1. Among 1000 university students who consent to participate in an experiment, 700 are undergraduates and 300 are postgraduate students. The students are randomly assigned into control group of size 700 and treatment group of size 300. The number of undergraduates in the treatment group is likely to be _____ the number of postgraduate students in the control group.

(A) less than

(B) equal to

(C) more than

No. undergraduates in treatment group = 300 x 0.7 = 210
No. postgraduates in control group = 700 x 0.3 = 210

2. A researcher reported that men have 36 times the odds of drinking wine as compared to women. Which of the following statements must also be true?

(A) Men have a higher risk of drinking wine as compared to women
(B) Men have a lower risk of drinking wine as compared to women
(C) Men have an equal risk of drinking wine as compared to women
(D) The relationship between risk of drinking wine among men and risk of drinking wine among women cannot be determined from the information

3. A researcher aims to examine the rate of Generalised Anxiety Disorder among all undergraduate students in Singapore. In order to do so, he obtains a list of all undergraduate students in the National University of Singapore, and randomly selects 1000 students for his sample. Given that the National University of Singapore is one of various universities in Singapore, which of the following statements is most likely to be true?

(A) The researcher has committed the atomistic fallacy
(B) The researcher has committed the ecological fallacy
(C) The sample is representative of the target population
(D) The sample is not representative of the target population

4. In order to determine public attitudes toward various social issues, a newspaper publication firm asks its readers to visit its website to participate in a poll. The poll addresses a different social issue each week. During the previous week, the following poll was found on the website:

Foreign companies should be allowed to participate in domestic political issues.
1 : Strongly disagree
2 : Disagree
3 : Neutral
4 : Agree
5 : Strongly agree

The following week, the newspaper publication firm announced that 6 in 10 respondents either disagreed or strongly disagreed with the statement. The population to which the results of this poll can be generalised to is:

A. All readers of the newspaper
B. All readers of the newspaper who have visited the website
C. All readers of the newspaper who have participated in this particular poll
D. All readers of the newspaper who have participated in at least one of the polls

5. In order to obtain a representative sample of all students in a particular university, a researcher positioned 10 research assistants at the entrance of each of the 10 faculties (the university has a total of 10 faculties), and asked them to collect a sample consisting of 1000 opinions by interviewing the first student who entered at the start of each 5 minute mark (length of each interview is less than 3 minutes).

After having collected the data, the researcher found that the demographics of the sample matched that of the target population on the following characteristics: age, gender, ethnicity, and major. All of these characteristics were argued to be important in influencing the results of the study. Which of the following statement is true?

(A) The sample is representative because of its demographic characteristics
(B) The sample is representative because of the sampling scheme used to obtain it
(C) The sample is not representative because of its demographic characteristics
(D) The sample is not representative because of the sampling scheme used to obtain it

6. In a study in which Singapore house addresses were used as the sampling frame, which of the following situations would prevent the researcher from generalising the results of his/her study to the population of Singapore?
(i) The proportion of addresses with no one living inside of them is 10%.
(ii) The proportion of addresses with no one willing to participate is 80%.
(iii) The sample size is 500.

(A) (i) only
(B) (ii) only
(C) (iii) only
(D) (ii) and (iii) only

7. In 2016, the Singapore population comprised 74% of Chinese ethnicity, 13% of Malay ethnicity, 9% of Indian ethnicity, and 3% of other ethnicities. A researcher stratified the population into these four ethnic groups, and obtained a simple random sample of 500 respondents within each group. Which of the following statements is true?

(A) The sample is representative of the Singapore population because the chance a Chinese individual is selected is the same as that a Malay individual is selected
(B) The sample is representative of the Singapore population because the chance a Chinese individual is selected is greater than that a Malay individual is selected
(C) The sample is not representative of the Singapore population because the chance a Chinese individual is selected is smaller than that a Malay individual is selected
(D) None of the above

8. Which of the following is/are true about p-value?
(i) P-value is a conditional probability computed based on the assumption that null hypothesis is true.
(ii) P-value gives the probability that null hypothesis is true.
(iii) P-value is dependent on the sample size of the hypothesis test.
(iv) A small p-value provides evidence that the null hypothesis is not true.
(v) A large p-value provides evidence that the null hypothesis is true.

(A) (i), (ii), (iii), (iv) and (v) only
(B) (i), (ii), (iv) and (v) only
(C) (i), (iii), (iv) and (v) only
(D) (i), (iii) and (iv) only

9. A multiple choice exam has 60 questions. Each question has 4 possible answers and only 1 answer out of the 4 possible answers is correct. To receive an A grade, one must answer 95% and above of the questions correctly. We know that 54 questions were answered correctly. What is the probability of receiving an A grade (rounding off to 3 decimal places), if one were to guess the remaining questions?

(A) 0.169
(B) 0.466
(C) 0.500
(D) 0.743

10. A biased coin has probability p of getting heads and suppose $0.4 < p < 0.5$. Let a, b and c be the probabilities of the following events:
a: Getting one head and one tail from two coin throws.
b: Getting two heads and one tail from three coin throws.
c: Getting two tails and one head from three coin throws.
Which of the following is true?

(A) $a < b < c$
(B) $b < c < a$
(C) $b < a < c$
(D) The relationship between a, b and c cannot be determined from the information given

11. In a certain day care class, 30% of the children have brown eyes, 20% of them have blue eyes and the other 50% have eyes that are in other colors. One day some of them play a game together. In the game, 45% of the children have brown eyes, 20% have blue eyes and 35% have other eye colors. Now, if a child is selected randomly from the class, and we know that he/she was not in the game, what is the probability that the child has blue eyes (rounding off to 2 decimal places)?

(A) 0.00
(B) 0.04
(C) 0.11
(D) 0.20

12. Assuming a telemarketer has a 20% chance of selling item A to each caller, a 40 % chance of selling item B to each caller, and the event of selling item A is independent to the event of selling item B. Each call in which the telemarketer sells both item A and item B takes 150 seconds, each call that sells only one item (item A or item B but not both) takes 60 seconds and each call that doesn't sell any items takes 30 seconds. If the telemarketer makes 45 calls, what is the average amount of time it takes?

(A) 53 seconds
(B) 2268 seconds
(C) 2376 seconds
(D) 2808 seconds

13. There is an epidemic. A person has probability 0.01 of having the disease. The authorities decide to test the population, but the test is not completely reliable. The sensitivity of the test is 0.98 and the specificity of the test is 0.95. Patrick was tested positive for the disease, what is the probability that Patrick has the disease (rounding off to 3 decimal places)?

(A) 0.010
(B) 0.020
(C) 0.165
(D) 0.198

14. Referring to Question 13, Patrick wants a second opinion: an independent repetition of the test (regardless of Patrick's disease status, outcomes of tests are independent). He went for a second test and was tested positive again. What is the probability that Patrick has the disease?

(A) 0.027
(B) 0.165
(C) 0.795
(D) 0.960

15. Eric, Freddie and Gavin are three students in a class of 55. A network is created with 55 vertices, and each vertex represents a student in this class. In this network, two vertices are adjacent if the corresponding students have each others phone numbers. Some of the centrality measures are shown below.

	Degree centrality measure	Closeness centrality measure
Eric	0.315	1.23
Freddie	0.537	1.71
Gavin	0.611	2.34

Let N be the number of vertices which are adjacent to both Freddie and Gavin. What is the minimum value that N can take?

(A) 2
(B) 5
(C) 7
(D) 9

16. Suppose you are going to open a new airline which operates a network of flights between 6 cities. In the network, a vertex represents a city, and two vertices are adjacent if there is a direct flight between the two cities.

You must ensure that any city can be reached from any other city either directly or via connecting cities. Also, due to the limited start-up fund, the size of the graph should be as small as possible. There are a few graphs satisfying the above requirements. Among all these graphs, what is the smallest (possible) closeness centrality measure of the cities?

(A) 1
(B) 1.2
(C) 1.4
(D) 1.6

17. A particular network has 4 vertices. There is one vertex u where there is exactly two vertices at distance 1 from u, and exactly one vertex at distance 2 from u. Moreover, all vertices have degree greater than 1. Which of the following statement is correct?

(A) The degree of u is 3
(B) It is not possible to determine the smallest closeness centrality measure based on the information given in the question
(C) It is not possible to determine the largest closeness centrality measure based on the information given in the question
(D) None of the above

18. A class has 45 students. A network is drawn with 45 vertices, each representing a student in the class. Two vertices are joined directly by an edge if the two corresponding students are friends. It is known that student A, student B and student C are friends with each other. Suppose there are two other students in class, student D and student E, who are not friends with each other but both are friends of student A. Which of the following is true about the betweenness centrality measure of student A?

(A) Bcen (student A) = 0
(B) 0 < Bcen (student A) < 1
(C) Bcen (student A) = 1
(D) None of the above

19. Another class has 43 students. The network of this class is drawn according to the criteria in Question 18. It is known that a student in this class, student H, has a unique social circle. For any two students in this class (excluding student H), say student S and student T, we have

d(student S, student T) = d(student S, student H) + d(student H, student T).

Which of the following is true about the betweenness centrality measure of student H?

(A) Bcen (student H) = 0
(B) 0 < Bcen (student H) < 1
(C) Bcen (student H) = 1
(D) None of the above

20. With reference to the discussion of the Bacon number of an actor, suppose the Bacon number of an actor A is 2. It is known that actor A, actor B and actor C were in the same movie "Frank". Moreover, actor A and actor C acted in only one movie. Which of the following statements must be correct?
(i) Bacon number of actor A is less than bacon number of actor B.
(ii) Degree of actor B is less than degree of actor C.
(iii) Closeness centrality measure of actor A is the same as the closeness centrality measure of actor B.
(iv) Betweenness centrality measure of actor A is the same as the betweenness centrality measure of actor C.

(A) (iv) only
(B) (i) and (ii) only
(C) (iii) and (iv) only
(D) None of the above

21. With reference to the discussion of the Bacon number of an actor, suppose the Bacon number of an actor C was initially computed to be 3. However, it was later discovered that there was one error in the database. The error was namely:
(i) There is a movie "Happiness", not involving Kevin Bacon, that was omitted from the database.
Suppose we know that everyone who acted in the movie "Happiness" has an initial Bacon number of at least 2. Which of the following statement must be correct after the error was corrected?

(A) The Bacon number of actor C is now 1
(B) The Bacon number of actor C is now 2
(C) The Bacon number of actor C is still 3
(D) It depends, the Bacon number of actor C now could be 2 or 3

22. In 2013, the sex ratio at birth of a certain country is 900. In 2014, the death rate of 0-year-old males is 10 per 1000 and the death rate of 0-year-old females is 7 per 1000. The 2015 sex ratio at 1 year old is (assuming there is no age 1 migration)

(A) less than 890
(B) between 890 to 895
(C) between 895 to 900
(D) more than 900
(E) We cannot determine based on the given information

23. In a peculiar town where everyone is above 19 years old and retires at 70 years old, the old age support ratio (OASR(70)) among males is two, and the old age support ratio (OASR(70)) among females is one. What are the odds x that an individual in this town is above the age of 69?

(A) $0.50 < x < 0.67$
(B) $0.33 < x < 0.50$
(C) $0.50 < x < 1.00$
(D) $0.67 < x < 0.75$

24. In 2015, OASR(75) $= \frac{1}{2}$ and OASR(76) $= \frac{2}{3}$. What is the rate (Age 75 population | Age $\geq$ 75 population) in 2015?

(A) $\frac{1}{3}$
(B) $\frac{1}{5}$
(C) $\frac{1}{10}$
(D) $\frac{2}{3}$

25. The age specific fertility rates (ASFR) for the year 2013 are given as follow:

Age group	15-19	20-24	25-29	30-34	35-39	40-44	45-49
ASFR	8.00	30.20	70.80	86.20	30.50	6.30	0.70

Suppose it is known that between 2013 and 2014, ASFR increases by 4% per year for the age group 20-24 and age group 30-34 while it increases by 2% per year for the age group 15-19 and age group 35-39. ASFR remains unchanged for all other age groups. Rounding off to 2 decimal places, the total fertility rate (TFR) for the year 2014 is

A. 1.15
B. 1.17
C. 1.18
D. 1.19

26. Assume that the Age X cohort of the following year arises from those who are Age (X-1) in the current year, minus deaths of Age (X-1), plus net migration of Age X. Given that:

- the Age (X-1) cohort of the current year is 740,

- the Age (X-1) deaths of the current year is 42,

- the Age X migration rate is 24 per 1000,

the Age X population (rounded off to whole number) for the following year is

(A) 715
(B) 716
(C) 722
(D) 727

27. Moving from the current year to the following year, the size of the Age $\geq$ 85 group in the following year is computed based on the data of the current year. In our model, we assume that the Age $\geq$ 85 group arises from the Age $\geq$ 84 group in the current year, minus deaths (of Age $\geq$ 84), plus migration (of Age $\geq$ 85). Given that:
- The Age 84 and Age $\geq$ 85 cohorts of the current year comprise 3000 and 10000 individuals respectively,
- The Age 84 and Age $\geq$ 85 death rates of the current year are 55 and 105 respectively, and
- The Age 84 and Age $\geq$ 85 migration rates of the current year are -2 and -5 respectively.
What is the Age $\geq$ 85 population of the following year?

(A) 11577
(B) 11726
(C) 11735
(D) 11844

28. Suppose Age X population in 2015 is smaller than the Age (X-1) population in 2014. We also know that Age X death rate in 2015 is smaller than Age (X-1) death rate in 2014 and Age (X+1) migration rate in 2015 is larger (in value) than Age X migration rate in 2014. Which of the following statements about Age (X+1) population in 2016 must be correct?
(i) It is larger than Age X population in 2015.
(ii) It is smaller than Age X population in 2015.
(iii) It is larger than Age (X-1) population in 2014.
(iv) It is smaller than Age (X-1) population in 2014.

(A) (i) only
(B) (i) and (iii) only
(C) (ii) and (iv) only
(D) None of the above