

Quiz 3

1. We call it “analysis of variance” instead of “analysis of means” because we compare the means of groups relative to the variance of the individual groups.
2. For the t-statistic, the numerator is the difference between the sample mean and the true mean, and the denominator is the standard deviation for the sample mean (or the difference in two sample means over the variability of the difference between the two sample means for the two-sample t-statistic). The F-statistic is measured as the size of the among-group’s variability over the within-group’s variability. Both statistics measure the relationship between response and predictor variables (to see if there is a relationship (F) or if something is an important predictor (t)).
3. Heteroscedasticity (inhomogeneity of variance) can increase the variability of the regression parameters. We can look for it by looking for the “megaphone” pattern in the residual scatterplot or “doglegs” in the normal probability plot.
4. In fitting a bivariate regression line we minimize the individual deviations of observations.
5. Confidence intervals assess the variability of the regression line and prediction intervals assess the variability of data.
6. If the assumptions of regression analysis are not violated, then the estimates are unbiased and have minimum variance. The assumptions are:
 - a. the prediction errors or residuals are assumed to be independent, identically normally distributed random variables, with a mean of 0 and a standard deviation of s ,
 - b. the X ’s (predictor or independent variables) are known without error,
 - c. the X ’s are not correlated.
 - d. the correct model has been specified
7. For ANOVA, the F-test tests the hypothesis that group means are not different, while for regression analysis the F-test tests the null hypothesis that the proportion of the variance of the response variable is not significant. The F-test and R^2 measure the strength of a relationship. The sum-of-squares and F-statistic grow larger when the between group variance grows more than the within-group variance, or when the regression line fits the data accurately.