GLM 1 Midterm Study Guide:

Lecture 1 – Introduction:

- Definitions:
  - **Descriptive Statistics:** is the summary of the information in a collection of data
  - **Inferential Statistics**: provides predictions about a population on the basis of a sample
  - **Population**: is the total set of units of interest.
  - **Sample**: is a subset of the population of interest.
  - **Parameter**: is a number that summarizes (tells us something about) a population.
  - **Statistic**: is a number that summarizes (tell us something about) a sample.
  - **Random Sampling**: is the idea that each member of a population has an equal chance of being selected to be part of a sample. This is a very crucial idea in statistics.
- Statistical models:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

Y: dependent (outcome) variable
$\hat{Y}$: predicted score of Y
Y- $\hat{Y}$: error

  - Measuring the error of the model:
    - The basic measure of a model's error is the **Sum of Squared Errors (SSE)** where you square each individual's error then add up all the squared errors
  - **Model comparison**: make incremental changes then see if the more complex model predicts Y sufficiently better, and if it does then we conclude that adding that increased complexity (adding the new predictor X) benefited the model
  - **General Linear modeling/general multivariate regression model**: a compact way of writing several multiple linear regression models in one linear model
    - *Used when the data is normally distributed*
    - Examples: t-test, regression, ANOVA

Data = Model + Error

$$Y = \beta_0 + \varepsilon$$

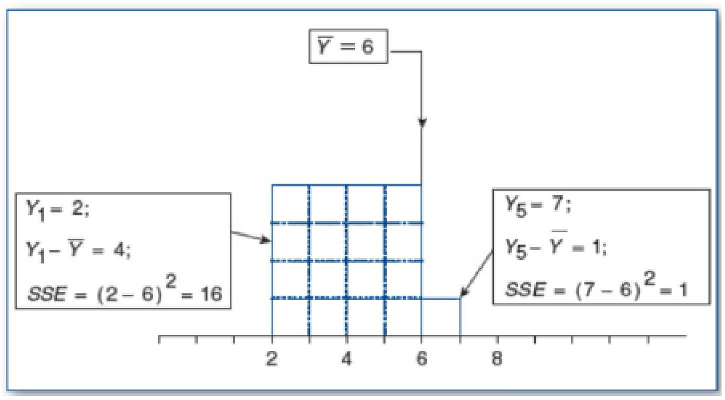
Lecture 2 – Predicting Scores:

- **Data**: (Y) the dependent variable or outcome variable. Specifically, values derived from scientific experiments
- **Model**: (Beta values) based on this model we can expect or anticipate the data. A model that predicts well implies we understand the phenomena. Beta sub 0 is the mean of Y aka it is the predicted value
  - **Mean only model:** The most basic model that predicts the mean for every subject
- **Error**: (epsilon) epsilon is the error that is left over when subtracting the mean predicted value (beta sub 0 aka Ŷ) from the actual observed value (Y)

- Steps to calculate error:
  - Version 1 (not preferred):
    - 1) Compute the mean (aka Ŷ, the observed value)
    - 2) Subtract Y – Ŷ to give you the deviations
      - HINT: all of these values combined should add up to 0
    - 3) Take the absolute value |Y – Ŷ|
    - 4) Σ |Y – Ŷ|
    - RESULT = error

OR

  - Version 2 (preferred) – calculating the **sum of squared errors (SSE)**:
    - 1) Compute the mean (aka Ŷ, the observed value)
    - 2) Subtract Y – Ŷ to give you the deviations
      - HINT: all of these values combined should add up to 0
    - 3) square the deviations $(Y - \hat{Y})^2$
    - 4) take sum of squared deviations aka sum of squared errors $\Sigma(Y - \hat{Y})^2$
    - RESULT = SSE
      - SSE represented visually:



- Conceptual vs Computational formulas to calculate SSE (error):
  - **Conceptual formula**:

$$\boxed{\text{Conceptual formula: } SSE = \sum(Y_i - \widehat{Y}_i)^2}$$

  - **Computational formula** (preferred):

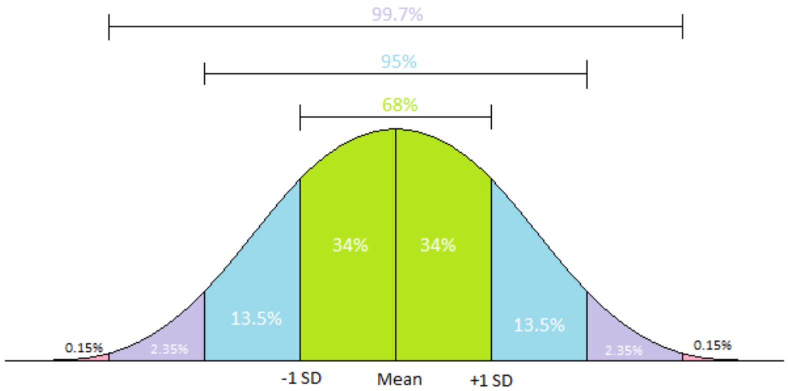
$$\boxed{\text{Computational formula: } SSE = \sum Y^2 - \frac{(\sum Y)^2}{N}}$$

- Descriptive statistics:
  - **Variance**: (S^2) a measurement of how far each number in a dataset is from the mean

$$\boxed{\text{Variance} = S_y^2 = \frac{\sum Y^2 - \frac{(\sum Y)^2}{N}}{N-1}} \quad \text{aka} \quad \boxed{\frac{SSE}{df}}$$

- **Standard deviation**: (S) a measure of how dispersed the data is in relation to the mean

$$\boxed{S = \sqrt{Variance}}$$



- **Standard error**: (SE) is the measure of statistical accuracy of an estimate

$$\boxed{SE = \frac{Standard\ dev.}{\sqrt{sample\ size}}} \quad \text{aka} \quad \boxed{SE = \frac{S}{\sqrt{n}}}$$

- Steps to compute descriptive statistics in SPSS:
  - Option 1: Analyze > descriptive statistics > descriptives > in new window click variable you want descriptives on and move that to the right > options > check descriptives of interest > click ok
    - Output looks like this:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
CEat	12	1.00	5.00	2.8333	1.26730
Valid N (listwise)	12				

Lab 2 – Descriptive Statistics:

  - Option 2: Analyze > descriptive statistics > frequencies > in new window click variable you want descriptives on and move that to the right > statistics > check descriptives of interest > click ok
  - Option 3: Analyze > descriptive statistics > explore > in new window click variable you want descriptives on and move that to the right > plots > check descriptives/plots of interest > continue > make sure “both” is checked under display > click ok

- <u>HINT</u>: checking "histogram" will also give you a <u>**normality plot aka a Q-Q plot**</u>

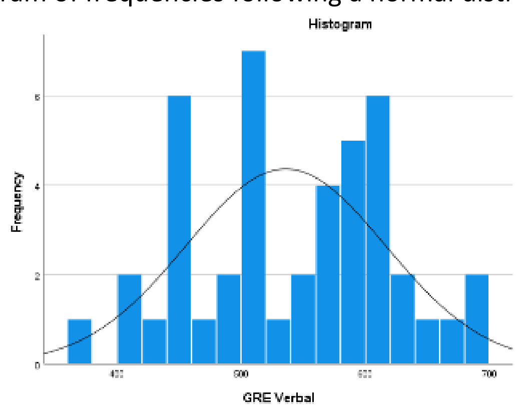
Descriptives	Frequency	Descriptives	Explore
Mean	Must select	Default	Default
Standard Deviation	Must select	Default	Default
Range	Must select	Must select	Default
Minimum	Must select	Default	Default
Maximum	Must select	Default	Default
Skewness	Must select	Must select	Default
Kurtosis	Must select	Must select	Default
Number of Missing	Default	Not given	Default
Frequency distribution	Selectable	Not available	Breakdown by factor
Charts/ Plots	Multiple available	Not available	Normality plots available

- ASSUMPTIONS: Checking normality:
  - Steps to test for normality in SPSS: Analyze > descriptive statistics > explore > move variables of interest over to dependent list > plots > check normality plots with tests > check histogram > continue > ok
  - Methods to check for normality:
    - Skewness value between -1 and 1
    - Kurtosis value between -2 and 2

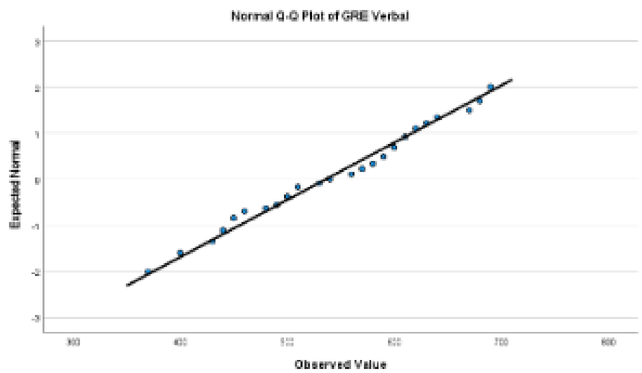
Descriptives

			Statistic	Std. Error
GRE Verbal	Mean		535.91	12.132
	95% Confidence Interval for Mean	Lower Bound	511.44	
		Upper Bound	560.38	
	5% Trimmed Mean		536.01	
	Median		540.00	
	Variance		6475.899	
	Std. Deviation		80.473	
	Minimum		370	
	Maximum		690	
	Range		320	
	Interquartile Range		135	
	Skewness		-.090	.357
	Kurtosis		-.810	.702

    - Histogram of frequencies following a normal distribution



- Q-Q plots with most points following along the diagonal (also normal p-p plots)



- Non-significant Shapiro-Wilk's test and Kolmogorov-Smirnov test (>0.05 means normally distributed)

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
GRE Verbal	.096	44	.200*	.973	44	.384
GRE Quant	.158	44	.007	.901	44	.001

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

- SAMPLE LANGUAGE: "The data is normally distributed according to the skewness values of .85 and .60 (within the normal range of -1 to 1), and the kurtosis values of -.31 and -.65 (within the normal range of -2 to 2). The results of the Kolmogorov-Smirnov test (significance values of 0.01 and 0.04 before and after intervention respectively, falling within the 0.05 range for significance) and the Shapiro-Wilk test (with values of .00 and 0.01 before and after the intervention respectively, falling within the 0.05 threshold) do not support these findings, but given that the skewness and kurtosis values are still within range, we can say this data meets the assumption of normality."
- Shapiro-Wilks: W(120) = .995, p = .950

Lecture 2 – Hypothesis Testing:

- **Hypothesis testing**: a decision-making process where two possible decisions are considered (null hypothesis – H0 and alternative/research hypothesis – H1) to describe a parameter or population
  - Steps of hypothesis testing:
    - 1) State the null and alternative hypotheses

    - For example:
        - Null: (H0) Children in the U.S. watch an average of 3 hours of TV per week (H0: m = 3) [usually written as H0: µ = 0]
        - Alternative: (H1) Children in the U.S. watch more or less than 3 hours of TV per week (H1: m ≠ 3) {this is a two tailed prediction, could also be one-tailed} [usually written as H1: µ =/= 0, H1: µ > 0, or H1: µ < 0]
- 2) Set the criteria for a decision/the level of significance
    - Usually, p < or = .05
        - If your p < or = .05 then you reject the null hypothesis and say that there is actually a difference/your alternative hypothesis is supported
- 3) Compute the test statistic
    - **Test statistic**: (z-score) how far a sample outcome is from the value stated in a null hypothesis. The larger the value, the further a sample mean deviates from the population mean
        - Type 1: z-score

$$z = \frac{x - M}{SD} \Rightarrow Z_{obt} = \frac{M - \mu}{\sigma_M}$$

            - *z:* z-score
            - *x*: sample mean
            - M: population mean
            - SD: standard deviation
            - µ: mean of means
            - σM: standard error (aka standard deviation of the mean)
        - Type 2: t-score (preferred)

$$t_{obt} = \frac{M - \mu}{s_M}$$

            - M: population mean
            - µ: mean of means (the number you are suspecting that your mean is different from in your hypotheses)
            - S sub M: standard error (aka standard deviation of the mean)
    - SPSS output of a one sample (nondirectional) t-test:

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
CEat	7.745	11	.000	2.83333	2.0281	3.6385

        - t: t value you compare to the t crit
        - df: N – 1
        - sig (2 tailed): tells you it's a 2 tailed test. This tells you whether to reject the null or not
        - mean difference: the absolute difference between the mean value in two different groups
        - 95% confidence interval: gives us the middle range that contains 95% of the data. If our t value is within this range then we retain the null, if it is outside of this range then we reject the null
      - Steps to compute in excel:
        - 1) compute the deviations between each score and the mean. So, score - mean
        - 2) square each individual deviation
        - 3) add up the sum of the squared deviations to get the sum of squares
        - 4) get the standard deviation by taking the square root of the sum of squares divided by n - 1 (aka the degrees of freedom).
        - 5) calculate the standard error by dividing the mean over the square root of the sample size
    - 4) Make a decision
      - Reject H0: if t value > tcrit
        - the sample mean is associated with a low probability of occurrence. (there is a less than 5% chance that this result was by chance, the results fell outside of the 2 standard deviations of the mean mark, so these results are significant)
      - Retain (fail to reject) H0: if t value < tcrit
        - the sample mean is associated with a high probability of occurrence. (your score fell within the 95% confidence interval/within 2 standard deviations of the mean, so it was not significantly different than the mean, so your results were not significant)
      - Use a t table for this. The df is N-1
- **Central Limit Theorem**: Assumption that with a large enough sample size, the scores form a normal distribution
- Types of error & power:
  - **Type I error**: when we say there is a difference and we reject the null, but there is truly no difference, and we should have retained the null (false positive – telling a man he is pregnant))
  - **Type II error**: we say there is no difference (aka we retain the null), but there is actually a difference, and we didn't detect it (false negative – telling a pregnant woman that she isn't pregnant)

- **Power**: when you say there is a difference (aka that you should reject the null) and there actually is a difference. The goal of statistical research is to increase this power to detect the difference while also reducing the type II error rate

Lecture 3 – Bivariate Regression:

- The simple mean-only model is what we have computed so far, but it isn't the best because it has a lot of error (SSE). Adding in a predictor variable (an X) to the model helps reduce error. This is now called a **Bivariate Regression** or **Simple Linear Regression**
- Contains a predictor variable/known variable/independent variable (X), and a criterion variable/to-be-predicted variable/dependent variable (Y)
- This describes a correlational relationship
- **Regression analysis** provides an equation that corresponds to a line that best predicts Y scores from X scores
- A population parameter (β), estimate of population parameter using sample data (*b*)
- Comparing the two models:

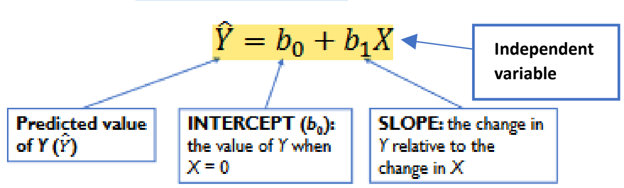
Simple (Mean-only) model	*vs.*	*Enhanced (bivariate regression) model*
Compact model (Model C)	*vs.*	*Augmented model (Model A)*
$\boxed{Y_i = \beta_0 + \varepsilon_i}$		$\boxed{Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i}$

  - The addition of the β1Xi variable in the bivariate regression model/Model A will help explain some of the unexplained variance from the mean only mode/Model C. Model A also has a lower amount of error than Model C (has the most amount of error)
  - β are the **regression coefficients**
- **Effect Size**: (r^2) tells us the amount/proportion of explained variance of one factor (Y) that can be explained by another factor (X) (mathematically equivalent to eta-squared)
- **Sum of Squared Error (SSE)**: the unexplained variance

Model	Source	Label	Error
$Y_i = \beta_0 + \varepsilon_i$	Model C	Total error	$SST\ (SSE_C)$
$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$	Model A	Residual error	SSE_A
	Model C – Model A	Explained error	SSR

*SST is almost always greater than SSE_A.

$SST = SSE_A + SSR$



  - **Explained error**: the amount of explained variance (error) by including the variable X
  - How to calculate the regression equation and SSEa:
    - 1) Calculate b_1:

Conceptual formula for computing b_1:	$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$
Computation formula for computing b_1:	$b_1 = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}}$

- 2) Calculate b_0:

$$b_0 = \bar{Y} - b_1\bar{X}$$

  - Variables with the line over it refer to the mean of that variable. You can use the means because the regression line must pass through the means (so without a predictor variable, b_0 is the mean of Y)
- 3) Put b_0 and b_1 into the equation format for Model A (above)
- 4) Plug the X values into the regression equation to give you the Ŷ values
- 5) Solve for SSEa using the following equation:

$$SSE = \sum(Y_i - \hat{Y}_i)^2$$

- RESULT: SSEa

- The explained error of the model is what you get when you subtract the SST aka SSEc and the SSEa.

$$SST = SSE_A + SSR$$

- To find the percentage of how much the regression model explained of the original SST in variable Y (R^2, aka how much of the error is reduced by adding in one X value – this is often used as an effect size), you follow this equation:

$$R^2 = \frac{SST - SSE_A}{SST} \quad \text{aka} \quad R^2 = \frac{SSR}{SST}$$

- Steps to running a bivariate regression/simple linear regression in SPSS: Analyze > Regression > Linear > move DV (Y) into dependent area, move IV (X) into block 1 of 1 > Statistics > check Durbin Watson, estimates, confidence intervals, model fit, descriptives, and collinearity diagnostics (not sure if all are needed here) > continue > plots > ZPRED in Y and ZRESID in X (this is to get the plot of errors, you are looking at the relationship between the predicted values and the residual error values) > check both histogram and normal probability plot (you only need one, but having both is helpful) > continue > save > check Cook's and unstandardized residuals
  - SPSS output for Bivariate Regression:

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	X[b]	.	Enter

a. Dependent Variable: Y

b. All requested variables entered.

- Table tells you what variables are used in summary

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.897[a]	.805	.785	1.24967

a. Predictors: (Constant), X

- R: correlation between X and Y (tells you strength and direction)
- R^2: indicates how much of the total variation in the dependent variable can be explained by the independent variable. This is a percentage (effect size)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	64.383	1	64.383	41.227	<.001[b]
	Residual	15.617	10	1.562		
	Total	80.000	11			

a. Dependent Variable: VAR1

b. Predictors: (Constant), VAR2

$$F(1,10) = 41.227, p < .001, R^2 = .805$$

- This table reports how well the regression equation fits the data (p <.05 indicates that, overall, the regression model statistically significantly predicts outcome the outcome variable)
- Regression and ANOVA are basically the same thing
- Sum of square total: SST value
- Sum of squares residual: SSEa value
- Sum of squares regression: SSR value
- HINT: When reporting the F statistics, be sure the italicize all of the variables!
- SAMPLE LANGUAGE FOR REPORTING F STATISTICS: A simple linear regression revealed a statistically significant relationship between x and y, F(1,10) = 41.227, p < .001, R2 = .805.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	10.266	.756		13.579	<.001
	X	-.776	.121	-.897	-6.421	<.001

a. Dependent Variable: Y

- Unstandardized b constant: b0 intercept value (indicates that if a x is 0, we would predict that y value will be 10.266)
- Unstandardized b X value: b1 value in the regression equation (indicates that if a x is 0, our best prediction would be that y score will be .776 points lower)
- Put together with these values, the equation would be Y = 10.266-.776X

Correlations

		VAR1	VAR2
Pearson Correlation	VAR1	1.000	-.897
	VAR2	-.897	1.000
Sig. (1-tailed)	VAR1	.	<.001
	VAR2	.000	.
N	VAR1	12	12
	VAR2	12	12

  - Will also get this table with the correlations, not necessary here

- EXAMPLE LANGUAGE – full APA: "A simple linear regression test was run to see to determine if there is a statistically significant relationship between mother's IQ and kid's cognitive score. It was hypothesized that higher mother's IQ leads to higher kid's cognitive scores. [assumption section was skipped] A simple linear regression revealed a statistically significant relationship between mother's IQ and kid's cognitive score, F(1,432) = 108.64, p < .001, R2 = .20. The intercept for our model, B = 25.8, indicates that we would predict a kid's score on our cognitive test to be 25.8 if their mother had an IQ score of 0. As this represents an unrealistic IQ score, this intercept does not make practical sense, but provides a general idea of how our values are anchored (e.g., where our slope could start) within the regression model. Our slope, B = .61, indicates that for every 1-point increase in mother's IQ, we would predict a .61 increase in the kid's cognitive score. According to our R2 value, R2 = .20, mother's IQ explains about 20% of our observed variance in kid's cognitive test score."
- Centering the predictor variable to its mean to make calculations easier:
  - Centering the predictor of X to its mean can generate a better estimate because then the estimated β0 is now the predicted value of Y when the predictor X is its mean
  - Subtract X – Xmean to get the centered X values (X'), then use this variable as your X in all future calculations
  - When using centering, the regression equation shows the same relationship between X and Y and the SSEa is the same

Lab 3 – Bivariate Regression/Simple Linear Regression:

- Symbol reminder:

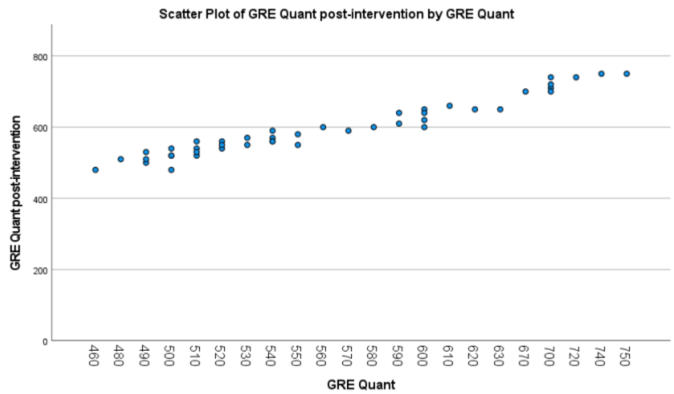
  $\bar{V}$ (Variable bar) = mean of the variable

  V_i = individual value of the variable

  $\hat{V}$ (variable hat) = predicted value of variable

- Bivariate Regression/Simple Linear Regression Assumptions (5 out of 7):
  - **Linearity**:
    - A linear relationship is needed between the 2 variables

- HOW: create a scatter plot then visually inspect for linearity



- LANGUAGE TO REPORT: “In terms of testing the assumption of linearity, a scatter plot was created to compare GRE quant scores both before and after the intervention and it shows a linear relationship between the two variables, therefore allowing us to say that the assumption of linearity is met (figure 3)”
- Steps to make scatter plot in SPSS: graphs > chart builder > click ok to pop up > in bottom left click "scatter/dot" > drag the first scatter plot picture to the window above > drag the y var to the y axis and the x var to the x axis > on far right check box for total under linear fit lines > click okay

○ **No significant outliers**:

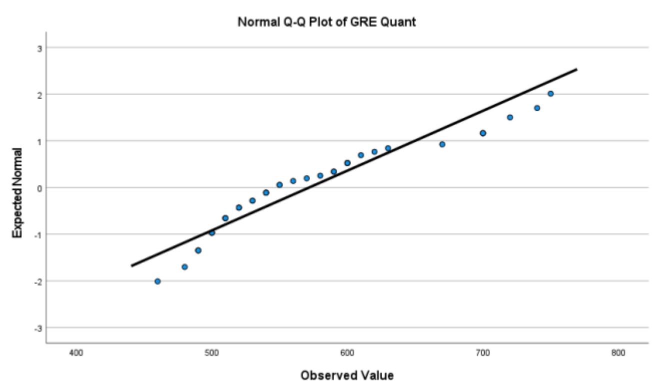
- HOW: Visual inspection of a scatter plot or normal P-P plot or normal Q-Q plot or box and whisker plot, or looking at Cook’s distance value (maximum should be under 1 for there to be no outliers present), or looking at a frequency histogram

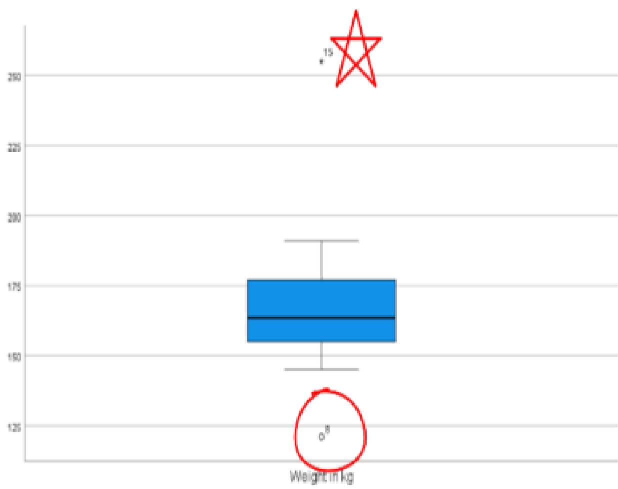
Residuals Statistics[a]

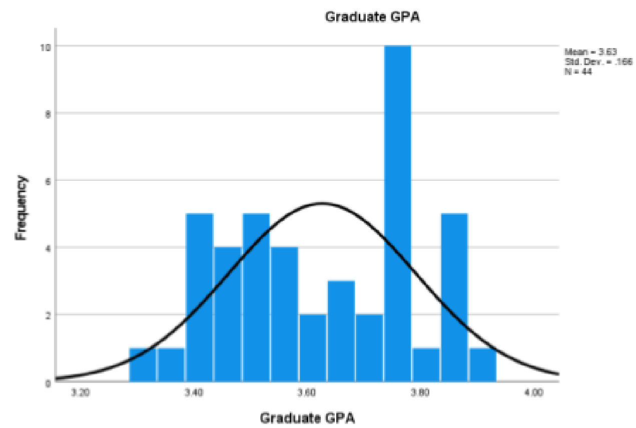
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2.51	9.49	6.00	2.419	12
Std. Predicted Value	-1.443	1.443	.000	1.000	12
Standard Error of Predicted Value	.366	.652	.501	.102	12
Adjusted Predicted Value	1.95	9.67	5.98	2.462	12
Residual	-2.164	1.491	.000	1.192	12
Std. Residual	-1.731	1.193	.000	[illegible]	[illegible]
Stud. Residual	-1.829	1.399	.007	[illegible]	[illegible]
Deleted Residual	-2.416	2.049	[illegible]	[illegible]	[illegible]
Stud. Deleted Residual	-2.128	1.479	[illegible]	[illegible]	[illegible]
Mahal. Distance	.026	2.082	[illegible]	[illegible]	[illegible]
Cook's Distance	.000	.366	[illegible]	[illegible]	[illegible]
Centered Leverage Value	.002	.189	.083	[illegible]	[illegible]

a. Dependent Variable: VAR1

Value greater than 1 is problematic

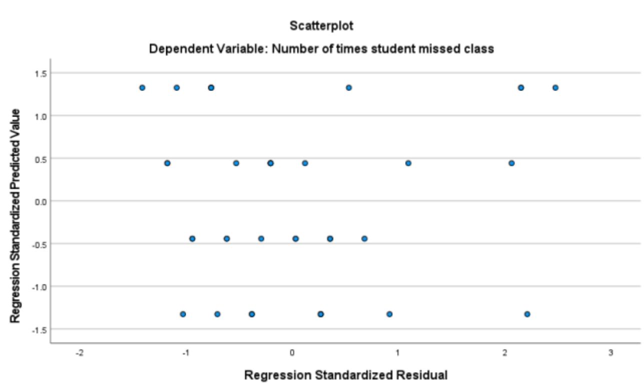


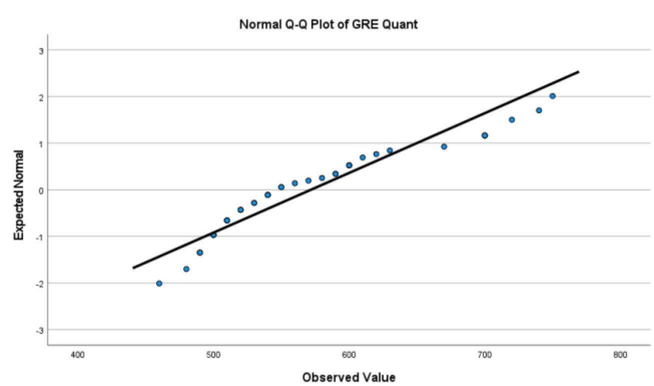




- LANGUAGE TO REPORT: "In terms of extreme outliers, neither Normal Q-Q plots of GRE Quant scores before and after the intervention show outliers upon visual review (see figures 1 and 2), so we can say the assumption of no extreme outliers is met."

- **Homoscedasticity**:
  - **<u>Homoscedasticity (aka homogeneity of variance)</u>**: A condition in which the variance of the residual, or error term, in a regression model is constant. Aka there is no pattern in the scatter plot, it is just data in a straight line, no curves
  - HOW: Visual inspection of a normal Q-Q plot/scatterplot of errors making sure there is no pattern in the scatterplot or curves in the normal Q-Q plot, also a nonsignificant Levene's test ($p > 0.05$ means assumption is met and variances are equal)





Tests of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Injury Severity (0-100)	Based on Mean	.891	3	26	.459
	Based on Median	.827	3	26	.491
	Based on Median and with adjusted df	.827	3	19.578	.495
	Based on trimmed mean	.889	3	26	.460

The result of **Levene's test is not significant, $F(3,26) = .891$, $p = .459$, at the $\alpha = .05$ significance level.** Therefore, we fail to reject H_0 and conclude **the variances are equal** between the groups.

- LANGUAGE TO REPORT: "homoscedasticity is present (as seen by the consistency in the variance of errors in the scatterplot; Figure 2)"
- LANGUAGE TO REPORT: "As Levene's test was not statistically significant (F = .001, p = .367), we retain the null hypothesis and assume the variances are equal. Thus, we will use the top row (Equal variances assumed) to report and interpret our results."

- **Errors are independent for all observations**
  - HOW: Examine values on the **Durbin-Watson test** which is used to detect the presence of autocorrelation in the residuals from a statistical model or regression analysis. Values should fall within the normal **range of 1.5 to 2.4**, values outside of this range may be a cause for concern. If the number is close to that range you can say that the assumption is assumed to be met because the number is close to the normal range. Record if any assumptions are not met

Model Summary[b]

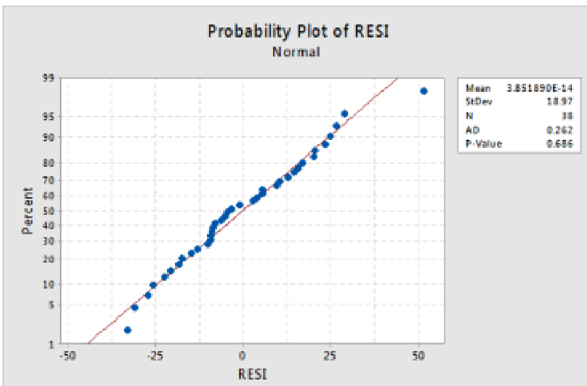
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.897[a]	.805	.785	1.250	1.482

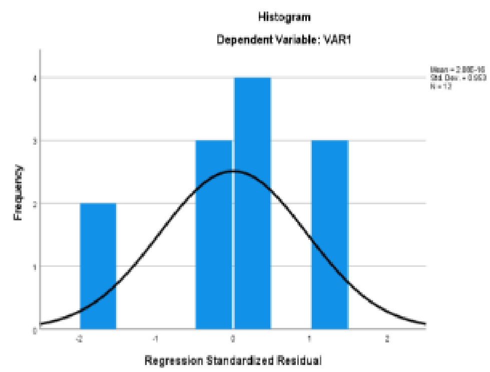
a. Predictors: (Constant), VAR2

b. Dependent Variable: VAR1

- LANGUAGE TO REPORT: "The assumption for the independence of errors for all observations is met as seen from the value from the Durbin-Watson test (1.996; within the normal range of 1.5-2.4)"

- **Errors are normally distributed**
  - HOW: Check either a normal P-P plot or a histogram or scatterplot of errors. If the random errors are normally distributed, the plotted points will lie close to straight line. This is similar to linearity, but we are checking it by looking at the errors (see more earlier in study guide)





  - LANGUAGE TO REPORT: "The errors are normally distributed as seen on the P-P plot/histogram, so we can say this data meets the assumption that errors are normally distributed."

Lecture 4 – Model Comparison:

- Model comparison requires that we have two models that predict the same data and that they are nested, meaning that each element of the smaller (compact) model must be present in the

$$Neuropsych = \beta_0 + \beta_1 Meth + \varepsilon$$

$$Neuropsych = \beta_0 + \varepsilon$$

Nested

larger (augmented) model. So, if any variable in the larger model is 0, then you are left with the smaller model. You can't directly compare two models if they aren't nested
- Steps to Compare Models:
  - 1) State the Model C and an augmented Model A.
    - Here Model C doesn't have to be mean only, it just has to be smaller than Model A. Terms with a β and a variable ($\beta_1 X_1$) are called regression coefficients/regression weights/slope coefficients

Model A:	$Y = \beta_0 + \beta_1 X_1 + \varepsilon$
Model C:	$Y = \beta_0 + \varepsilon$

  - 2) Identify the null hypothesis (H_0).

- Here that would be that there is no difference between Model C and Model A. The typical H_1 would say the variables are related. Basically, this means that if the b1 term in model A is 0 then model A becomes the same thing as model C. So, if the b1 term shows no significant value then that basically means the two models are equivalent and you can accept the null hypothesis

H_0:	$\beta_1 = 0$

- 3) Count the number of parameters estimated by each model.
  - Parameter: any coefficient in a model that was calculated using data. Parameters are indicated with the Greek letter, beta (β)
  - In an augmented model, you have at least 2 parameters (b_1 and b_0)

- 4) Calculate the regression equation

Conceptual formula for computing b_1: $b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$

Computation formula for computing b_1: $b_1 = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}}$

$$b_0 = \bar{Y} - b_1\bar{X}$$

- 5) compute the total sum of squares (SSEc aka SST). Use the Ymean for the Y predicted here

$$SSE = \sum(Y_i - \widehat{Y}_i)^2$$

- 6) Compute the sum of squares for Model A (SSEa aka SSB)

$$SSE = \sum(Y_i - \widehat{Y}_i)^2$$

- 7) Compute sum of squares reduced (SSR aka SSW)
  - This tells us how much error was reduced by adding X to augment the model, aka it tells us about the amount of explained variance between Y with the impact of X vs without based on our model. A large SSR means our model predicts the data well, so we want a high SSR (more explained variance)

$$SST = SSE_A + SSR$$

- 8) Compute the **proportional reduction in error** (PRE aka R^2).
  - This is a percentage of "variance explained," aka this is the "effect size" (how much overlap there is between the two variables

$$R^2 = \frac{SSR}{SST}$$

- 9) Compute the summary table (an ANOVA table):

Source	SS	df	MS	F	R^2
Model comparison	SSR	PA – PC	MS_{model} = SSR ÷ (PA – PC)	$MS_{model} \div MS_{residual}$	SSR ÷ SST
Residual	SSE_A	N – PA	$MS_{residual}$ = SSA ÷ (N – PA)		
Total	SST	N – PC			

  - MS column: gives us the ratio between the mean of squares
  - PA – PC: # of parameters of Model A – # of parameters for Model C

- 10) Decide about H_0
  - If you are rejecting the null then you would say β_1 =/= 0 saying that there is a significant difference between the two variables
  - The closer F is to 1, the less chance you have of rejecting the null. You want a large F
  - The F ratio generated in the table above is the ratio between the residual and explained variances, So the closer the value is to 1, then the less difference there is between the models with and without the predictor variable and therefore the less chance you have of getting significant findings and rejecting the null. If the F ratio is much larger than 1, then you are much more likely to get a significant relationship between X and Y

- Types of Correlation:
  - **Pearson's R Correlation** (most common):
    - **Correlation**: statistical procedure used to describe the strength and direction of the linear relationship between 2 variables
      - r = .2 to .4 -> weak correlation
      - r = .4 to .6 -> medium/moderate correlation
      - r = .6 to .8 -> strong correlation
    - Values range from -1, to 1, with both of those values insinuating a perfect correlation (rare). Values of r = 0 doesn't mean the variables aren't related, it just means they aren't LINEARLY related.
    - Correlations are normally represented on a scatter plot
    - If you square root the proportional reduction in error (R^2), then you get Pearson's R value (r). The full formula is as follows:

$$\boxed{r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}}$$

    - You can calculate Pearson's r in SPSS following these steps: Analyze > correlate > bivariate > move over the variables of interest > check Pearson's and two-tailed test > options > can check means and standard deviations if you want > continue > click Pearson's R > ok

- Output is as follows:

Correlations

		Pretest achievement scores	Posttest achievement scores
Pretest achievement scores	Pearson Correlation	1	.654**
	Sig. (2-tailed)		<.001
	N	40	40
Posttest achievement scores	Pearson Correlation	.654**	1
	Sig. (2-tailed)	<.001	
	N	40	40

**. Correlation is significant at the 0.01 level (2-tailed).

▪ Assumptions for using Pearson's r (all must be met):
- How to assess the assumptions of the correlation in SPSS: Analyze > Descriptive Statistics > explore > move over dependent variables > statistics > check descriptives > continue > plots > check histogram and normality plots with tests > continue > display for both > ok
  - 1) Association must be linear
    - How to check in SPSS: graphs > chart builder > click ok > pull the first scatter plot in the bottom left into the big window at the top > drag the DV and IV from the top left onto the dotted areas of the graph > click ok > add a line
      - When you look at the scatter plot, if the points look linearly related, then you can say the assumption of linearity is met
        - **<u>OR</u>**
    - How to check in SPSS: analyze > compare means > means > move the variables of interest to dependent list and independent list > check test for linearity > continue
      - If the value under deviation from linearity and sig is > .05, then the variables ARE linearly related. If < .05 then they are NOT linearly related

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Pretest achievement scores * Posttest achievement scores	Between Groups	(Combined)	7475.817	28	266.993	2.665	.045
		Linearity	3672.763	1	3672.763	36.658	<.001
		Deviation from Linearity	3803.054	27	140.854	1.406	.282
	Within Groups		1102.083	11	100.189		
	Total		8577.900	39			

  - 2) No extreme outliers

        - look at the normal Q-Q plot to see if there are extreme outliers that can wrongfully inflate the correlation coefficient
        - see if the data point in question is beyond 3 standard deviations away from the mean
    - 3) Independent observations for X and Y, no overlap (but there should be pairs of datapoints for each ID, an X and a Y for each person)
        - Can visually examine data and look at N values of SPSS output

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Pretest achievement scores	40	100.0%	0	0.0%	40	100.0%
Posttest achievement scores	40	100.0%	0	0.0%	40	100.0%

        - I.e., Ron's education score should be separate from Sally's. However, if Sally and Ron are both working on something together and there is something impacting both of them that makes them different from the rest of the class, then the assumption of independent observations is NOT met and you can't use a pearson's R
    - 4) For now, only continuous variables
        - If you look at the "values" section under the "variable view" of a dataset and it says "none" then you can assume that the variable is continuous. If the variables are continuous then you can run Pearson's R. If the variable is dichotomous, aka there are 2 options shown under the "values" section then you can't use Pearson's R. However, if you have a dichotomous variable in SPSS and you click Pearson's R, then it still works because it automatically switches it to a biserial test for you
    - 5) Variables are normally distributed
        - Under the section tests of normality, if the significance value under the Kolmogorov-Smirnov test is <0.05 then you assume that the data is NOT normally distributed. You can report the Shapiro-Wilk test statistic as well that also has the range of < 0.05.
    - 6) Sample must be similar to population of interest
    - 7) Homoscedasticity: the variance of the errors in Y is constant for all values of X

        - If levene's test of equality of variance is NOT significant, then you can follow the assumption that equal variances are assumed

- **Point Biserial Correlation**:
    - Assumptions:
        - Used when one variable is dichotomous
        - No outliers
        - Normally distributed
        - Equal variances (levene's test)
    - If you have a dichotomous variable in SPSS and you click Pearson's R, then it still works because it automatically switches it to a biserial test for you. Output will still say Pearson correlation though. The table output setup is the same as for a pearson's.
    - SAMPLE LANGUAGE: "A point-Biserial correlation was run to determine the relationship between engagement in an internet advertisement and gender. There was a negative correlation between engagement and gender, which was statistically significant (rpb= -.358, n = 40, p = .02)."
- **Spearman's p (Rho):**
    - Assumptions:
        - Used when at least one measure is ordinal (in a series/in order)
        - Variables are ordinal, interval, or ratio scale
        - Paired observations
        - Linear relationship
    - Same steps to conduct a pearson's correlation, just check spearman's instead
    - SAMPLE LANGUAGE: "A Spearman's rank-order correlation was run to determine the relationship between 10 students' English and maths exam marks. There was a strong, positive correlation between English and maths marks, which was statistically significant (rs(8) = .669, p = .035)"
- **Kendall's tau-b**:
    - Assumptions:
        - Used when normality assumption is NOT met
        - Continuous or ordinal variables
        - Linear relationship
    - Same steps to conduct a pearson's correlation, just check spearman's instead
    - SAMPLE LANGUAGE: "A Kendall's tau-b correlation was run to determine the relationship between income level and views towards income taxes amongst 24 participants. There was a strong, positive correlation between income level and the view that taxes were too high, which was statistically significant (τb = .535, p = .003)."
        -
- **Phi Coefficient**:
    - Not very common
    - Used when both variables are dichotomous
- **Cramer's V:**

- Not very common
- Used when both variables are nominal (named), but one has more than 2 levels

Variable Y\X	Quantitiative X	Ordinal X	Nominal X
Quantitative Y	Pearson r	Biserial r_b	Point Biserial r_{pb}
Ordinal Y	Biserial r_b	Spearman rho/Tetrachoric r_{tet}	Rank Biserial r_{rb}
Nominal Y	Point Biserial r_{pb}	Rank Bisereal r_{rb}	Phi, L, C, Lambda

Lecture 5 – Correlation: Traditional and Regression Approaches:

- Steps for hypothesis testing in correlation:
  - 1) State the hypotheses:
    - For two tailed test:
      - H0: r = 0 [Y is not related with X.]
      - H1: r ≠ 0 [Y is related with X.]
    - For one tailed test:
      - H0: r = 0 [Y is not related with X.]
      - H1: r > 0 [Y is positively correlated with X]
      - H1: r < 0 [Y is negatively correlated with X]
  - 2) Set the criteria for a decision:
    - "We will compute a two-tailed test at a .05 level of significance"
    - Find the critical value in a Pearson Correlation r table
    - df for a correlation are n – 2 (it is 2 here because there are 2 sets of information for each variables' mean)
  - 3) Compute the test statistic:

$$\boxed{\text{Pearson } r = \frac{SS_{XY}}{\sqrt{SS_X SS_Y}} = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2 \times \Sigma(Y-\bar{Y})^2}} = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\left[\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right]\left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right]}}} = \frac{\textit{Sum of cross products}}{\textit{Combined variance}}$$

    - SS: Sum of squares, aka sum of squares of error (SSE)
    - Numerator: the extent to which values on the x-axis (X) and y-axis (Y) vary together. This common/shared error variance is called the **covariance** of X and Y
    - Denominator: extent to which values of X and Y vary independently, or separately
  - 4) Make a decision:
    - If the computed value is GREATER than the r crit value, then there is a significant correlation. If it is less than the crit value, then you can say the relationship is not significantly different than 0
    - SAMPLE LANGUAGE: "Using the Pearson correlation coefficient, there was found to be a strong, negative, statistically significant relationship between X and Y, (r = −.897, p<.05)."

- 5) Calculate effect size:
  - Coefficient of determination (r^2 or R^2) aka the effect size –mathematically equivalent to eta squared, and is used to measure the proportion of variance of one factor (Y) that can be explained by known values of a second factor (X) aka this tells you how much overlap there is between the 2 variables
  - SAMPLE LANGUAGE: "We conclude that about 80.5% of the variability in the variable Y can be explained by the predictor X"
- How to make a correlation scatter plot in excel: highlight over the data for X and Y > click insert > the scatter plot picture > the first scatter plot. Either make sure the X data is off the left and Y data is on the right, or you can right click on the chart and click select data then edit then make sure the correct data is highlighted for X and Y. you can also click the plus in the top corner of the add chart elements

- The Traditional T-test: Concepts and Demonstration:
  - **<u>T-test</u>**: Experimental and quasi-experimental research designs investigate whether a categorical independent variable (X) is related to a continuous dependent variable (Y). We often compare mean scores from two groups using the t-test
    - Many categorical variables are dichotomous
    - Significant relationship between the dichotomous variable (X) and continuous variable (Y) = Significant mean difference of Y between two X groups
    - A large t value and a small p value means the groups are significantly different from each other, so their bell curves won't overlap much
    - The t-statistic puts everything into the same scale so you can compare them
    - Steps to compute a t-test:
      - 1) State H_0 and H_1 (can be directional < or > or = or nondirectional = or ≠):
        - H_0: $\mu1-\mu2 = 0$ [aka there is no difference between the means of the groups – can also be written as Y with a line over it to signify mean]
        - H_1: $\mu1-\mu2 \neq 0$
      - 2) Set the criteria:
        - Two-tailed test with $p<.05$ with df being total sample size – 2, then find tcrit value.
      - 3) Compute test statistic:

$$\boxed{t = \frac{\bar{Y}_1 - \bar{Y}_2}{S^2_{pooled}}} \qquad \boxed{S^2_{pooled} = \sqrt{\frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_1^2}{N_1 + N_2 - 2}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}$$

      - 4) Make a decision:
        - If our t value is greater than the crit value, then the results are significant, and we can reject the null hypothesis because this shows there is a significant difference between the two groups
        - Reject or retain the null and what does that mean
        - t-test: $t(238) = -14.32$, $p < .001$, $d = 1.85$
    - How to compute a t-test in SPSS: analyze > compare means > independent samples t test > test variable = DV

- All of the data needs to be put into one variable then the group membership needs to be dichotomized to -1 and 1 then you can run these stats

Lab 5 – Correlation and T-test:

- Types of T-tests:
  - **Independent samples t-test**: comparing means for 2 independent groups (I.e., intervention and control group)
    - Steps to perform an independent samples t-test:
      - Step 1:
        - RQ: Is there a mean level difference in post-intervention GPA between the intervention and control curriculum?
          - H0: µ Int.GPA = Cnt.GPA
          - H1: µ Int.GPA ≠ Cnt.GPA
      - Step 2:
        - criteria: a = 0.05
        - Assumptions:
          - Continuous DV
          - Two categorical independent groups in IV (i.e., male and female)
          - Independent Data
            - No participants are in more than one group
          - Normally distributed for both groups
            - Steps to do a test of normal distribution in SPSS: Analyze > descriptive statistics > explore > move over DV into dependent list > move over grouping variable into factor list > plots > normality plots with tests > check histograms, uncheck stem and leaf
            - Can check skewness/kurtosis, Shapiro-Wilk test, Kolmogorov Smirnov test, histogram) – see breakdown earlier in study guide
          - Homoscedasticity/Homogeneity of variance (HOV) – aka we want the variances of the populations being compared to be equal
          - No significant outliers
      - Step 3:
        - Steps to compute an independent samples t-test in SPSS: analyze > compare means > independent samples t-test > move over DV into test variable > move and define grouping variable aka IV (use -1 and 1) > check box for “estimate effect sizes” to get Cohen’s d
      - Step 4:
        - Make a decision

- **Effect Size (Cohen's d):** assesses the magnitude of difference between the means of two groups using the following series of equations:

$$\boxed{d = \frac{|\bar{X}_1 - \bar{X}_2|}{s_p} = t \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \qquad s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}}$$

Verbal Label	Magnitude of *d*
No effect	*d* = 0
Small effect	*d* ≤ .20
Medium effect	*d* ≈ .50 (e.g., *d* between .20 and .79)
Large effect	*d* ≥ .80

- Output in SPSS looks like this:

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Post intervention GPA	Cohen's d	.16263	-1.849	-2.150	-1.545
	Hedges' correction	.16314	-1.843	-2.143	-1.540
	Glass's delta	.16232	-1.852	-2.198	-1.505

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

- To compute Cohen's d based on descriptives follow this:

Group Statistics

	Math Curriculum Used	N	Mean	Std. Deviation	Std. Error Mean
Post intervention GPA	Control	120	3.2006	.16294	.01487
	Persistence Based	120	3.5013	.16232	.01482

$$d = \frac{|\bar{X}_1 - \bar{X}_2|}{s_p}$$

where:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$d = \frac{|3.5013 - 3.2006|}{.1626277945}$$

where:

$$s_p = \sqrt{\frac{(120 - 1)(0.162318)^2 + (120 - 1)(0.162937)^2}{120 + 120 - 2}}$$

d = 1.849

OR

Independent Samples Test

		Levene's Test for Equality of Variances F	Sig.	t-test for Equality of Means t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Post intervention GPA	Equal variances assumed	.001	.976	-14.321	238	<.001	<.001	-.30067	.02100	-.34203	-.25931
	Equal variances not assumed			-14.321	237.997	<.001	<.001	-.30067	.02100	-.34203	-.25931

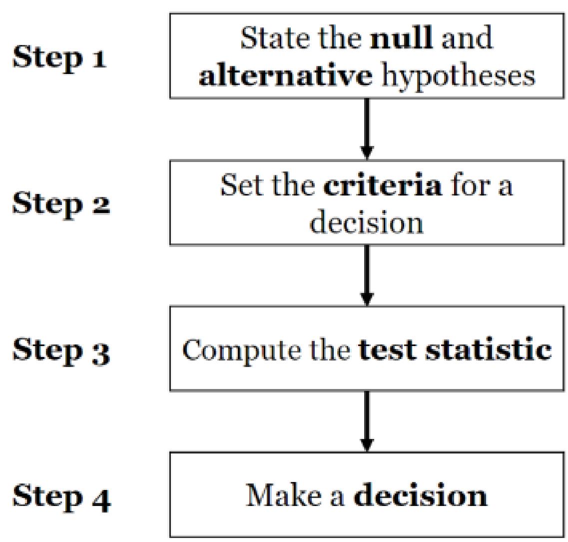
$$d = t \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \qquad d = -14.321 \cdot \sqrt{\frac{120 + 120}{120 \cdot 120}} \qquad d = -1.849$$

Group Statistics

	Math Curriculum Used	N	Mean	Std. Deviation	Std. Error Mean
Post intervention GPA	Control	120	3.2006	.16294	.01487
	Persistence Based	120	3.5013	.16232	.01482

        - Full APA write-up example for independent t-test:
            - RQ: Is there a mean level difference in post-intervention GPA between the intervention and control curriculum?

              "An independent-samples t-test was used to assess the effectiveness of a recently developed persistence-based curriculum. It was predicated that 8th grade students who were instructed using the persistence-based curriculum would attain a higher mean GPA than those in the control group by the end of the term. The assumption of normality was upheld for both groups via the visual examination of histograms, non-significant Shapiro-Wilk's tests [persistence curriculum ($W(120) = .995$, $p = .950$); control curriculum ($W(120) = .984$, $p = .165$)], and skewness [persistence = -.124; control = .061] and kurtosis [persistence = -.784; control = -.047] values within an acceptable range. Similarly, the assumption regarding homogeneity of variance was upheld by a non-significant Levene's test ($F(238) = .001$, $p = .976$), indicating equal variances can be assumed. It was found that students in the persistence group ($M = 3.50$, $SD = .16$) were observed to have overall higher post-intervention GPA than those in the control group [($M = 3.20$, $SD = .16$); $t(238) = -14.32$, $p < .001$, $d = 1.85$)]. Given the substantial observed difference and large effect size, our results support the conclusion that the persistence-based curriculum is associated with gains in 8th students' academic achievement in the form of GPA."
    - **<u>Paired samples t-test</u>**: comparing means from the same group at different times
    - **<u>One sample t-test</u>**: comparing mean of a single group against a known mean to determine if they are significantly different



        - I.e., Step 1: RQ: Is the sample pre-intervention mean GPA different from the school GPA of 3?
            - H0: μ Int.GPA = 3
            - H1: μ Int.GPA ≠ 3
        - Step 2:
            - criteria: a = 0.05
            - Assumptions:
                - Continuous DV
                - Independent Data
                - Normally distributed
        - Step 3:
            - Steps to do a one-sample t-test in SPSS: Analyze > compare means > one-sample t-test > move over variable of interest > ok

One-Sample Test

	Test Value = 3						
			Significance			95% Confidence Interval of the Difference	
	t	df	One-Sided p	Two-Sided p	Mean Difference	Lower	Upper
Pre interventuion GPA	.088	239	.465	.930	.00092	-.0197	.0216

- $t(239) = .088$, $p = .930$

  - Step 4:
    - Decision: retain the null, the sample pre-intervention mean GPA is not significantly different from the school GPA of 3

- Writing Results in APA:
  - What to include:
    - RQ and hypotheses
    - The test ran
    - Assessment of assumptions
    - Test statistic
    - Your interpretation of the results
  - Example write up:
    - "RQ: Is there a mean level difference in post-intervention GPA between the intervention and control curriculum?

      An independent-samples t-test was used to assess the effectiveness of a recently developed persistence-based curriculum. It was predicated that 8th grade students who were instructed using the persistence-based curriculum would attain a higher mean GPA than those in the control group by the end of the term. The assumption of normality was upheld for both groups via the visual examination of histograms, non-significant Shapiro-Wilk's tests [persistence curriculum ($W(120) = .995$, $p = .950$); control curriculum ($W(120) = .984$, $p = .165$)], and skewness [persistence = -.124; control = .061] and kurtosis [persistence = -.784; control = -.047] values within an acceptable range. Similarly, the assumption regarding homogeneity of variance was upheld by a non-significant Levene's test ($F = .001$, $p = .976$), indicating equal variances can be assumed. It was found that students in the persistence group ($M = 3.50$, $SD = .16$) were observed to have overall higher post-intervention GPA than those in the control group [($M = 3.20$, $SD = .16$); $t(238) = -14.32$, $p < .001$, $d = 1.85$)]. Given the substantial observed difference and large effect size, our results support the conclusion that the persistence-based curriculum is associated with gains in 8th students' academic achievement in the form of GPA."

Lecture 6 – One-Way ANOVA:

- The F statistic for an ANOVA is equivalent to the t statistic for a t-test

- **<u>Analysis of Variance (ANOVA):</u>** used to test if two variables are related. It compares means by looking for variability between those means and tells you how much variation there is between groups and within groups:
    - DV (Y) is continuous
    - IV (X) is categorical (the grouping variable)
- Equivalent to a t-test but can have more levels (groups) tested at once
- If the means of k different groups are all the same, then the variance between them is 0
- Assumptions for using an ANOVA:
    - Continuous DV and categorical IV
    - No significant outliers
    - Normality/normally distributed:
        - Look at the distribution of outcomes for each group
    - Random sampling:
        - Random sampling is required to avoid bias in sampling
    - Independent data:
        - Look out for any clustering if scores, that could indicate that the observations may not be independent of each other
    - Homogeneity of variance/homoscedasticity:
        - This is the assumption that the variance within groups is consistent for different groups, so they all vary relatively the same way/amount
        - if this assumption is not met, then the variance between groups in the numerator of the F ratio/test statistic may be inflated leading to an increased likelihood of committing a Type I error (when you say there is a significant difference but there really is not)
- Steps to doing an ANOVA:
    - 1) State the null and alternative hypotheses:
        - H_0: $\mu_1 = \mu_2$… [μ represents the population mean for each group]
        - H_1: At least one mean is different
    - 2) check assumptions above
    - 3) Compute SST:

$$\boxed{SSE = \sum(Y_i - \widehat{Y}_i)^2}$$

- 4) Compute the df:

Degrees of freedom between groups

$$df_{BG} = k - 1$$

k: number of levels/groups/conditions

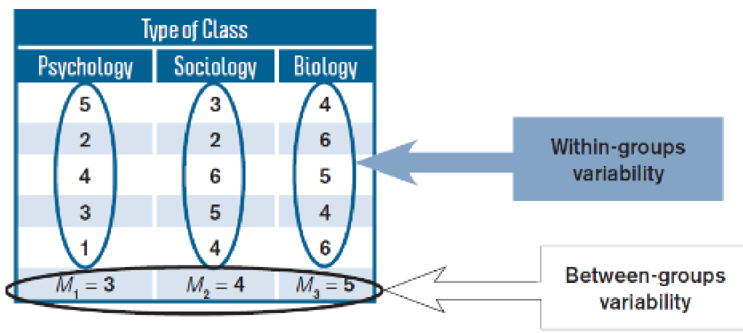
Degrees of freedom within groups (error)

$$df_E = N - k$$

Sum of degrees of freedom

$$df_{Tot} = N - 1$$

- 5) Compute F ratio:



$$F_{obt} = \frac{MS_{BG}}{MS_E} = \frac{\text{Variance between groups}}{\text{Variance within groups}}$$

Proportion of variability in *Y* explained by *X*

Proportion of variability in *Y* *un*explained by *X*

$$df_{BG} = k - 1$$

$$df_E = N - k$$

n = number of participants per group
N = total number of participants in a study
k = number of groups

- MS: the mean square. So instead of getting the sum of squares of error, we will be getting the mean square
- BG: between group
- E: the error

$$F_{obt} = \frac{\boxed{SSB/df}}{\boxed{SSW/df}}$$

$$SSB = \left[\frac{\left(\sum Y_{Group1}\right)^2}{N_{Group1}}\right] + \left[\frac{\left(\sum Y_{Group2}\right)^2}{N_{Group2}}\right] - \frac{(\sum Y)^2}{N_{Total}}$$

$$SSW = \sum Y^2 - \left\{\left[\frac{\left(\sum Y_{Group1}\right)^2}{N_{Group1}}\right] + \left[\frac{\left(\sum Y_{Group2}\right)^2}{N_{Group2}}\right]\right\}$$

  - HINT: values for the F statistic should all be positive. F = 1 means no mean difference because the between and within group variabilities would be the same. The higher your F the greater chance of it being significant (would mean there is a much higher variability between groups than you see within groups)

- 6) Complete F table:

1	*2*	*3*	*4*	*5*
Source	*SS* ÷	*df* =	*MS*	*F*
Sex	SSB	*K* − 1	SSB/*df*	MS_B/MS_W
Residual	SSW	*N* − 1 − (*k* − 1)	SSW/*df*	
Total	SST	*N* − 1		

- 7) Compare the calculated F value against the F crit value in a table. If F value is GREATER than F crit then we can say there is a significant difference between the means of the groups. When looking at the F crit table, the df between is horizontal and df within is vertical
- 8) Make a decision and write up F values (I.e., F(1,10) = 41.227, p < .001). A bar graph can be used to visually depict the differences as well
  - EXAMPLE LANGUAGE: "We conducted the analysis of variance to examine whether there was a mean difference of depression scores between subjects who were diagnosed with depression and who were not. The ANOVA results showed that there was a significant mean difference between the two groups (F(1, 22) = 117.42, p<.001)."

- Steps to compute an ANOVA in SPSS:

- Version 1: analyze > general linear models > univariate > DV (continuous) goes in DV spot, IV goes in fixed factor spot > options > select homo of var. (HOV) test and effect size > ok

Univariate Analysis of Variance

[DataSet0]

Between-Subjects Factors

		N
VAR00003	-1.00	12
	1.00	12

Tests of Between-Subjects Effects

Dependent Variable: VAR00002

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1014.000[a]	1	1014.000	117.411	<.001
Intercept	4056.000	1	4056.000	469.642	<.001
VAR00003	1014.000	1	1014.000	117.411	<.001
Error	190.000	22	8.636		
Total	5260.000	24			
Corrected Total	1204.000	23			

a. R Squared = .842 (Adjusted R Squared = .835)

- Version 2: analyze > compare means > one-way ANOVA > DVs (continuous) go to dependent list and IV (group) goes to factor list > option > Homogeneity of variance (homoscedasticity) test > check effect size

Oneway

ANOVA

VAR00002

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1014.000	1	1014.000	117.411	<.001
Within Groups	190.000	22	8.636		
Total	1204.000	23			

ANOVA Effect Sizes[a]

		Point Estimate	95% Confidence Interval Lower	Upper
VAR00002	Eta-squared	.842	.675	.898
	Epsilon-squared	.835	.661	.893
	Omega-squared Fixed-effect	.829	.651	.889
	Omega-squared Random-effect	.829	.651	.889

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

  - Point estimate eta-squared (η^2) tells you the effect size (same as the R^2)

- EXAMPLE LANGUAGE (test-statistic write-up): "A one-way ANOVA suggests that there are mean level differences in children injuries depending on their costume type ($F(3, 26) = 8.32$, $p < .001$, $\eta 2 = .49$). The moderate effect size suggests that 49% of the observed variance in injury score can be explained by the costume type worn when the child was admitted to the emergency room."
- EXAMPLE LANGUAGE (full APA write-up): "A one-way ANOVA was used to test for mean differences in children's injury score depending on the type of Halloween costume they were wearing when checked into the emergency room. It was predicted that there

would be differences among the costume types, though no specific group differences were made a priori. Assumptions of the one-way ANOVA were upheld. For example, data was generally normally distributed with histograms following a normal curve and a significant Shapiro-Wilk's test ($W(30) = .97$, $p = .536$). Similarly, HOV was upheld via a significant Levene's test ($F(3,26) = .891$, $p = .459$). The results of the omnibus test suggests that there are mean level differences in children injuries depending on their costume type ($F(3, 26) = 8.32$, $p < .001$, $\eta 2 = .49$). Notably, the large effect size suggests that 49% of the observed variance in injury score can be explained by the costume type worn then the child was admitted to the emergency room."

- You can get the same information by running an independent samples t-test (The second table of this output labeled "independent samples test" gives you the F statistic as well under the "equal variances assumed" line and levene's test tells you if the value is significant)
- You can also get the same information from running a linear regression (in the output under "model summary," the R value is the correlation coefficient. under the "coefficients" table, the table under "unstandardized B for the constant" is the intercept, and the value below that next to the name of the variable is the slope coefficient) (????)

- Testing two groups using model comparison:
  - You can use a regression model to get these results as well (in the beginning of the semester we used the approach for a continuous variable, but it can also be used for a categorical variable. See steps from lecture 4, page 14 of study guide