

Introduction to Biostatistics (BIOS 4120)  
Breheny

Quiz 2 (Practice)

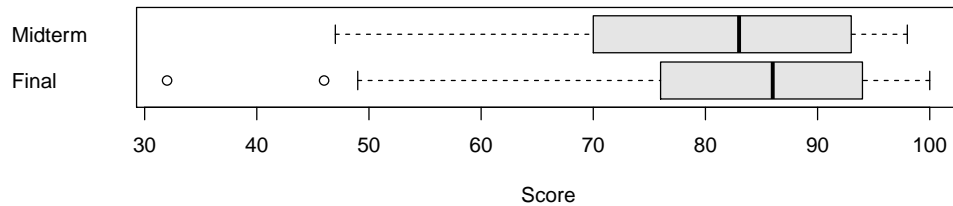
1. A dental cary is an invasion of the tooth by microorganisms. Caries develop gradually over time, starting out minor but eventually developing into cavities. I will refer to a cary that has not yet become a cavity as a “minor cary”. Let  $A$  denote the event that a randomly selected 5-year-old child has a cavity, and  $B$  denote the event that the child has a minor cary. A dental study conducted at the University of Iowa estimated that the probability that a randomly selected 5-year-old child has a cavity is .242, the probability that the child has a minor cary is .229, and the probability that the child has cavities, given that the child has minor caries, is .544.
  - (a) Are  $A$  and  $B$  independent? Justify your answer with numbers.
  - (b) Find the probability that a randomly selected child has both cavities and minor caries.
  - (c) Are  $A$  and  $B$  mutually exclusive? Justify your answer with numbers.
  - (d) Find the probability that a randomly selected child has cavities or minor caries or both.
  - (e) Find the probability that a randomly selected child has no minor caries, given that he or she has no cavities.

2. Blood-based tests for the presence of HIV are very accurate, but somewhat invasive. Less invasive saliva-based tests have also been developed. Below are the results of a Canadian study examining the accuracy of saliva based tests for HIV status:

Saliva test	HIV Status	
	Positive	Negative
Positive	358	2
Negative	10	886

- (a) What is the sensitivity of the saliva test?
- (b) What is the specificity of the saliva test?
- (c) Suppose that, in the population of interest, the prevalence of HIV is 240 per 100,000 individuals. If a randomly drawn individual from that population has a positive saliva test, what is the probability that he/she is actually HIV-positive?
3. A 2007 study performed a regression of quality of care (measured on a 100-point scale where 0 is the worst and 100 is the best) on number of chronic conditions. The study obtained an intercept of 55 and a slope of 2.2 per condition.
- (a) What is the predicted quality of care for a patient with no chronic conditions?
- (b) What is the predicted quality of care for a patient with 2 chronic conditions?
- (c) The correlation between quality of care and number of chronic conditions (i) is positive (ii) is negative (iii) could be either positive or negative – it is impossible to tell from the information given

- (d) You do not have enough information to calculate the correlation coefficient from the slope of the regression line. What additional information would you need to calculate it?
4. Below is a box plot of students' scores on the midterm and final exams from an "Introduction to Biostatistics" course I taught in Fall 2009 (this was back in the day when I gave a midterm instead of quizzes):



- (a) What was the median score on the midterm?
- (b) Were scores higher on the midterm or final?
- (c) What was higher, the mean score on the final, or the median score on the final?
- (d) Did anyone get a 100 on the midterm?
- (e) What percent of students scored below a 70 on the midterm?
- (f) Is the correlation between students' scores on the midterm and final likely to be (i) close to -1 (ii) somewhat negative (iii) close to 0 (iv) somewhat positive (v) close to 1
- (g) Suppose a student scores 1 standard deviation above average on the midterm. On the final, would you expect that student to score (i) more than 1 standard deviation above the mean (ii) less than one standard deviation above the mean (iii) right around 1 standard deviation above the mean?