

GER1000 2018 Sem 2
Quiz 10 and solutions

1. By “elderly”, we mean a person who is more than 65 years old. In Singapore, the percentage of elderlies among women is higher than the percentage of elderlies among men. Consider the following statements:

- I. In Singapore, the percentage of women among elderlies is higher than the percentage of women among the non-elderlies.
- II. In Singapore, the percentage of women is higher than the percentage of men among elderlies.

Which of the statement(s) is/are true?

- A. I only**
- B. II only
- C. I and II
- D. None of the above

Explanation: From the given information, $\text{rate}(\text{elderlies} \mid \text{women}) > \text{rate}(\text{elderlies} \mid \text{men})$, and thus women and elderlies are positively associated. This also means that $\text{rate}(\text{women} \mid \text{elderlies}) > \text{rate}(\text{women} \mid \text{non-elderlies})$, which is expressed by statement I. We cannot determine the percentage of women and men among the elderlies with the information given.

2. Tom tossed a coin into the air 7 times to test whether the coin is biased. The observed outcome was HHHTHHH. Consider the following statements:

- I. Our opinion about the null hypothesis is the same if the outcome was THHHHHH.
- II. We do not reject at the 5% significance level that the coin is unbiased.

Which of the statement(s) is/are true?

- A. I only
- B. II only
- C. I and II**
- D. None of the above

Explanation: The null hypothesis is that the coin is unbiased. The p-value is used as evidence against the null hypothesis (see Chapter 6, Unit p-values, slide 6), it does not give evidence for the null hypothesis. Since the p-value of the observation is 0.0625, we do not have enough evidence to reject the null hypothesis. The outcome THHHHHH is as extreme as the observed outcome and therefore the p-value is also 0.0625.

3. A researcher wanted to find the correlation between heights of father-and-son pairs. After collecting and analyzing his data, he realized that the device he had been using to measure height suffered from significant systematic bias causing every measurement to be too high by 10cm. He then corrected the values of all his analyses. After the correction, which values of the new data are expected to change significantly?

- A. The correlation coefficient between heights of father-and-son pairs.
- B. The standard deviation of son's height and father's height.
- C. The average son's height and average father's height.
- D. All of the above.

Explanation: The measurements are systematically off the mark in the same direction. The correlation coefficient does not change by adding a number to all the values of a variable (Chapter 2, slide 59). The spread or the SD deviation of the variables should not change either. Only the average of the heights would change.

4. The transcutaneous bilirubinometer is a new device that relies on flashes of light to calculate a baby's bilirubin level. In a sample of baby girls, the correlation coefficient between bilirubin level from transcutaneous bilirubinometer and the traditional heel prick test is 0.99. In a sample of baby boys, the correlation coefficient is also 0.99. If we combine the two samples together and calculate the correlation, what would the correlation coefficient be in the combined sample?

- A. More than 0.99
- B. 0.99
- C. Less than 0.99
- D. Cannot be determined

Explanation: We cannot determine the actual correlation coefficient without knowing the actual data points or the pattern of relationship between the two variables.

5. In a large department, it was discovered that, out of 64 people who have disease X, 48 had eaten tuna casserole from the canteen. Further investigation found that the same department also has 60 people who do not have disease X, although 20 of those had eaten tuna casserole from the canteen. What is the odds ratio of having disease X (for those who have eaten tuna casserole, relative to those who have not eaten)?

- (A) 6
- (B) $3/2$
- (C) $2/3$
- (D) $1/6$

Explanation: The odds of X among those who have eaten tuna is $48/20$, while the odds of X among those who have not eaten tuna is $16/40$. The odds ratio is 6.

	Having X	Not having X	Overall
Eat tuna	48	20	68
Do not eat tuna	16	40	56
Overall	64	60	124