***COURSERA: STATS W/ R SPECIALIZATION***

***COURSE 2 - Inference***

**WEEK 2- Inference and Significance**

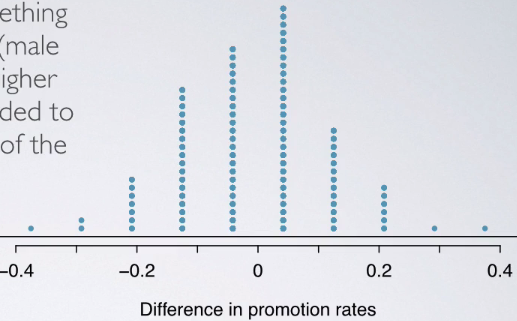
***4.2 Hypothesis Testing***

**Another Introduction to Inference**

* Remember:

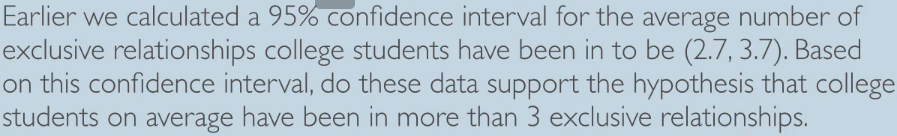


* Experiment done on male bank managers presented the same employee file/resume w/ all qualifications being the same as original but ½ were labeled male + ½ were labeled female.
* In EDA, the % of males promoted were 21/24 = 88% of males.
* However only 14/24 = 58% of females were promoted.
* We were clearly able to see a difference between % of males and females promoted.
* However, instead of jumping the gun, we said that there could be 2 competing claims that actually explain what's going on here:
* 1) Nothing 🡪 promotion + gender = independent 🡺 observed difference in proportions due to chance = **Null**
* 2) Something is going on 🡪 promotion + gender = dependent 🡺 observed difference in proportions due to gender discrimination (not due to chance) = **Alternative**
* Did a simulation-based inference under the assumption of the null being true/assumption of independence (leaving everything up to chance)
* For each simulation, we recorded this observed difference in promotion proportions:
* In our original observed data, this was 88 – 58 = 30%.
* We looked to see if 30% was an unusual outcome for this difference



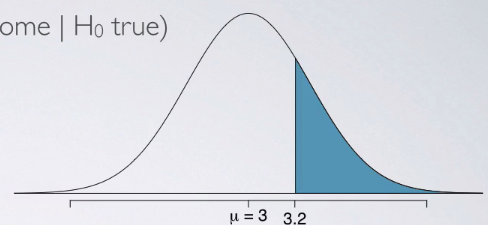
* The actual data value of 30% seems very unlikely based on the dot plot above 🡪 reject the null in factor of the alternative
* Framework 🡺 Start w/ Null H0 (status quo) 🡪 Create alternative H1 to represent research question (what we’re testing for) 🡪 conduct hypothesis test under assumption the null is true either via simulation or theoretical methods that rely on the CLT)
* If test results suggest the data do NOT provide convincing evidence for the alternative, stick w/ the null 🡺 If they do, reject the null in favor of the alternative.

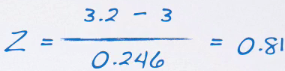
**Hypothesis Testing (for a Mean)**

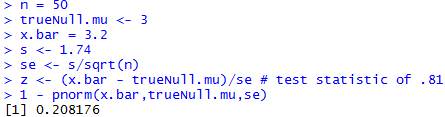
* **Null** = H0 = skeptical perspective or claim to be tested set to a parameter value
* **Alternative** = H1 = claim under consideration often represented by a range of possible parameter values (<, >, or != the null)
* Skeptics will not abandon H0 unless evidence in favor of H1 is *so strong* we can reject H0 in favor of H1
* 
* H0: μ = 3 (students have been in 3 exclusive relationships on average)
* H1: μ > 3 (students have been in > 3 exclusive relationships on average)
* Our null is included in our CI, so therefore μ *could* = 3
* CI says any value w/in it is the possible true population mean, so therefore 3 is a possible solution so we cannot reject H0 in favor of H1
* But this doesn’t tell us the p-value (likelihood of certain outcomes under the null), based on which we can make decision on the hypotheses
* \*\*Note: Hypotheses are always about **population parameters**
* Remember p = the probability of an observed (or more extreme) outcome value given H0 is true
* Here, this is x.bar > 3.2 | H0: μ = 3
* Since we assume H0 is true, we can use it to construct the sampling distribution based on the CLT

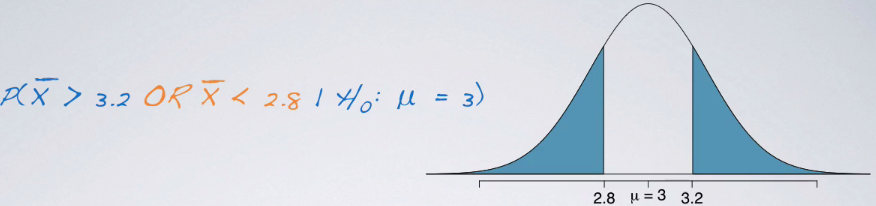




* X.bar is nearly normally distributed w/ mean = 3 + SE = .246
* Now draw the curve + shade area of interest corresponding to the p-value
* 





* So, **p = .209**
* We used the **test statistic** z to calculate **p-value** = *probability of observed data at least as favorable to the alternative as our current dataset, if the null were true*
* If p is low (< significance level **α** usually set at 5% or .05), we say it’s very unlikely to get the observed data if the null were true 🡪 reject it
* If p is high (> significance level **α** usually set at 5% or .05), we say it’s likely to get the observed data if the null were true 🡪 fail to reject it
* Our p-value is high (> .05) 🡺 even though we have a sample mean x.bar > 3, there is not enough evidence to reject the null that the population average # of exclusive relationships for college students = 3 🡪
* If in fact college students have been in = 3 exclusive relationships on average, the p-value tells us there is a 21% chance a random sample of 50 students would yield a sample mean >= 3.2
* Since this is a pretty high probability, we’d think a sample mean >= 3.2 is likely to happen simply by chance
* We fail to reject the null + these data do not provide convincing evidence college students have been in > 3 relationships on average + the difference between the null of 3 relationships + the observed sample mean of 3.2 relationships is simply due to chance or **sampling variability**.
* Often, instead of looking for a divergence from the null in a *specific direction* (< or >), we are interested in divergence in any direction 🡪 **two-sided/tailed hypothesis tests**,
* Definition of a p-value = the same regardless of doing a 1 or a 2-sided test.
* Calculation becomes slightly different + ever so slightly more complicated 🡪 need to consider at least as extreme as the observed outcome in BOTH directions away from the mean.
* Now set p-value now to be x.bar > 3.2 OR x.bar < 2.8, given the null is true = population mean = 3
* *Use 2.8 b/c it’s the same distance from 3 as 3.2*
* 
* Already know the upper tail = 0.209 + since this is a symmetric distribution, lower tail also = 0.209
* Therefore p-value = probability on upper tail + probability on the lower tail = ~41.8%.

