Study Questions: Multi-Population Statistics

CHAPTER 9. TWO SAMPLE HYPOTHESIS TESTS

- What is 2-sample hypothesis testing used for?
- Are X and Y independent in 2-sample testing?
- Explain when to use z-test and t-test for two sample testing.
- If $X \sim \text{Normal}(\mu_X, \sigma_X)$ and $Y \sim \text{Normal}(\mu_Y, \sigma_Y)$ then what is the distribution of $\left(\overline{X} - \overline{Y}\right)$?
  (Assume X and Y are sampled n and m times)
- Explain when to use F-test and t-test for two sample testing.
- How many degrees of freedom are there in a two-sample t-test if X and Y are **both** sampled n times?
- What is the "pooled sample variance" of $\{X_i\}$ and $\{Y_j\}$?
- How many degrees of freedom are there in a **pooled variance** two sample t-test if X and Y are **both** sampled n times?
- When is pooled sample variance testing used?
- Why use pooled sample variance 2-sample testing? (instead of non-pooled variance)
- How does the pooled sample variance compare to the individual sample variances of $\{X_i\}$ and $\{Y_j\}$?
- How can you check whether pooled sample variance testing is appropriate?
- When is paired sample testing used, and why use it?
- What is blocking, and why use it?

CHAPTER 10. ANALYSIS OF VARIANCE

- What does ANOVA stand for?
- When do you use ANOVA?
- Why use ANOVA instead of multiple 2-sample t-tests?
- When is ANOVA equivalent to a 2-sample t-tests?
- What hypothesis does the F-test in ANOVA test?
- What are the "Residuals" in ANOVA?
- In ANOVA, what does MSF, MSE, and MST stand for?
- In ANOVA, what is the relationship between SSF, SSE, and SST?
- In ANOVA, how many degrees of freedom are there for SSF, SSE, and SST?
- In ANOVA, what is the relationship between the degrees of freedom for SSF, SSE, and SST?
- In ANOVA, what are the distributions of SSF, SSE, and SST?
- Give a formula for the test statistic F used in ANOVA. What is its distribution?
- Give a formula for "effect size" η^2 in ANOVA.
- What values are possible for η^2 in ANOVA? Which values are better?
- What does Tukey's HSD Test compute? Why use it?

CHAPTER 12. SIMPLE LINEAR REGRESSION

- When is regression analysis used?
- Are X and Y independent in regression analysis?
- What is the "Linear Regression Model"?
- What is the relationship between ANOVA and regression analysis?
- What does regression "SSR" correspond to in ANOVA?
- What does Regression "R^2" correspond to in ANOVA?
- In regression analysis we write $Y = (\beta_0 + \beta_1 x) + \varepsilon$. What are β_0, β_1, and ε?
- What does "regression line" mean in regression analysis?
- Write a formula for the regression line $\hat{y}(x)$ in terms of the sample means $\bar{x}$, $\bar{y}$.
- Name one point that all regression lines go through.
- Give a formula for $\hat{\beta}_1$ in terms of Variance and Covariance.
- What is the relationship between β_1 and Correlation?
- What hypothesis does the F-test in regression analysis test?
- In regression analysis, what does MSR, MSE, and MST stand for?
- In regression analysis, what is the equation relating SSR, SSE, and SST?
- In regression analysis, how many degrees of freedom are there for SSR, SSE, and SST?
- What are the "residuals" in regression analysis?
- What are the "observed" and "fitted" values in regression analysis?
- Give a formula for the coefficient of determination R^2.
- What does it mean when R^2 is big?
- What values are possible for R^2 in regression analysis? Which values are better?
- How many degrees of freedom do $\hat{\beta}_0$ and $\hat{\beta}_1$ have?
- Name four things that affect the size of the confidence interval around the regression line.
- At which value is the confidence interval for a regression line the most narrow?
- Why does the confidence interval for a regression line get wider as x moves away from $\bar{x}$?