Final Exam Cheat Sheet

Critical p-values:

- P-values must be **equal to or smaller** than these levels to reject the null:
  - * $p \leq 0.05$
  - ** $p \leq 0.01$
  - *** $p \leq 0.001$
- If p-values are **larger** than these levels, we "accept" the null.

Correlation (r) – Strength & direction of the relationship

- 2 interval/ratio (I/R) scale variables
- Direction: (+) or (–) r-value
- Strength: r-values vary from -1 to 0 to +1. In either direction +/- :
  - 0.0 = no relationship
  - .01 to 0.19 very weak
  - 0.20 to 0.40 weak
  - 0.41 to 0.70 moderate
  - 0.71 to .90 strong
  - 0.91 to 0.99 very strong
  - r = 1.0 or r = - 1.0 indicates a *perfect* positive or negative relationship
- Correlation Hypotheses:
  - Null hypothesis- The correlation for the two variables equals zero (r=0).
  - H_R (2-tailed)- The correlation for the two variable does not equal zero ($r \neq 0$, 2 tailed).
  - H_R (1-tailed)- The correlation for the two variables is either positive ($r>0$) or negative ($r<0$).
- r^2 x 100 = % of the variation of the DV that is explained by the IV.

OLS Regression – The line of best fit

- Dependent Variable (DV) – I/R Scale; Independent Variables (IV) – I/R Scale or Dummy (0,1)
- r^2 x 100 = % of the variation of the DV that is explained by the IV. Use adjusted r^2 if multi.
- Examine whole model significance with ANOVA (F-test) - testing H0: $r^2 = 0$; HR: $r^2 > 0$
- Examine if the b coefficient is statistically significant (T-test, is t-value greater than critical t) - testing H0: b= 0 ; HR b > 0 (one-tail positive) HR b< (one-tail negative) HR $b \neq 0$ (two-tail)
- If b's significant, the model predicts:
  - I/R scale IV: For each unit increase in IV, the DV increases or decreases by b-value.
  - Dummy IV: Included group(s) increase or decrease by b-value compared to excluded group.
  - For multivariate: Use controlling for all other variables language.
- Making predictions – use values of IVs to predict the DV (assume all b's are significant)
  - Simple regression: $Y' = a + bx$
  - Multi-regression: $Y' = a + b_1x_1 + b_2x_2 + b_3x_3 \ldots$
  - If IV is I/R: value if x is actual value of IV
  - If IV is 0/1: Value of x is 1 (yes) or 0 (no)

OLS Regression with Log Dependent Variables

- Uses the same procedures as regular OLS regressions.
- Multiply b's by 100 - Interpret as % increases or decreases

Logistic Regression – Probability of being in the YES category

- DV - Dummy (0,1); IVs - I/R Scale or Dummy
- Pseudo r^2 - Cox & Snell and Nagelkerke
- Omnibus Test of Model Coefficients Box: Model chi-square for model significance
- If b's are significant:
  - Manipulate b's into odds ratios: e^b = Exp(B)
  - If Exp(B) is greater than 1, then [Exp(B)-1] x 100 = % more likely
  - If Exp(B) is less than 1, then [Exp(B)-1] x 100 = % less likely
  - I/R scale IV: The model predicts for each unit increase in the IV, "DV condition" is more or less likely.
  - Dummy IV: Included group(s) more or less likely compared to excluded group.

- Odds Ratio Comparison: $\frac{e^{bx}}{e^{bx}} = OR$ To interpret: (OR - 1) * 100
  - More or less likely compared to another to be in the DV condition

- Probabilities: $\frac{e^{a + b1x1 + b2x2 + b3x3 + bkxk}}{1 + e^{a + b1x1 + b2x2 + b3x3 + bkxk}} = P$ To interpret: P*100
  - Likelihood that an individual with a set of characteristics will be in the DV condition

Test	Independent Variable	Dependent Variable	Testing
OLS Regression	0/1 or I/R	I/R	B indicates differences (with dummy variables) or changes (with I/R variables) in the DV for changes in the IV.
OLS Regression w/log DV	0/1 or I/R	I/R w/log DV	Interpret the b's by multiplying them by 100 to get % change (for I/R variables) or % difference (for 0/1 dummy variables)
Logistic Regression	0/1 or I/R	0/1	The likelihood of being in the group (say, poverty or having mental health problems). Interpret the odds ratio Exp(B). You know it's a logistic regression analysis when there is the column with exp(B), the odds ratio.