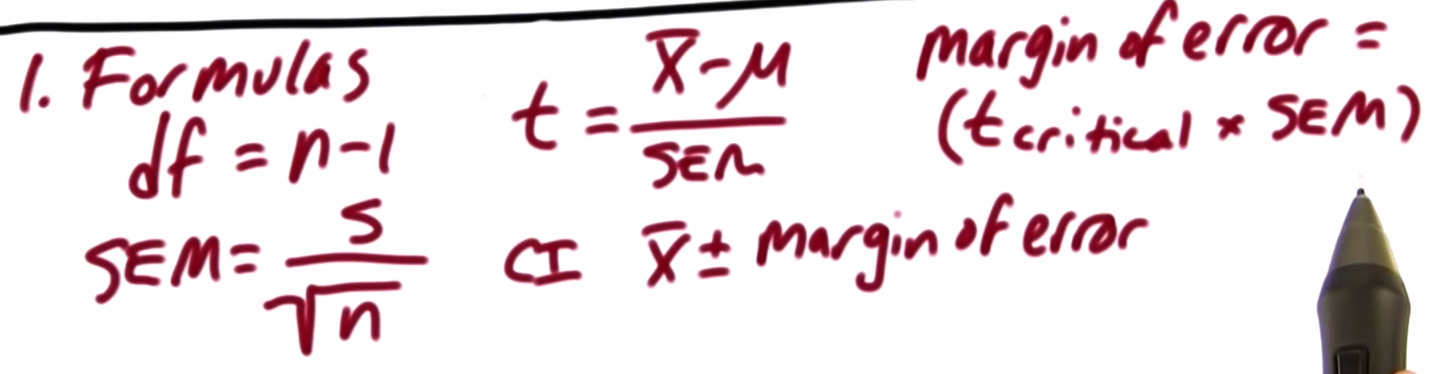
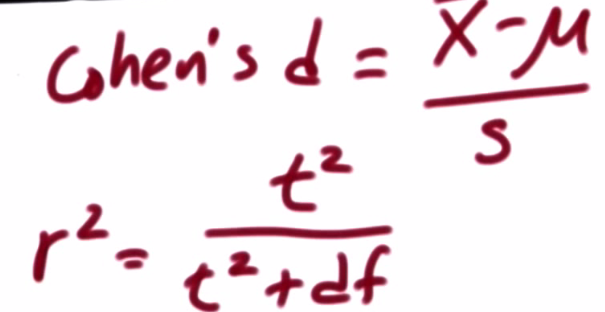
***Udacity Data Analyst Track***

**I. Into to Inferential Stats**

5. T-tests Part 2

* **Effect size** 🡪 size of the effect of our treatment
* In non-experimental studies, this may refer to the strength of the relationship between variables
* In z and t tests, simplest measure of effect size **= mean difference**
* Z or 1-tail t-test mean difference **= x – mu**
* Good for variables w/ easy-to-understand meanings (no specialized training)
* If we don’t know what a variable is telling us/it doesn’t have an everyday meaning, the mean difference may not be the best informative measure, so statisticians developed other effect size measures
* 2 main groups
* Difference measures
* **mean difference**
* **Standardized difference** 🡪 **Cohen’s D** (mean difference but in standardize units)
* Correlation measures
* **R2** 🡪 proportion/% of variation in a variable that is related to/explained by another variable
* Can explain a variation in 1 variable by knowing the value of the explanatory variable
* i.e. measures proportion of a difference in means that can be explained by the IV
* **Statistical Significance** 🡪 rejected the null + results not likely due to chance or sampling error
* *DOESN’T mean it’s important, large in size, or meaningful*
* How to tell if results were meaningful
* 1) What was measured?/What were the variables?
* 2) Did the variables have practical application, social issue, or theoretical (theory) importance?
* If not, the results are clearly not important, no matter what they are
* 3) Effect size = how large were the results?
* Small does not = unimportant and large does not = important
* Small effect sizes can be very important
* 3) Can we rule out random chance/sampling error as an explanation?
* Does not guarantee meaningful results, but goes a long way in helping us asses meaningfulness
* 4) Can we rule out alternative explanations (**lurking/confounding variables**) for our results?
* Jacob Cohen 🡪 **Cohen’s D** 🡪 difference measure giving the *standardize mean difference (x – mu)*
* When standardized, we take x – mu / sample SD to get the mean difference in SD units
* Tells us how far the means are in SD units
* **R2/Coefficient of Determination** 🡪 strength of the relationship between 2 variables
* Goes from 0 (not related at all) to 1 (perfectly related)
* R2 can be derived from t-test values 🡪 **t^2 / (t^2 + dF)**
* When computing statistical results from a research study, they are put into a **results section**
* 1) Descriptive Statistics (mean, SD, etc.) 🡪 the heart of the study, exactly what happened in it, should always be included + reported as text, graphs, or tables
* 2) Inferential Statistics 🡪 hypothesis test, CI, or both, telling reader what kind of test was conducted (1 or 2-sample t-test) + the t-statistic, dF’s, p-value, and direction of test (1 tail or 2)
* Reported in American Psychological Association (APA) format guide for research papers (i.e. no leading zeroes if the value cannot be greater than 1 like in p or r2)
* Ex: reporting t-test in APA 🡪 t(dF) = x.xx, p = x.xx, direction 🡪 t(24) = -2.50, p < .05/p = .01, one-tailed
* Ex: reporting CI’s in APA 🡪 CI is on what? (on a single mean or on the difference between 2 means); Confidence lvl = (lower limit, upper limit)
* Confidence Interval on the mean difference; 95% = (4,6­)
* 3) Effect size measures (d, r2, etc.) 🡪 d = x.xx, r2 - .xx
* FULL 1 SAMPLE T-TEST
*  
* Data – US families spend $151 on average per week on food in 2012 (Gallup survery) 🡪 mu
* Food co-op FoodNow! wants to reduce cost of food for members via cost-saving programs (buying from local growers, etc.)
* DV 🡪 $ spend per week on food, Treatment 🡪 cost-saving programs, h(0) = program did not change cost of food, h(a) = program significantly reduced cost of food
* H(0) 🡪 mu(p) = 151 H(a) 🡪 m(p) < 151
* Need 1 tailed t-test in negative direction
* N = 25, dF = 24
* t-critical for alpha = 0.05 🡪 -1.711
* s = 50, SEM = (50/sqrt(25)) = 10
* x = 126, mean difference = 126 – 151 = -25
* t = (-25)/10 = -2.5 🡪 t falls w/in critical region 🡪 p < .05 = statistically significant differences
* Whether or not our results are meaningful depends…. Is saving $25 a week trivial for you, or is it significant/very important (depends on income for members of the co-op)
* Cohen’s d = -25/50 = -1/2 🡪 these 2 means are ½ of a SD apart
* R2 = (-2.5)^2/((-2.5)^2 + 24) = 0.20661157024 = .21 = 21%
* 21% of the differences in food prices for those 25 co-op members are due to the treatment/cost-savings program
* Margin of error for a 95% CI (2-tailed, alpha = .05) 🡪 new t-critical = 2.064 🡪 2.064/\*10 = 20.64
* CI = (126-20.64, 126+20.64) = (105.36, 146.64) 🡪 If the co-op implemented this cost-savings program for all of its members, we are 95% sure they’d pay between $105.36 and $146.64/week
* Researchers are interested in effects of a law to penalize talking on a phone while driving, so they samples 10 counties in a state before + after passing the law + measures total accidence due to phone usage the year before + after the law + hypothesized the law would reduce the # of accidents
* Dependent-samples t-test + pre/post-test, and NOT longitudinal since not sampling *individuals*
* DV = # of accidents due to phone usage, treatment = law, 1-tailed t-test in negative direction
* H(0) = law will have no effect 🡪 mu(law) = mu
* mu(d) >= 0
* H(a) = law will reduce accidents due to phone 🡪 mu(law) < mu
* mu(d) < 0

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| FULL 2 SMPL T TEST |  |  |  |  |  |  |  |  |  |  |  |
| mu(d) | -3 |  |  |  |  |  |  |  |  |  |  |
| n | 10 |  |  |  |  |  |  |  |  |  |  |
| df | **9** |  |  |  |  |  |  |  |  |  |  |
| alpha | 0.05 |  |  |  |  |  |  |  |  |  |  |
| t-crit | **-1.833** |  |  |  |  |  |  |  |  |  |  |
| s(d) | 1.33 |  |  |  |  |  |  |  |  |  |  |
| SEM | **0.420582929** |  |  |  |  |  |  |  |  |  |  |
| x | 126 |  |  |  |  |  |  |  |  |  |  |
| t | **-7.13295712** | <-- t w/in critical region = statistically sig diff = reject null = p < .05 | | | | | | |  |  |  |
| cohen's d | **-2.25563909** | <-- the 2 means are 2.26 SD's apart | | | |  |  |  |  |  |  |
| r2 | **0.849697083** |  |  |  |  |  |  |  |  |  |  |
| t-crit for 95% CI | **2.262** |  |  |  |  |  |  |  |  |  |  |
| Margin of err for 95% CI | **0.95135858** |  |  |  |  |  |  |  |  |  |  |
| 95% CI (2 tail) | **-3.95135** | **-2.0486** | <-- 95% confidence that if we apply the law to all counties,  they'd reduce accidents by 2 to 4 | | | | | | | | |

* Ex: Pizza company A wants to know if they deliver faster than Company B

|  |  |
| --- | --- |
| A | B |
| 20.4 | 20.2 |
| 24.2 | 16.9 |
| 15.4 | 18.5 |
| 21.4 | 17.3 |
| 20.2 | 20.5 |
| 18.5 |  |
| 21.5 |  |

* Use Cohen’s d to measure the effect size between the two times.