

Review Sheet for Test 2

HERE ARE MOST OF THE IMPORTANT THINGS YOU NEED TO KNOW:

1. that scatterplots, correlation and regression are only appropriate when both the explanatory and response variables are quantitative (Case III)
2. how to describe and interpret features of scatterplots, especially direction, strength and form
3. that correlation only describes the *linear* relationship between two quantitative variables
4. the basic properties of the correlation coefficient—unitless, ranges between -1 and 1 , strength depends only on magnitude, not sign
5. that Simpson's paradox can occur in Case III as well as in Case II
6. how to find the equation of the least squares regression line if the means and sds of each variable, along with the correlation coefficient, are given
7. how to use a regression line to predict the value of the response variable from a given value of the explanatory variable
8. what extrapolation is, and that it is unreliable
9. how to interpret the slope of the regression line in context
10. what the interpretation of r^2 is
11. how to calculate and interpret residuals
12. how to interpret a residual plot
13. what regression outliers and influential observations are
14. that association does not imply causation, and what lurking variables are
15. what the meaning of the term *statistical inference* is (part of the Big Picture)
16. the meaning of the terms population, sample, sampling frame, parameter and statistic
17. the meaning of *bias*, and the various types of bias that can occur
18. the meaning of the term *sampling variability*, and how it is related to sample size (but *not* to population size, unless the sample is a large fraction of the population)
19. the different types of studies (see Topic 13)—their definitions and their strengths and weaknesses
20. the features of a good experiment—comparison, randomization, blindness (where applicable)
21. how to make a graphic to show the design of an experiment
22. what a placebo is, and what the placebo effect is
23. how to sketch a bell curve labeled with the mean, sd and values on the axis
24. how to estimate the mean and sd of a normal distribution from its graph
25. how to determine probabilities for a normally distributed variable (“forwards problems”)—less than, more than, between
26. how to determine the value of a normally distributed variable that achieves a given percentage above or below that value (“backwards problems”)