

## STATISTICS 8 - REVIEW FOR FIRST EXAM (CHAPTERS 1 TO 6)

### Know how to find, construct or calculate:

1. A two-way table (explanatory variable as rows) and corresponding conditional percents
2. A bar chart comparing two groups or conditions
3. All of the values in a five-number summary (median, quartiles, low and high)
4. Range and interquartile range
5. A stem and leaf plot
6. A bell-shaped curve with the Empirical Rule percentages marked
7. A standardized score, given a value, mean and standard deviation
8. A predicted (or estimated) value of  $y$ , given a regression equation and an  $x$ -value
9. The residual for a given individual, given that individual's values and a regression equation
10. Risk, relative risk, increased risk, odds and odds ratio
11. Expected counts for a chi-square test (compute and interpret)
12. The conclusion for a hypothesis test, based on the  $p$ -value
13. (Conservative) margin of error and 95% confidence interval

### Understand:

1. Reasons for outliers and what to do about them
2. How outliers affect various measures like mean, median, range, interquartile range
3. How to interpret a boxplot and side-by-side boxplots
4. Various shapes, how to recognize them and the influence on the mean versus median
5. The concept of standard deviation and how it is used for bell-shaped curves
6. How standard deviation and range are related for bell-shaped curves
7. What a standardized score represents and how it is useful
8. How to determine which is the explanatory variable and which is the response variable
9. A deterministic relationship versus a statistical relationship
10. How to interpret a scatter plot - look for linear trend (or not) and outliers
11. How to interpret correlation and  $r^2$ , and how correlation relates to slope (same sign)
12. How to interpret the slope and intercept in a regression equation and when they are useful (or not)
13. The two purposes for which regression is used, and when each is used (predicting  $y$  for an individual  $x$  or estimating the mean of all  $y$ 's at a given value of  $x$ )
14. The concept of the "least squares" line
15. How different types of outliers influence correlation and the regression equation
16. How combining groups distorts correlation and regression
17. Simpson's Paradox
18. How you can be misled by statistics about risk (pages 201-202)
19. The purpose of hypothesis testing
20. How to write the hypotheses for a chi-square test
21. The interpretation of a  $p$ -value
22. The possible conclusions for hypothesis tests, and why the null hypothesis should not be accepted
23. How the sample size affects the results of a hypothesis test (pages 213-215)
24. Types of samples and the advantages and disadvantages of them; when results extend to a population
25. How to interpret a confidence *interval* and a confidence *level*
26. Sources of bias in surveys and samples
27. Observational studies versus randomized experiments, the role of confounding variables and what conclusions can be made (cause-and-effect or not)
28. What defines a confounding variable and why they are more of a problem in observational studies
29. Principles of good experiments: randomization, blinding, blocking, placebo, control group
30. What is meant by a "treatment" in an experiment
31. Issues related to observational studies: retrospective, prospective, case-control
32. Overcoming difficulties encountered in experiments and observational studies (Section 4.4)
33. Interpretations of an observed correlation or relationship (Section 5.5)