Review

- Regression, χ^2, ANOVA
- What Test or Confidence Interval?
- Top Ten Mistakes
- Thoughts on Probability

Regression, χ^2, ANOVA

- **Regression:** Two quantitative variables, interested in predicting y given x or see how y changes as x changes.
- **Goodness of Fit:** Given a frequency table, does it fit a known distribution?
- **Test for Independence:** Are two Categories Independent or Dependent?
- **Test for Homogeneity:** Two frequency tables given. Do they have the same distribution?
- **1-Way-ANOVA:** Are all (more than 2) means the same? More than 2 quantitative variables.

What Test or Confidence Interval?

Estimate or Decision	Mean or Proportion	Sigma Known or Unknown	1 or 2 Samples	Dependent or Independent	**Conclusion**
Estimate	Mean	Known	1	NA	ZInterval
Estimate	Mean	Unknown	1	NA	TInterval
Estimate	Prop	NA	1	NA	1PropZInt
Estimate	Mean	Known	2	Independ.	2SampZInt
Estimate	Mean	Unknown	2	Independ.	2SampTInt
Estimate	Mean	Known	2	Depend.	ZInt (L1-L2)
Estimate	Mean	Unknown	2	Depend.	Tint (L1-L2)
Estimate	Prop	NA	2	NA	2PropZInt

What Test?

Estimate or Decision	Mean or Proportion	Sigma Known or Unknown	1 or 2 Samples	Dependent or Independent	Conclusion
Decision	Mean	Known	1	NA	ZTest
Decision	Mean	Unknown	1	NA	TTest
Decision	Prop	NA	1	NA	1PropZTest
Decision	Mean	Known	2	Independ.	2SampZTest
Decision	Mean	Unknown	2	Independ.	2SampTTest
Decision	Mean	Known	2	Depend.	ZTest (L1-L2)
Decision	Mean	Unknown	2	Depend.	TTest (L1-L2)
Decision	Prop	NA	2	NA	2PropZTest

Top 10 Mistakes

1. Say $n > 30$ for a proportion.
2. Say $np, nq > 5$ for a mean.
3. Say $n > 30$ to justify Z instead of T.
4. Use definitive language for regression.
5. Confuse the three χ^2 tests.
6. Accepting H_0 in the conclusion statement.
7. Refer to the sample instead of the population when interpreting the CI or Hyp test.
8. Forgetting to subtract $P(A \text{ and } B)$ when finding $P(A \text{ or } B)$.
9. Trying to multiply probabilities when reading from a table.
10. Forgetting to divide by the square root of n when finding a probability involving means.

Thoughts on Probability

General Probability

$$P(A) = \frac{\#A}{\#S}$$

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

A and B independent:

$$P(A \text{ and } B) = P(A)P(B)$$

$$P(A \mid B) = P(A)$$

Tables Row A, Column B

$$P(A \text{ and } B) = \frac{AB \text{ cell}}{\text{Grand Total}}$$

$$P(A) = \frac{\text{Row } A \text{ Total}}{\text{Grand Total}}$$

$$P(B) = \frac{\text{Column } B \text{ Total}}{\text{Grand Total}}$$

$$P(A \mid B) = \frac{AB \text{ Cell}}{\text{Column } B \text{ Total}}$$