

MAT 2379, Introduction to biostatistics**Assignment 1**

Due date: Wednesday September 29, 2021 at 11:59 p.m. EST

Total 100 marks

- Solve the following problems from the textbook “Expect the Unexpected: a First Course in Biostatistics” (second edition), using a Faculty of Science standard calculator (TI30, TI34, Casio fx-260 and Casio fx-300):

1.2, 2.8, 3.6, 6.2, 6.6

The statement of each problem is included below.

- Each problem is worth 20 marks.
- Please complete all the questions in this assignment, in the order given above. However, due to our limited TA resources, it is possible that not all the questions will be marked. You will not be informed beforehand which questions will be marked. If not all the problems are marked, the total number of marks will be adjusted accordingly.

Problem 1.2 In humans, the presence of freckles is a dominant trait, and their absence is a recessive trait. The eyebrow shape is also a genetically inherited trait, with separated eyebrows being dominant and joined eyebrows being recessive. In a couple, the woman has freckles and separated eyebrows, and she is heterozygous for both traits. The man does not have freckles but he has joined eyebrows. What is the probability that their offspring has freckles and separated eyebrows? Draw the tree diagram or Punnett square to justify your answer.

Problem 2.8 Consider 1150 students that were enrolled in both Biology and Chemistry. Among these students only 50 got an A+ in Chemistry. However, 375 students got an A+ in Biology. There were 45 students that got an A+ in both Biology and Chemistry. We select a student at random from these 1150 students.

- (a) What is the probability that the selected student got an A+ in Biology but not in Chemistry?
- (b) What is the probability that the selected student got an A+ in Chemistry but not in Biology?
- (c) What is the probability that the student did not get an A+ in Biology or did not get an A+ in Chemistry?
- (d) What is the probability that the student got an A+ in at least one of these two courses?
- (e) What is the probability that the student did not get an A+ in Biology and did not get an A+ in Chemistry?

Problem 3.6 In the 1960s, Warner and Greenberg proposed ways to maintain the confidentiality of the respondent on a survey question by using a technique called *randomized responses*. If a question concerns a sensitive issue (e.g. criminal behavior), the respondent might not be

truthful. If we can convince the respondent that the answer will be confidential, then this person may answer truthfully. Warner proposed that a question concerning a sensitive issue be formulated as a question with Yes/No question and that the respondent be offered the question and its negation. Suppose that we would like to estimate the prevalence of bullying in a high school. We then ask the student the following two questions:

- (A) Have you bullied a classmate in the last two years? (Yes/No)
- (B) Have you not bullied a classmate in the last two years? (Yes/No)

We ask the respondent to throw a die and answer version (A) of the question if the result on the die is 1, and version (B) otherwise. Since we do not know the result on the die, we cannot tell if the respondent has answered version (A) or version (B) of the question. Assume that the respondents are answering truthfully.

- (a) Suppose that the probability that a respondent has bullied a classmate in the last two years is 13%. What is the probability that the respondent will answer “Yes” to the question.
- (b) Suppose that the probability that a respondent has bullied a classmate in the last two years is p . Express p as a function of the probability of answering “Yes” to the question.
- (c) Suppose that 275 out of 395 respondents answered “Yes” to the question. Approximate the probability that a respondent has bullied a classmate in the last two years.

Problem 6.2 One of the common symptoms of Alzheimer’s disease is the memory loss that disrupts daily life. This is just one of the 10 warning signs of Alzheimer’s disease. Another frequently encountered sign consists in having difficulty to complete familiar tasks. In a large group of early stage Alzheimer patients, 85% of them experience memory loss, 78% have difficulty completing familiar tasks and 67% show both signs.

- (a) What is the probability that an early stage Alzheimer patient shows one of these signs?
- (b) What is the probability that an early stage Alzheimer patient does not show any of these two signs?

Problem 6.6 A lab has a population of 25 fruit flies, of which 5 are black and 20 are grey. A sample of 2 flies is selected. Is the fact that the second selected fly is black independent of the first one being black?

- (b) Suppose now that there are 10,000 flies in the lab, of which 2,000 are black and 8,000 are grey. A sample of 2 flies is selected. Is the fact that the second selected fly is black independent of the first one being black?