GER1000 QUANTITATIVE REASONING
(Special Term : AY2015/2016)
TEST
4 June 2016
10:30—11:30 am

Instructions:

(i) This test paper contains 7 pages (including this page) and 14 questions. Answer <u>all</u> questions. Give your answers in the bubble answer sheet provided. Follow the instructions on the sheet. Fill in the bubbles neatly and completely, using a 2B pencil.

(ii) <u>Only</u> the bubble answer sheet will be marked. Make sure you have <u>written and bubbled</u> in your correct matriculation number (the number starting with "A") in the answer sheet.

(iii) Only one answer per question is allowed. Each correct answer is given one (1) mark. An incorrect/incomplete/missing answer receives 0 mark. The maximum total mark for this test is 14.

(iv) This is a closed book exam. Calculators are allowed. Computers, tablets and mobile devices are not allowed.

1. DES was given to pregnant women to prevent miscarriage. A literature review found 3 randomised controlled experiments and 5 nonrandomised studies with control groups. The rate of miscarriages was about the same in all the 8 treatment groups and in the control groups in the 3 randomised controlled experiments. However, the rate was substantially higher among the control groups in the 5 nonrandomised studies.

    A. DES is effective.
    B. DES is not effective.
    C. The result is inconclusive.

2. In a US General Social Survey, 2,726 adults were asked about education level and feelings about the future. The responses are as follows:

	Optimistic	Pessimistic	Row sum
Without high school diploma	748	1336	2084
With high school diploma	138	504	642
Column sum	886	1840	2726

    Which of the following justifies the statement "Among the survey participants, not having high school diploma was associated with optimism about the future."?

    (I) 1336/2084 > 748/2084
    (II) 748/2084 > 138/642
    (III) 748/886 > 1336/1840

    A. Only (I).
    B. Only (II).
    C. Only (III).
    D. Only (II) and (III).
    E. All of them.

3. Since the 1950's, many observational studies found consistent association between high-density lipoprotein-cholesterol (HDL-C) and heart attacks: people with higher levels of HDL-C had lower rates of heart attacks, even after controlling for many confounders. A double-blind randomised controlled experiment involving 12,000 subjects was done from 2012 to 2016 to study a drug that increases HDL-C level. Subjects in the treatment group had higher HDL-C compared to the control group, but the rate of heart attacks were the same in both groups.

    (I)   The observational studies suggested that increasing HDL-C prevents heart attacks.
    (II) The randomised experiment proved that increasing HDL-C prevents heart attacks.

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

4. The Salk vaccine trial in Chapter 1 consisted of two parts: the NFIP study and a randomised experiment. The second part is "randomised" because

(I) eligible subjects were randomly assigned to treatment and control.
(II) eligible subjects were randomly chosen from the population.

A. Both (I) and (II).
B. Only (I).
C. Only (II).
D. Neither (I) nor (II).

All 674 employees of a company were asked if they found work stressful. Employees were also classified as old (age 50 years or more) or young (less than 50 years old). The data are summarised in the table. The next two problems are based on this situation.

Sex	Age group	Stressful	Not stressful	Row sum
Female	Old	53	414	467
	Young	11	37	48
Male	Old	0	16	16
	Young	4	139	143
Total	Old	53	430	483
	Young	15	176	191

5. For the question of whether age has an effect on work stress, sex is a confounder. This conclusion is justified by

(I) the overall rate of stress among old workers was higher than the overall rate of stress among young workers, i.e., 53/483 > 15/191.
(II) the rate of stress among females, 64/515, is different from the rate of stress among males, 4/159.

A. Both (I) and (II).
B. Only (I).
C. Only (II).
D. Neither (I) nor (II).

6. The statement "The rate of stress among old workers is lower than the rate of stress among young workers." is true among

(I) female workers.   (II) male workers.   (III) all workers.

A. (I) only.
B. (II) only.
C. (III) only.
D. (I) and (II) only.
E. (I), (II) and (III).

7. In 2015, 57% of US companies offered health insurance to employees. Among companies that had 100 or more employees, 97% offered health insurance. Among companies employing less than 100 people, the percentage that offered health insurance

(I) can be calculated from the information given.
(II) must be less than 57%.

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

8. In the Pearson's father-son data set (shown below), a researcher computes the average height of fathers who are between 65 inches (inclusive) and 66 inches (exclusive), and also the average height of their sons. This yields a single point plotted on a new graph. He proceeds to do the same for other intervals: 66-67, 67-68, etc., as well as 64-65, 63-64, etc., obtaining a new scatter diagram with about 25 points. The correlation of the new scatter diagram is ___ the correlation of the original data set.

Pearson Data

Figure (1)
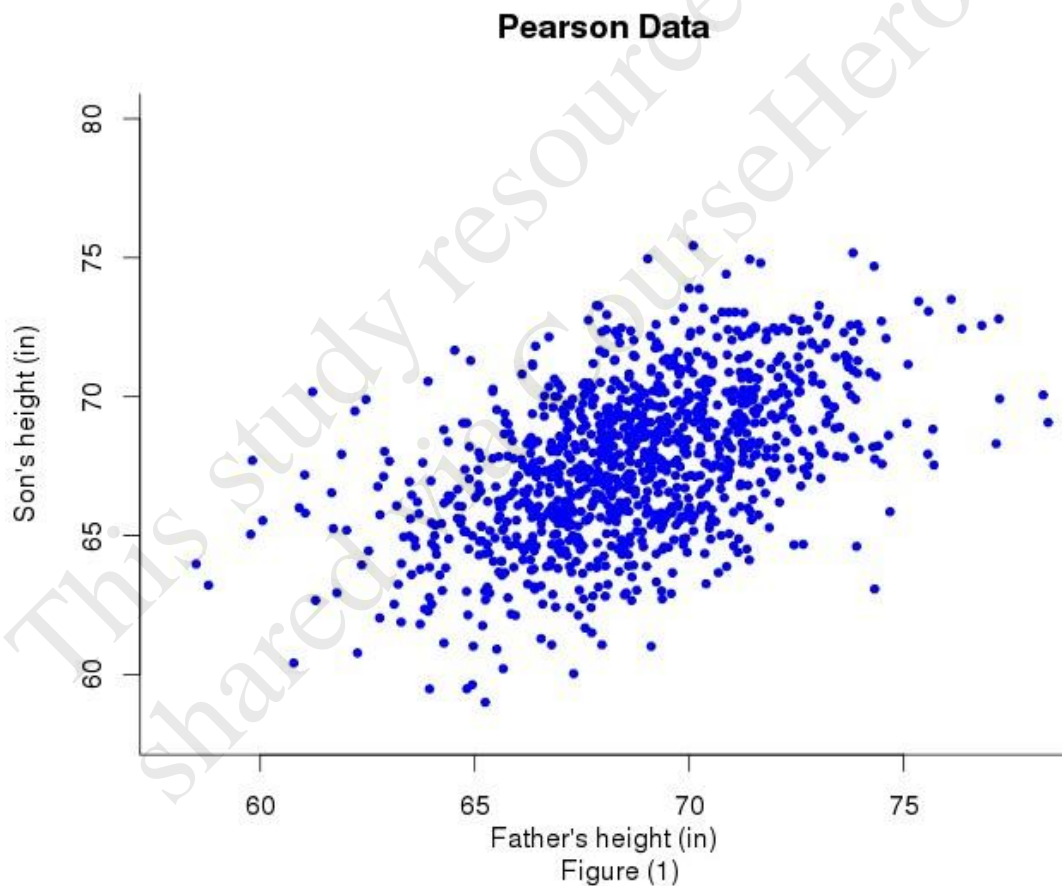
    A. more than
    B. similar to
    C. less than

9. Students of a university fill out questionnaires giving their year of birth, age (in years), age of father, and so forth. The correlation between student's age and year of birth is closest to

    A.   1
    B.   0.5
    C.   0
    D.   -0.5
    E.   -1

10. From the men in a large country, 13,191 were randomly selected. From the women in the same country, 11,482 were randomly selected. All 24,673 subjects were examined for diabetes. Data are summarised in the table.

 | | Diabetes | No diabetes | Row sum |
    |------------|----------|-------------|---------|
    | Men | 1251 | 11940 | 13191 |
    | Women | 512 | 10970 | 11482 |
    | Column sum | 1763 | 22910 | 24673 |

    (I)   The risk ratio for diabetes of men to women in the country is roughly 2.13.
    (II) The odds ratio for diabetes of men to women in the country is roughly 2.24.

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

11. As part of their training, air force pilots made two practice landings with their instructors and were rated on their performance. The instructors discussed the performance and ratings with the pilots after each landing. An analysis showed that pilots who made poor landings the first time tended to do better the second time. Conversely, pilots who had good landings the first time tended to do worse the second time. Choose the most appropriate option among the following:

    A.   This shows that criticism helps the pilots improve their landings, while praise makes them do worse.

    B.   The instructors now decided to criticize all first landings, regardless of actual performance. This decision will lead to better ratings on the pilots' second landing.

    C.   This can be explained by the effect of regression towards mediocrity.

    D.   These results are exceptional. Normally, we expect pilots with good ratings on the first landing to do even better the second time, and pilots with poor ratings on their first landing will do worse the second time.

12. The following plot shows an outlier in both the X and Y directions. What will happen if we remove the outlier?
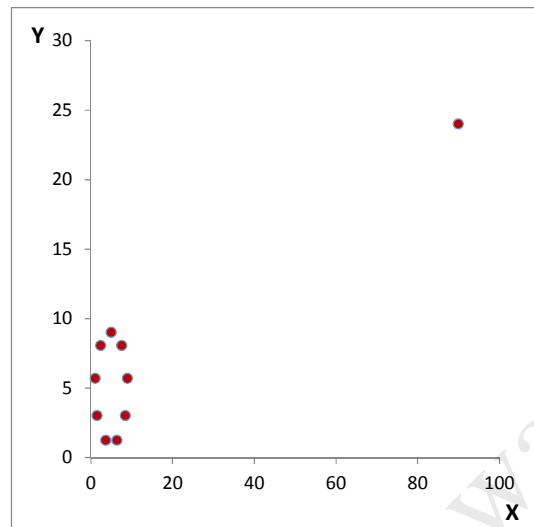
A. The correlation coefficient between X and Y will remain roughly the same.

B. The correlation coefficient between X and Y will decrease.

C. The correlation coefficient between X and Y will increase.

D. The correlation coefficient may increase or decrease, depending on the scales of measurement for X and Y.

E. It is not possible to know what will happen to the correlation coefficient between X and Y.

13. The correlation between height and weight among men age 18-74 is about 0.40.

(I) Taller men tend to be heavier.
(II) If someone eats more and puts on 10 kg, he is likely to get somewhat taller.

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

14. The following table lists the number of pages, price and the type of 15 books. H means hardcover while S means softcover. The scatter diagram is shown below, together with the regression line of price against page.

Page	104	188	220	264	336	342	378	385	417	417	436	458	466	469	585
Price	32.95	24.95	49.95	79.95	4.5	49.95	4.95	5.99	4.95	39.75	5.95	60	49.95	5.99	5.95
Type	H	H	H	H	S	H	S	S	S	H	S	H	H	S	S

Pages Vs Price
y = -0.06x + 51.19
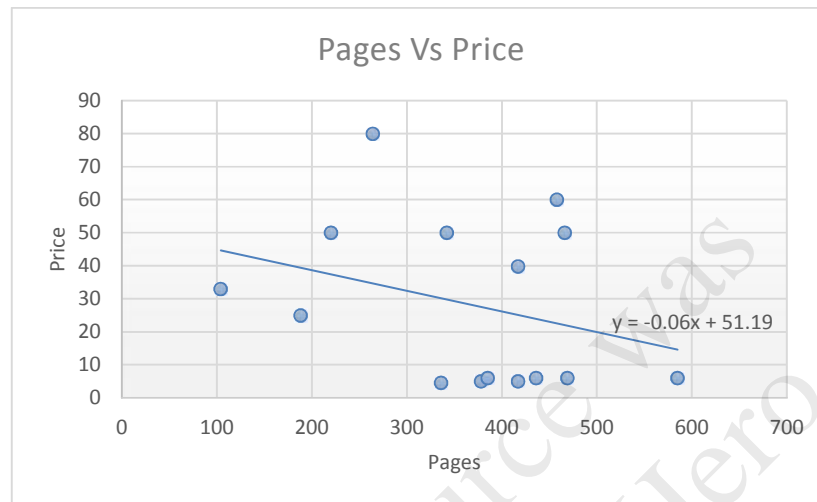
(I) For each increase of one page there is an average increase in the price of the book by $0.06.
(II) The regression is used to predict the price of a book of 500 pages. If the book is actually hardcover, the prediction is likely too low.

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

END OF PAPER

Risk and Odds

Suppose in a population of size n, s people have a disease. The risk of the disease is r = s/n, and the odds of the disease is r/(1-r) which is the same as s/(n-s).

Solutions:

1B. A controlled experiment with randomisation is more reliable than without. In the three randomised experiments, there was no difference in miscarriage rates between control and treatment groups. [Chapter 1 Units 4, 5]

2D. (I) states rate(pessimistic | no diploma) > rate(optimistic | no diploma). This is true, but does not show association. (II): rate(optimistic | no diploma) > rate(optimistic | diploma), and (III): rate(no diploma | optimistic) > rate(no diploma | pessimistic) both show association. [Chapter 1 Unit 6]

3B. (I) is true: this is like how smoking was suspected to cause ill health. (II) is false since the drug failed to reduce heart attacks in the treatment group.

4B. Eligible subjects were not randomly chosen, since parental consent was needed. [Chapter 1 Unit 4]

5D. (I) says age and stress are associated, but nothing about sex. (II) says sex and stress are associated, but this is incomplete: we also need an association between sex and age. [Chapter 1 Unit 6]

6D. (I) is true, since 53/467 < 11/48. (II) is true, since 0/16 < 4/143. But (III) is false: 53/483 > 15/191. This illustrates Simpson's paradox. [Chapter 1 Unit 9]

7C. It is impossible to know the percentage, but since the percentage in the big companies is larger than 57%, the percentage in the small companies must be less than 57% in order to have an overall percentage of 57%. This is like question 4 in Quiz 1.

8A. This is an ecological correlation. [Chapter 2 Unit 8]

9E. The correlation must be negative: the older students were born earlier. The vast majority of students lie on a straight line, so D can be ruled out.

10A. This is a cohort study, so both RR and OR can be estimated from the data. RR ~ (1251/13191)/(512/11482) = 2.13. The odds for men and women are respectively 1251/11940 and 512/10970, so the ratio is about 2.24. Or we can use the "cross-calculation" to get the answer directly. [Chapter 1 Unit 10, Chapter 2, Unit 3]

11C. This is like taller fathers (good landing on first try) tend to have relatively short sons (poorer second landing). [Chapter 2 Unit 9]

12B. Without the outlier, correlation is almost 0: as we go along the x axis, there is little change in the average y value.

13B. A positive correlation implies (I) in general. But the causal interpretation in (II) doesn't follow. "Association does not mean causation." [Chapter 2 Unit 3]

14C. (I) is false, since the slope should be negative. (II) is true since the hardcover books are more expensive than softcover books, and the regression line tries to fit in between.