditions
- Independent - yes because safe to assume that 13.6K is < total population
- Given that it's nearly normal

Step 4: Draw point estimate distribution
-
- Mean of distribution is set to 0
- +- 11.1 set as the test statistics on either end of 0

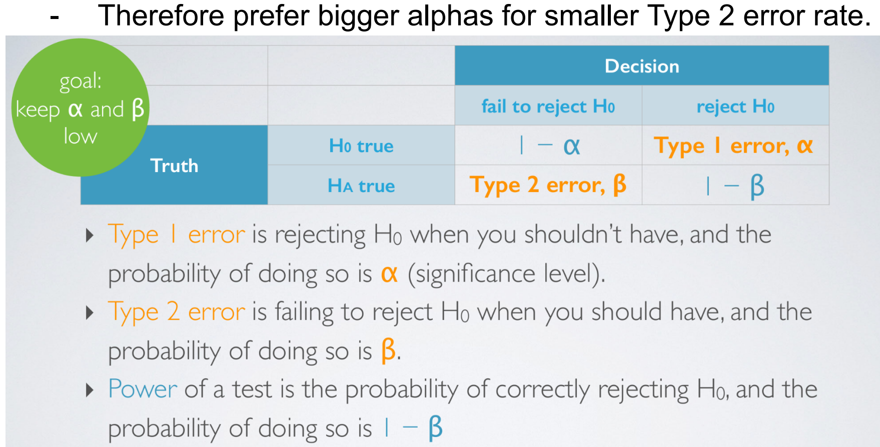
Step 5: Find z-score
- Draw distribution
$H_0: \mu_{men} = \mu_{women} \rightarrow \mu_{men} - \mu_{women} = 0$
$H_A: \mu_{men} \neq \mu_{women}$
- Null hypothesis is equal to 0
$Z = \dfrac{-11.1 - 0}{0.114} \approx -97.36$
- The score (pos/neg doesn't matter) shows the percent to the left and it's really high, so 1-z-table percent to show the right is really low and nearly 0
</td>
</tr>
</table>

<table>
<tr><td></td><td>Step 6: Find p-value
- P-value = 2* nearly 0 = 0


</td></tr>
<tr><td>Decision Errors (Type 1 and Type 2)</td><td>


H0 = defendant is innocent
HA = defendant is guilty

Type 1: rejected H0 but shouldn't have (guilty when innocent)
Type 2: failed to reject H0 but shouldn't have (innocent when guilty)
</td></tr>
<tr><td>Error Rate</td><td>Generally Type 2 is considered worse than Type 1.

Type 1 Error Rate: (alpha=5%)
- Want to be cautious about rejecting H0 when it is right,
- There is about 5% chance of rejecting H0 when H0 is true
- Therefore prefer smaller alphas
Type 2 Error Rate: (alpha = 5%)
- Want to be cautious about approving H0 when it is actually false
- There is about a 95% chance of approving H0 when H0 is false.
- Therefore prefer bigger alphas for smaller Type 2 error rate.


Where B depends on the effect size (difference between point
</td></tr>
</table>

	estimate and null value) - Bigger effect size, bigger B, higher chance of Type 2 error
Significance vs Confidence Level	- A two sided hypothesis with threshold of α is equivalent to a confidence interval with $CL = 1 - \alpha$. - A one sided hypothesis with threshold of α is equivalent to a confidence interval with $CL = 1 - (2 \times \alpha)$. - If H_0 is rejected, a confidence interval that agrees with the result of the hypothesis test should not include the null value. - If H_0 is failed to be rejected, a confidence interval that agrees with the result of the hypothesis test should include the null value. - If H0 is rejected….it makes sense, if it's rejected then it null value should not fall within the confidence interval range or else we're saying H0 is acceptable
Statistical vs Practical Significance.	