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.

   No; $P(A) = .242 \quad P(A|B) = .544$

   (b) Find the probability that a randomly selected child has both cavities and minor caries.

   $P(A \cap B) = P(B)\,P(A|B) = .229(.544) = .125$

   (c) Are A and B mutually exclusive? Justify your answer with numbers.

   No; $P(A \cap B) \neq 0$

   (d) Find the probability that a randomly selected child has cavities or minor caries or both.

   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

   $= .242 + .229 - .125 = .346$

   (e) Find the probability that a randomly selected child has no minor caries, given that he or she has no cavities.

   $\frac{654}{758} = .863$

Cary

	Cavity Y	Cavity N	
Y	125	104	229
N	117	654	771
	242	758	1000

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:

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

(a) What is the sensitivity of the saliva test?

$$\frac{358}{368} = .973$$

(b) What is the specificity of the saliva test?

$$\frac{886}{888} = .998$$

(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?

$$P(D|+) = \frac{P(D)P(+|D)}{P(D)P(+|D) + P(D^c)P(+|D^c)}$$

$$= \frac{.0024(.973)}{.0024(.973) + .9976(.002)} = .539$$

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?

55

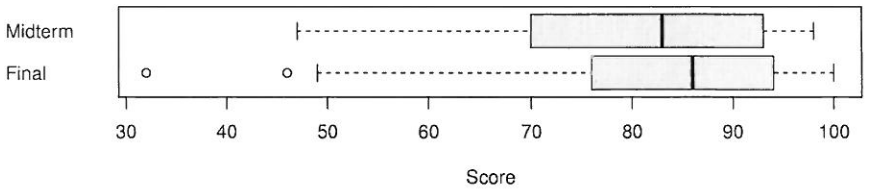
(b) What is the predicted quality of care for a patient with 2 chronic conditions?

$$55 + 2(2.2) = 59.4$$

(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 addition information would you need to calculate it? The SDs of quality of care and # of conditions

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?

83

(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?

No

(e) What percent of students scored below a 70 on the midterm?

25%

(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?