GER1000 Quiz 3

1. A professor wants to know the percentage of left handed students in NUS. Since he is teaching a module in NUS this semester, he decides to do a survey in his class. From the single survey, he concluded that twenty percent of students in NUS are left handed. Which of the following fallacy was committed by the professor?

(A) Atomistic fallacy
(B) Ecological fallacy
(C) None of the above

Answer: C

Atomistic fallacy occurs when a person generalizes the correlation about individuals towards ecological correlation. Ecological fallacy occurs when a person deduces the inferences on correlation about individuals based on ecological correlation. In this question, the professor merely generalizes the result of the class to the entire NUS. Note that correlation and ecological correlation were not computed by the professor. Hence, both fallacies were not committed.

2. From all smokers in country A, 1,121 were randomly selected. From all non-smokers in the same country, 13,224 were randomly selected. All 14,345 subjects were examined for lung cancer. Data are summarised in the table.

	Lung cancer	No lung cancer	Row sum
Smokers	109	1012	1121
Non-smokers	115	13109	13224
Column sum	224	14121	14345

(I) The risk ratio for lung cancer between smokers and non-smokers in the country is estimated as 11.2.
(II) The odds ratio for lung cancer between smokers and non-smokers in the country is estimated as 12.3.

(A) (I) and (II) are true
(B) Only (I) is true
(C) Only (II) is true
(D) (I) and (II) are false

Answer: A

Since this is a cohort study, both population risk ratio and population odds ratio can be estimated from sample data. From the sample data, sample risk ratio = (109/1121) / (115/13224) = 11.2 and sample odds ratio = (109/1012) / (115/13109) = 12.3. Thus, the population risk ratio is estimated as 11.2 and the population odds ratio is estimated as 12.3.

3. In town B, men always marry women who are shorter than they are. Imagine a scatter diagram with Y axis representing men's height and X axis representing women's height where each point represents a married couple. Which of the following must be true about married couples from town B?

(I) The correlation between men's height and women's height is positive.
(II) The relationship between men's height and women's height may best be described by a line.

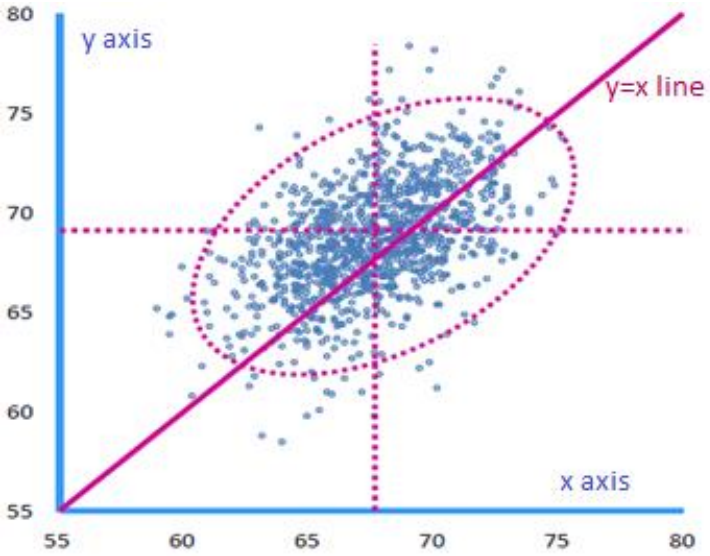
(A) (I) only
(B) (II) only
(C) (I) and (II)
(D) Neither (I) nor (II)

Answer: D

The data points of married couples from town B can be plotted on a scatter diagram with Y axis representing men's height and X axis representing women's height. Since men always marry women who are shorter than they are, all points will lie above the Y=X line. With the single condition that all points lie above Y=X line, we are unable to conclude the direction of correlation between X and Y. Moreover, we also cannot determine whether a line or a curve is more appropriate in describing the relationship between X and Y.

4. From the scatter diagram shown below, we can say that

(A) The Y=X line cuts through the data points in half, with 50% of the data points on either side of the line
(B) The average of X variable is larger than the average of Y variable
(C) The average of Y variable is larger than the average of X variable



Answer: C

The intersection of average lines (dotted lines) of X and Y lie above the Y=X line. This shows that the average of Y variable is higher than the average of X. The Y=X line does not cut through the data points in half, it only cuts through the X-Y graph in half.

5. Which one of the following is a property of correlation coefficient r?

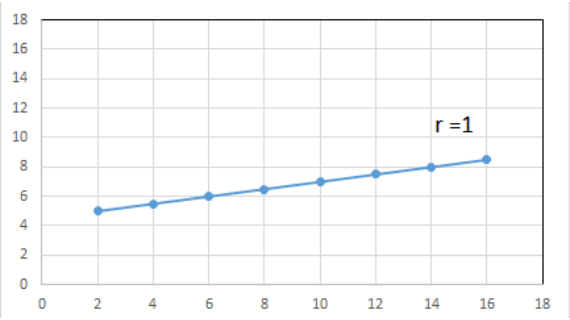
(A) r has units of measurements
(B) r indicates the direction of the linear relationship
(C) r is affected if we add 3 to all the values of one variable
(D) r is affected if we multiply all the values of one variable by 2

Answer: B

The positive or the negative sign of the correlation coefficient shows the direction of the linear relationship. The correlation coefficient is not affected if we add/multiply the same positive number to all the values of one variable. It is a pure number, without any units.

6. The figure below shows perfect positive correlation. If we model a linear regression, the slope of the regression line will be:

(A) 1 (Positive slope)
(B) -1 (Negative slope)
(C) 0
(D) None of the above



Answer: D

Correlation coefficient only measures the strength of the linear relationship. It is not the same as the gradient or slope of the line. We can have two lines of different slopes, but with the same correlation.