Midterm 2 Study Guide

This midterm is about probability concepts, as well as linear modeling.

This is meant to be a representative sampling of the key concepts you will need to know, and it is not meant to be exhaustive. You should make sure that you are comfortable with Quizzes 4-9 and Homework Assignments 7-14.

1. What graphs are appropriate when we want to examine the relationship between two variables? What types of variables does each apply to?
2. What terms do we use to describe the relationship between two scalar variables?
3. What does the correlation coefficient r measure?
4. When we have a regression line fitted to some data, explain the following terms: predicted value, actual value, residual, sum of squared residuals.
5. What is the key property that makes the "*least squares regression line*" special?
6. What is the meaning of "r-squared" in the context of the least squares regression line?
7. How does the *residual plot* work? What do we expect from it if we have a suitable fit?
8. Outliers far in the x direction and only far in the y direction affect the least squares regression line in very different ways. Explain.
9. What are the key characteristics of a *random phenomenon*?
10. What can we say about the probabilities of outcomes in a random phenomenon / probability model? Can they be any numbers? There are some restrictions to them (both on what numbers they can be, and on what their sum must be).
11. What is an *event* and how do we define its probability?
12. What is the *complement* of an event? What is a formula for its probability?
13. What is the *union* of two events, what is the *intersection* of two events? What rule must their probabilities obey?
14. What is the definition of *conditional probability*? Include a formula and also a more meaningful verbal definition.
15. What does the *multiplicative rule* say?
16. When do we say that two events are *independent* of each other? Provide some examples both of events that are independent of each other and events that are not independent of each other.

17. In general when do we tend to multiply probabilities? When do we tend to add them?

18. When can I say that $P(A \cap B) = P(A)P(B)$?

19. When can I say that $P(A \cup B) = P(A) + P(B)$?

20. What are the different parts in a tree diagram / decision tree? What probabilities do we associate with each part?

21. What differentiates *random variables* from other random phenomena? Give examples of random variables

22. What conditions must we meet in order to be in the *binomial setting*?

23. What is the formula for the *binomial distribution*?

24. How do we compute the mean and the standard deviation of a random variable?

25. How do we compute the mean and standard deviation of the binomial distribution?

26. What can we say about the mean and standard deviation of a linear transformation of a variable?

27. What can we say about the mean and standard deviation of a sum of two variables?

28. What do X and $\hat{p}$ represent in the binomial setting?