1. To investigate whether controlling mothers tend to have obese children, an observational study was conducted and it was found that controlling behaviour in mothers is positively associated with obesity in their children: rate(obese child | controlling mom) = 71% while rate(obese child | non-controlling mom) = 39%. It was suspected that the health status of the father plays a role in the phenomenon, so the sample was 'sliced' according to whether the fathers were obese or not:

	Non-controlling Mother			Controlling Mother		
	Number	Obese	Rate (%)	Number	Obese	Rate (%)
Obese father	87	78	90%	270	234	87%
Non-obese father	263	60	23%	80	16	20%
Overall	350	138	39%	350	250	71%

Which of the following statements is/are true:
(i) Obesity in the father is a confounder, this is an example of Simpson's paradox.
(ii) Obesity in children is positively associated with controlling mothers when the father is obese.

(a) (i) only.
(b) (ii) only.
(c) Both (i) and (ii).
(d) Neither (i) nor (ii).

2. There are 3 families X, Y, Z. The families have 2 children each. Family X has 1 boy and 1 girl. Family Y has 2 girls. Family Z has 2 boys. 1 child is randomly selected among the 6 children. If the selected child is a boy, the probability that he is from family X is _____.

(a) 0
(b) 1/6
(c) 1/3
(d) 1/2

1

Questions 3 and 4 are based on the following information about the success rates of students and professionals in turning up for their scheduled appointments. They receive only one of the prompts - either text or email – before their scheduled appointments. There was a success rate of 30% among students who were prompted through text messages, but only 10% when they were reminded through emails. Professionals on the other hand, stuck to the appointments 40% of the time when reminded through texts and 80% when through emails.

3. From the information given above, which of the following must be true?
   (i)   Overall success rate for those prompted through text message is between 30% and 40%
   (ii) Overall success rate for those prompted through text message is 35%

   (a) (i) only.
   (b) (ii) only.
   (c) Both (i) and (ii).
   (d) Neither (i) nor (ii).

4. Let X be the overall success rate for students and Y the overall success rate for professionals. Which of the following is true?

   (a) X < Y
   (b) X = Y
   (c) X > Y
   (d) The relationship between X and Y cannot be determined from the information above

5. In a certain year, it is known that the prevalence of diabetes among Singapore residents is 10%. And the prevalence of diabetes among old (age 60 and above) Singapore residents is 30%. It was suggested that sex is a possible confounder in the observed association between age and diabetes among Singapore residents. After further analysis, the researchers concluded that sex is not a confounder, and there is an association between sex and age. Which of the following statements is/are true?
   (i)   rate(Diabetes|Male)= rate(Diabetes|Female)
   (ii)   rate(Male|Diabetes)= rate(Female|Diabetes)
   (iii)   rate(Diabetes|Female)=10%

   (a) (i) only.
   (b) (ii) only.
   (c) (i) and (ii) only.
   (d) (i) and (iii) only.
   (e) (i), (ii) and (ii).

6. In a recent study of diabetes treatments, rabbits were used extensively. Researchers began by ordering and shipping 10 rabbits each from various countries. When a package of 10 live rabbits arrives at the laboratory, all of them were fed lots of sugar for a month. Then a fair coin is tossed and if it lands on heads, all rabbits of the package will enter group A where they were given cultivated insulin. Otherwise, they will enter group B where they were given a newly designed compound. What is the type of study?

   (a) An observational study.
   (b) A controlled experiment, with historical controls.
   (c) A controlled experiment, without randomisation.
   (d) A controlled experiment, with randomisation.

Questions 7 and 8 refer to the following scenario. NUS has opened two classes, A and B, for GER1000 and each class has 100 students. In class A, every student scores 1 point higher in the final than in the midterm. In class B, every student scores 1 point lower in the final than in the midterm. For the midterm test, the average score is 50 and SD is 20 for both classes A and B, respectively.

7. Which of the following statements is/are true?
   (i) The correlation coefficient between the final score and the midterm score in class A is 1.
   (ii) The correlation coefficient between the final score and the midterm score in class B is -1.

   (a) (i) only.
   (b) (ii) only.
   (c) Both (i) and (ii).
   (d) Neither (i) nor (ii).

8. Consider the three correlation coefficients between midterm and final in class A, class B, and class A and B combined together. Which of the following statements is/are true?
   (i) The correlation coefficient of class A and B combined is smaller than the correlation coefficient of class A.
   (ii) The correlation coefficient of class A and B combined is larger than the correlation coefficient of class B.

   (a) (i) only.
   (b) (ii) only.
   (c) Both (i) and (ii).
   (d) Neither (i) nor (ii).

9. Which one of the following is **false** about a non-zero correlation coefficient?
   (a) The correlation does not change when we add 5 to all the values of one variable.
   (b) The correlation is positive when the regression slope is positive.
   (c) The correlation does not change when we multiply all the values of one variable by 2.
   (d) A correlation of -0.3 is stronger than a correlation of -0.8.

10. The Registry of Marriages is interested to see the relationship between husband's and wife's age in City X. They randomly sampled 1000 pairs of husbands and wives from the population of City X and obtained data of their ages (in years). Looking through the data, they found that men always marry women who are younger than them. Consider the following statements:
    (I) The average age of the husbands is more than the average age of the wives.
    (II) The SD of husband's age is more than the SD of wife's age.
    Which of the above statements is/are true based on the information given?

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

11. A group of researchers would like to study the effect of electronic cigarettes (e-cigarettes) use on the risk of getting lung cancer. 500 subjects were randomly sampled from e-cigarettes users and 500 subjects were randomly sampled from non-users for monitoring over 5 years. After that, they obtained the number of subjects with and without lung cancer in the two groups. Is the risk ratio of lung cancer between e-cigarettes users and non-users estimated from this sample likely to be a good estimate of the population risk ratio?
    (a) Yes
    (b) No

12. A college has 60 male and 60 female students taking a quantitative reasoning module in a particular semester. The following table depicts the distribution of sex (male vs. female) and grades (S vs. U) among these students.

	S (Satisfactory)	U (Unsatisfactory)	Row Total
Male			60
Female			60
Column Total			120

Suppose that rate(S|Female) was 0.5, and the odds ratio for S grade between male students to female students was also 0.5. What is the rate(Male|S)?

(a) 0.2
(b) 0.3
(c) 0.4
(d) 0.5

Questions 13 to 16 are linked to one another and refer to the following scenario. Oak, an NUS student, is interested in studying the satisfaction level of exchange students in the school. Unfortunately, he does not have access to the class list but is allowed to survey all the seven residences on campus.

13. Assuming all exchange students have to board with campus residences, what can be said of this sampling frame?
    (i)   As the sampling frame is larger than the target population, the sample is biased.
    (ii) As the sampling frame is larger than the target population, it is costly to implement.

    (a) (i) only.
    (b) (ii) only.
    (c) Both (i) and (ii).
    (d) Neither (i) nor (ii).

14. Oak decides to randomly select one residence out of the 7 residences and conduct door to door interviews to all residents of the selected residence only. What sort of sampling scheme is this?

   (a) Volunteer sampling
   (b) Convenience sampling
   (c) Simple random sampling
   (d) Cluster Sampling
   (e) Quota sampling

15. However, Oak injures his leg before he can implement his study, preventing him from conducting door to door interviews. He decides instead to send the survey to the whole NUS student population via email. 5% of exchange students eventually responded. To whom can the results of his survey be applicable?

   (a) All exchange students.
   (b) All NUS students.
   (c) All respondents of the survey.

16. Oak is acquainted with the "Exchange Students in NUS - AY2019/20" Facebook group and decides to also disseminate his survey via this channel. Out of 1378 members in the group, 1184 are exchange students while the rest are local NUS students. 496 members responded to the survey, of whom responses from 27 local students were excluded. Which of the following statements is/are true?
   (i)   It is known that about 58% of exchange students joined the Facebook group, hence the final sample is likely to be biased.
   (ii)   The Facebook group is infiltrated by local students who will eventually be excluded, so the final sample is likely to be biased.

   (a) (i) only.
   (b) (ii) only.
   (c) Both (i) and (ii).
   (d) Neither (i) nor (ii).

17. Which of the following statements is/are true?

(i) A simple random sample of size 500 is taken 100 times from the same population, and the 95% confidence interval of the population parameter is created for each of the 100 samples. About 95% of the intervals should contain the population parameter.

(ii) A decrease in sample size is likely to lead to a decrease in the 95% confidence interval's range.

(a) Neither (i) nor (ii)
(b) (i) only
(c) (ii) only
(d) Both (i) and (ii)

18. Researchers A and B are attempting to study the average height of people in country X. Researcher A conducted a simple random sample of size 1000, and produced a 95% confidence interval. Researcher B conducted a simple random sample of size 200, and produced a 95% confidence interval. Researcher A's confidence interval has a greater confidence of containing the population parameter, compared to researcher B's confidence interval.

(a) True
(b) False

19. Two fair six-sided dice are rolled once each. The dice are labelled Die-1 and Die-2. The faces of each die are labelled 1 to 6. The outcome of each die is the face it lands on. Assume the outcomes of the two dice are independent. P(sum of the outcomes of the two dice = 2) is equal to:

(a) P(Die-1 lands on 1) * P(Die-2 lands on 1)
(b) P(Die-1 lands on 1) + P(Die-2 lands on 1)
(c) P(Die-2 lands on 1)
(d) None of the above

20. In 2009, some UK researchers randomly selected some schools in a city and banned their students from using mobile phones. Then the average performance for each selected school was compared with the average performance of all the other schools in the same city. The researchers claimed that the longer a student stays away from the mobile phone, the better his/her academic performance. Which of the following fallacies could these researchers have possibly committed?

(a) Atomistic fallacy
(b) Ecological fallacy

21. Jane suspects that sleep duration is associated with a child's mathematical ability. She conducts a study to test her hypothesis. What should be her null hypothesis?

(a) There is insufficient evidence to conclude sleep duration is associated with a child's mathematical ability.
(b) Sleep duration is not associated with a child's mathematical ability.
(c) Sleep duration is associated with a child's mathematical ability.
(d) There is sufficient evidence to conclude sleep duration is associated with a child's mathematical ability.

22. Choose the best statement regarding a hypothesis test.
(a) The alternative hypothesis is always assumed to be true.
(b) The null hypothesis is what we are trying to prove.
(c) The alternative hypothesis is accepted unless there is sufficient evidence to say otherwise.
(d) The null hypothesis is not rejected unless there is sufficient evidence to reject it.

23. Allen is recently interested in buying stock X. He observes that within one day the stock price increases by 10% with probability of 40% and decreases by 10% with probability of 60%. Today's stock price is $10. Assume that the stock prices fluctuate independently over days. Which of the following statements is/are correct?

  (i)   The expected stock price in two days is higher than today's stock price.

  (ii) The stock price in two days will be higher than today's stock price.

  (a) (i) only.
  (b) (ii) only.
  (c) Both (i) and (ii).
  (d) Neither (i) nor (ii).

24. Among all the people in country A, 40% have blue eyes and 38% are left-handed. Among the people who do not have blue eyes, 30% are left-handed. If a person was picked at random from the people who have blue eyes, the chance that he/she is right-handed is _____.

  (a) 50%
  (b) 60%
  (c) 62%
  (d) 70%

25. A coin manufacturer produced a biased coin that supposedly gives heads with probability of 0.4 and tails with probability of 0.6. Alan tosses the coin 7 times and obtains 6 heads and 1 tail. Denote the probability of getting heads in a toss as P(H). He suspects that P(H) is more than 0.4 and carries out a hypothesis test. Which of the following statements is true?

(A) Alan's null hypothesis should be P(H) = 0.4 and he can reject the null hypothesis at 5% level of statistical significance based on the calculated p value.

(B) Alan's null hypothesis should be P(H) > 0.4 and he can reject the null hypothesis at 5% level of statistical significance based on the calculated p value.

(C) Alan's null hypothesis should be P(H) = 0.4 and he cannot reject the null hypothesis at 5% level of statistical significance based on the calculated p value.

(D) Alan's null hypothesis should be P(H) > 0.4 and he cannot reject the null hypothesis at 5% level of statistical significance based on the calculated p value.

Answer Key

1. A
2. C
3. A
4. A
5. D
6. D
7. A
8. A
9. D
10. B
11. A
12. C
13. B
14. D
15. C
16. A
17. B
18. B
19. A
20. B
21. B
22. D
23. D
24. A
25. A