GER1000 AY2018/19 Semester 1
Quiz 9 and solutions

Q1

Some scientists have found that drinking coffee is associated with students' ability to sleep (enough vs not enough sleep). Sex was also found to be a confounder. This means that:

(I) Percentage of coffee drinkers among males is different from the percentage of coffee drinkers among females.
(II) Percentage of males among students who have enough sleep is different from the percentage of males among students who do not have enough sleep.

(a) (I) only
(b) (II) only
(c) Both (I) and (II)
(d) Neither (I) nor (II)

Answer: (c)

To be a confounder, sex must be associated with both drinking coffee and sleep.

(I) says rate (coffee drinker | male) ≠ rate (coffee drinker | female), which expresses an association between sex and drinking coffee.

(II) says rate (males | students with enough sleep) ≠ rate (males | students without enough sleep), which expresses an association between sex and sleep.

Q2

Suppose you are interested in a population of 100 people. Which of the following methods could be used to select a simple random sample of 10 people from this population?

(i) Assign without replacement to each person a random number from 1 to 100 such that none of them share the same number. Choose the people assigned numbers 1 to 10.

(ii) Write the names on equal-sized pieces of paper and put the papers in a hat. Shake the hat, mix the papers well, and draw out 10 names.

(iii) Sort the names based on alphabetical order and choose the first 10.

(a) i and ii only

(b) i and iii only

(c) ii and iii only

(d) i, ii and iii

Answer: (a)

In simple random sampling, every unit in the population of interest shall have the same chance of being selected in the sample. The procedure described in (iii) does not satisfy this requirement.

Q3

To join a certain tennis club, an applicant pays $100 and goes through a selection test by playing a game against with club member A and a game against club member B. To be accepted by the club, the applicant must win at least one of the two games, in which case, the applicant gets 50% refund of the fee. Otherwise, there is a 25% refund. Based on past records, A has a 60% chance of winning, while B has a 20% chance of winning. Assume the results of the two games are independent. How much will the tennis club make from each applicant on average? (Choose the closest option.)

(a) 40

(b) 48

(c) 53

(d) 58

(e) 72

Answer: (c)

P(applicant loses two games) = P(A win)*P(B win) = 0.6 * 0.2 = 0.12.

P(applicant wins at least one game) = 1 – 0.12 = 0.88.

Average value = $50*0.88 + $75*0.12 = $53.

Q4

A school consists of two faculties: arts and science, and enrolls both local and international students. The table shows some information about the rates of males. For example, among the local students, 34% are males, among the local students in arts faculty, 26% are males, etc.

Which of the following statement(s) is(are) correct?

	Local	International
Rate(males)	34%	68%

	Local	International
Rate(males|arts faculty)	26%	20%
Rate(males|science faculty)	76%	70%

(i) Among local students, the number of science students is the same as the number of arts students.
(ii) Within each faculty, males are more common among the local students than among the international students. But in the whole school, males are less common among the local students than the international students. There must be something wrong with the calculations.

(a) Only (i)
(b) Only (ii)
(c) Both (i) and (ii)
(d) None

Answer: (d)

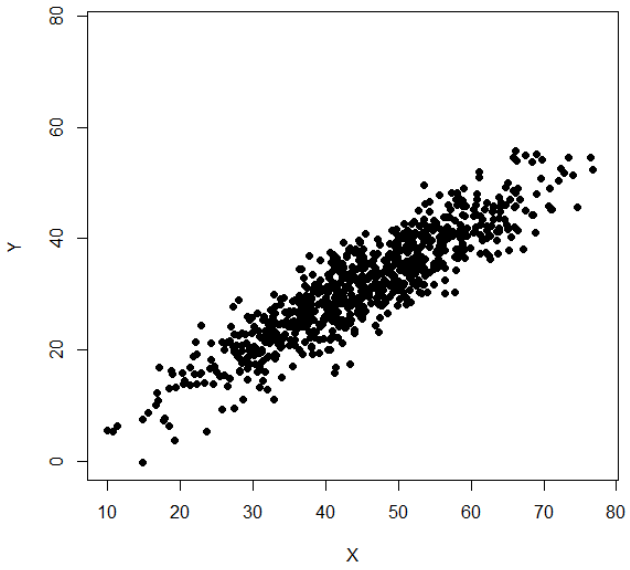
Firstly, suppose that among local students, there is an equal number of arts and science students. For example, let there be 100 arts and 100 science students. Then there would be 26 local male arts students and 76 local male science students. That is, there will be altogether 102 local male students. But 102/200 x 100% = 51%, not 34%. Thus (i) is false.

Secondly, Simpson's Paradox could have occurred because the overall rate of males among local students is lower than that of international students, but the association is reversed after slicing the students into arts and science faculties. That is, there is nothing wrong with the calculations. Hence (ii) is false.

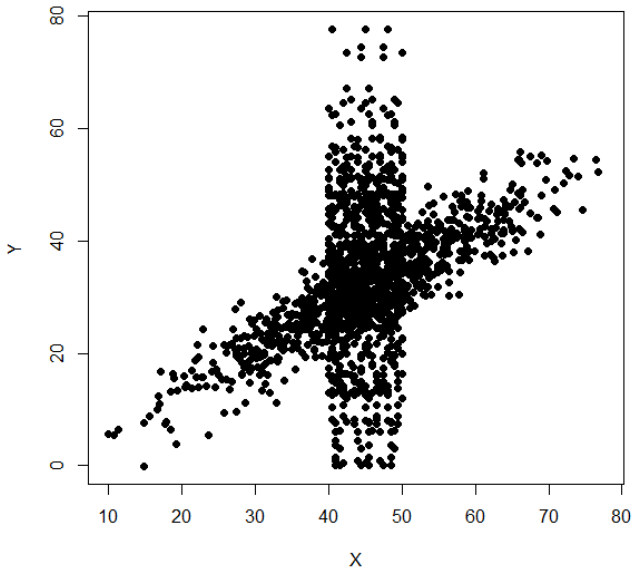
Q5

If X is restricted to the range 40 to 50, in which of the following scatter plots will the standard deviation of Y become less than before the restriction?

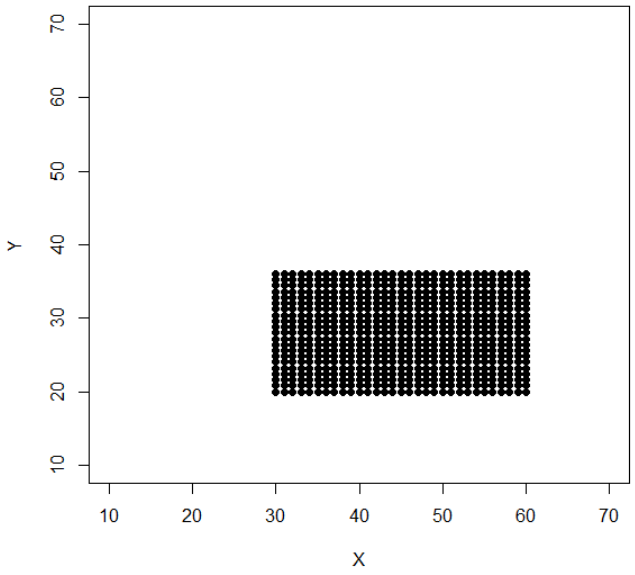
(a)



(b)



(c)



Answer: (a)

(a) The standard deviation in Y decreases after restriction. Please refer to Tutorial 2 Question 1(c).
(b) Data points outside the restricted range are closer to the average level of Y. The Y values are a lot more spread out in the restricted range.
(c) The standard deviation in Y is constant within any range on the x-axis.

Q6

A coin is tossed 8 times and all the outcomes are heads. Which of the following statement(s) is(are) correct?

i. If the null hypothesis is P(H)=0.6, we should reject it at the significance level of 5%.
ii. If the coin is tossed one more time, the outcome must be tail.

(a) Only (i)
(b) Only (ii)
(c) (i) and (ii)
(d) None

ANSWER: (a)

i. True. If P(H)=0.6, p-value = P(HHHHHHHH) = 0.6^8 = 0.017 < 0.05. Therefore, the null hypothesis is rejected.
ii. False. We are not sure whether the outcome of tossing the coin one more time is tail, unless we know that P(T) = 1.