

UNIVERSITY OF JAFFNA
FACULTY OF ENGINEERING

END SEMESTER EXAMINATION- January 2019

PROBABILITY AND STATISTICS

MC 3020

Writing Time: TWO Hours

Permitted Materials: Calculators; Notes from the class

Instructions

1. This is an **OPEN book** exam.
2. This paper contains **TWO (2)** parts:
 - (a) **Part 1** contains **25** questions: Each question in this part is a **multiple choice** with five answer choices. Read each question and answer carefully and choose the ONE best answer. There will be **NEGATIVE MARKING** (for each correct answer you will get 2 marks and for each wrong answer you will loose 1 mark.) for the wrong answers.
 - (b) **Part 2** contains **04** questions.
3. Answer **all** questions in the answer book.
4. Read all the problems first before beginning to answer any of them. Start with the one you feel most comfortable with, and only move on to the next problem when you are certain you have completed it perfectly.
5. If you have any doubt as to the interpretation of the wording of a question, make your own decision, but clearly state in the script.
6. This examination accounts for **60%** of module assessment. Total maximum mark attainable is **100**.
7. Write your **registration number, the module code** and the **title of the paper** in the answer book. Also write your registration number on each additional sheet attached.

Part 1[50 marks]

Answer all questions in this part and you are advised to spend 60 minutes answering the questions in Part 1. You must mark your answers on the separate answer sheet.

1. We would like to test the hypothesis that $\mu = 20$ versus the alternative that $\mu \neq 20$. From a sample of 30 subjects we calculate the test statistic to be $t = 2.3$. The p-value would be
 - (a) 29
 - (b) 0.014
 - (c) 0.028
 - (d) 0.02
 - (e) None of the above
2. Suppose you conduct a significance test for the population proportion and your p-value is 0.184. Given a 0.10 level of significance, which of the following should be your conclusion?
 - (a) accept H_a
 - (b) do not reject H_0
 - (c) reject H_0
 - (d) do not reject H_a
 - (e) None of the above
3. A company wants to estimate the proportion of employees that are sick on a randomly selected day. What sample size is required if the company wishes to be at least 99% confident that their estimate will be correct to within 0.05
 - (a) 271
 - (b) 385
 - (c) 664
 - (d) 663
 - (e) None of the above
4. A student takes a standardized exam. The grader reports the students standardized score (z-score) as -1.8 . This indicates:
 - (a) The student scored lower than the average.
 - (b) The student scored less than one standard deviation from the average.
 - (c) A mistake has been made in calculating the score, since a standard score can never be negative.
 - (d) Both a and b, but not c.
 - (e) None of the above

5. Assume the cholesterol levels in a certain population have mean $\mu = 200$ and standard deviation $\sigma = 24$. The cholesterol levels for a random sample of $n = 9$ individuals are measured and the sample mean \bar{x} is determined. What is the z-score for a sample mean $\bar{x} = 180$?

- (a) -3.75 (d) 2.50
 (b) -2.50 (e) None of the above
 (c) -0.83

6. A study was conducted by the Department of Zoology at the Virginia Tech to estimate the difference in the amounts of the chemical orthophosphorus measured at two different station on the James River. Orthophosphorus was measured in milligrams per liter. Fifteen samples were collected from station 1, and 12 samples were obtained from station 2. The 15 samples from station 1 had an average orthophosphorus content of 3.84 milligrams per liter and a standard deviation of 3.07 milligrams per liter, while the 12 samples from station 2 had an average orthophosphorus content of 1.49 milligrams per liter and a standard deviation of 0.80 milligram per liter. Find a 95% confidence interval for the difference in the true average orthophosphorus contents at these two station.

- (a) $0.40 < \mu_1 - \mu_2 < 2.10$ (d) $0.06 < \mu_1 - \mu_2 < 2.10$
 (b) $0.04 < \mu_1 - \mu_2 < 4.10$ (e) None of the above
 (c) $0.60 < \mu_1 - \mu_2 < 4.10$

7. Let X be a random variable with density function

$$f(x) = \begin{cases} cx^2 & \text{if } -1 < x < 2 \\ 0 & \text{if elsewhere} \end{cases}$$

Find c and the expected value of $4X + 3$.

- (a) $0, 4$ (d) $\infty, 0$
 (b) $0.333, 8$ (e) None of the above
 (c) $3, 0.125$

8. You take a random sample from some population and form a 96% confidence interval for the population mean μ . Which quantity is guaranteed to be in the interval?

- (a) 0 (d) 0.04
 (b) μ (e) None of the above
 (c) \bar{x}

9. An electrical manufacturer produces light bulbs that have a life, before burn-out, that is normally distributed with mean equal to 800 hours and a standard deviation of 40 hours. Find the probability that a bulb burns between 778 and 834 hours.
- (a) 0.8023 (d) -0.5101
 (b) 1.0935 (e) None of the above
 (c) 0.5111
10. An industry report on cafes states that the strength of a cup of coffee (measured by the number of beans ground to make the cup of coffee) is normally distributed, with a mean of 20 beans per cup. You take a random sample of 81 cups, and observe a mean strength of 18.4 beans per cup with a sample standard deviation of 9 beans per cup. If you test the hypothesis that the average cup of coffee weaker than reported at the 5% level of significance, which of the following statements is true?
- (a) The test statistic will be -1.6, and we will not reject the null.
 (b) The test statistic will be -1.6, and we will reject the null.
 (c) The test statistic will be 1.6, and we will reject the null.
 (d) The test statistic will be 1.6, and we will not reject the null.
 (e) None of the above
11. In a certain assembly plant, three machines A , B and C , make 30%, 45% and 25%, respectively, of the products. It is known from past experience that 2%, 3% and 2% of the products made by each machine, respectively, are defective. Now, suppose that a finished product is randomly selected. What is the probability that it is defective?
- (a) 0.0135 (d) 0.006
 (b) 0.0011 (e) None of the above
 (c) 0.0245
12. Professors at a local university earn an average salary of 80,000 dollars with a standard deviation of 6,000 dollars. The salary distribution is approximately bell-shaped. What can be said about the percentage of salaries that are less than 68,000 dollars or, more than 92,000 dollars?
- (a) It is about 5% (d) It is about 95%
 (b) It is about 32% (e) None of the above
 (c) It is about 68%

13. The average growth of a certain variety of pine tree is 10.1 inches in three years. A biologist claims that a new variety will have a great three-year growth. A random sample of 25 of the new variety has an average three-year growth of 10.8 inches and a standard deviation of 2.1 inches. The appropriate null and alternate hypotheses to test the biologist's claim are:
- (a) $H_0 : \mu = 10.1$ against $H_a : \mu > 10.1$
 - (b) $H_0 : \mu = 10.8$ against $H_a : \mu \neq 10.8$
 - (c) $H_0 : \mu = 10.1$ against $H_a : \mu \neq 10.1$
 - (d) $H_0 : \mu = 10.8$ against $H_a : \mu > 10.8$
 - (e) None of the above
14. The following are percentages of fat found in 5 samples of each of two brands of baby food: $A : 5.7, 4.5, 6.2, 6.3, 7.3$ $B : 6.3, 5.7, 5.9, 6.4, 5.1$ Which of the following procedures is appropriate to test the hypothesis of equal average fat content in these two brands ?
- (a) Paired t-test with 5 degree of freedom
 - (b) Paired t-test with 4 degree of freedom
 - (c) Two sample t-test with 8 degree of freedom
 - (d) Two sample t-test with 9 degree of freedom
 - (e) None of the above
15. Which of the following p-values will lead us to reject the null hypothesis if the confidence level of the test is 95%?
- (a) 0.15
 - (b) 0.10
 - (c) 0.025
 - (d) 0.06
 - (e) None of the above
16. A 95% confidence interval for the mean number of televisions per American household is (1.15, 4.20). Which of the following statements about the above confidence interval is true?
- (a) The probability that μ is between 1.15 and 4.20 is 0.95.
 - (b) We are 95% confident that the true mean number of televisions per American household is between 1.15 and 4.20.
 - (c) 95% of all American households have between 1.15 and 4.20 televisions.
 - (d) The probability that \bar{x} is between 1.15 and 4.20 is 0.95.
 - (e) None of the above

17. For a random sample of 9 women, the average resting pulse rate is 76 beats per minute, and the sample standard deviation is 5. The standard error of the sample mean is
- (a) 0.745 (d) 2.778
 (b) 0.778
 (c) 1.667 (e) None of the above
18. Suppose a 95% confidence interval for the proportion of Americans who exercise regularly is 0.29 to 0.37. Which one of the following statements is **FALSE**?
- (a) It is reasonable to say that more than 25% of Americans exercise regularly.
 (b) It is reasonable to say that more than 40% of Americans exercise regularly.
 (c) The hypothesis that 33% of Americans exercise regularly cannot be rejected.
 (d) It is reasonable to say that fewer than 40% of Americans exercise regularly
 (e) None of the above
19. Suppose that we wanted to estimate the true average number of eggs a queen bee lays with 95% confidence. The margin of error we are willing to accept is 0.5. Suppose we also know that standard deviation is about 10. What sample size should we use?
- (a) 1536 (d) 2650
 (b) 1537
 (c) 2653 (e) None of the above
20. If the coefficient of determination is 0.81, then the correlation coefficient
- (a) is 0.6561 (d) must be negative
 (b) could be either +0.9 or -0.9
 (c) must be positive (e) None of the above
21. Given the following data pairs (x, y) , find the regression equation.
 (1, 1.24), (2, 5.23), (3, 7.24), (4, 7.60), (5, 9.97), (6, 14.31), (7, 13.99), (8, 14.88), (9, 18.04), (10, 20.70)
- (a) $y = 0.490x - 0.053$ (d) $y = 0.49x$
 (b) $y = 2.04x$
 (c) $y = 1.98x + 0.436$ (e) None of the above

22. A national random sample of 20 **ACT** scores from 2010 is listed below. Calculate the sample mean and standard deviation.
 29, 26, 13, 23, 23, 25, 17, 22, 17, 19, 12, 26, 30, 30, 18, 14, 12, 26, 17, 18
- (a) 20.50, 5.79 ~~(d) 20.85, 5.94~~
 (b) 20.50, 5.94
 (c) 20.85, 5.79 (e) None of the above
23. Provided that the **ACT** is reasonably normally distributed with a mean of 18 and standard deviation of 6, determine the proportion of students with a score of 33 or higher.
- ~~(a) 0.0062~~ (d) 0.0217
 (b) 0.0109
 (c) 0.0124 (e) None of the above
24. Using the data in question 22, calculate the 95% confidence interval for the mean **ACT** score based on the t-distribution.
- (a) $-\infty$ to 23.05 (d) 18.22 to 23.48
 (b) $-\infty$ to 23.15
~~(c) 18.07 to 23.63~~ (e) None of the above
25. A hypothesis test is done in which the alternative hypothesis is that more than 10% of a population is left-handed. The p-value for the test is calculated to be 0.25. Which statement is correct?
- (a) We can conclude that more than 10% of the population is left-handed.
 (b) We can conclude that more than 25% of the population is left-handed.
 (c) We can conclude that exactly 25% of the population is left-handed
~~(d) We cannot conclude that more than 10% of the population is left-handed~~
 (e) None of the above

Part 2[50 marks]

Answer all questions in this part and you are advised to spend an hour answering the questions in Part 2. You must write your answers on the answer book provided.

Question 1[10marks]

1. A factory produces components of which 1% are defective. The components are packed in boxes of 10. A box is selected at random.
 - (a) Find the probability that the box contains exactly 2 defective component.
 - (b) Using a suitable approximation, find the probability that a batch of 250 components contains between 1 and 4(inclusive) defective components.
2. A report from the Secretary of Health and Human Services stated that 75% of single vehicle traffic fatalities that occur at night on weekends involve an intoxicated driver. If a sample of 15 single -vehicle traffic fatalities that occur at night on a weekend is selected.
 - (a) Find the probability that at least 12 involve a driver who is intoxicated.
 - (b) What is the expected and variance number of driver who is intoxicated?

Question 2[07 marks]

1. Patients arrive at a hospital accident and emergency department at random at a rate of 6 per hour. A patient arrives at 11.30a.m. Find the probability that the next patient arrives before 11.45a.m.
2. Airline passengers arrive randomly and independently at the passenger-screening facility at a major international airport. The mean arrival rate is 420 passengers per hour.
 - (a) What is the probability that at least three passengers will arrive in a 30 seconds period?
 - (b) What is the expected number of passengers arrived in given day?

Question 3[18 marks]

1. Customers using a self-service soda dispenser take an average of 12 ounces of soda with a standard deviation of 4 ounces. Assume that the amount would be normally distributed.
 - (a) What is the probability that a randomly selected customer takes between 13 and 14 ounces of soda?

- (b) What is the probability that the next 100 customers will take an average amount less than 12.24 ounces?
2. The IQ scores of 1800 applicants for admission to a tuition free graduate school are normally distributed with mean of 130 and a standard deviation of 12.
- (a) If an applicant is chosen at random, what is the probability that the IQ score is less than 142?
- (b) What is the 36th percentile of these scores?
- (c) If the admission policy is to refuse entry to any applicant with an IQ below 100, how many applicants will be admitted?
3. The average weekly income of information workers in private industry is 777 dollars. If the standard deviation is 77 dollars, what is the probability that a random sample of 50 information workers will earn, on average, more than 800 dollars per week?

Question 4[15 marks]

1. In a psychological testing experiment, 25 subjects are selected randomly and their reaction time, in seconds, to a particular stimulus is measured. Past experience suggests that the variance in reaction times to these types of stimuli is 4sec^2 and that the distribution is approximately normal. The average time for the subjects is 6.2 seconds. Find a 95% confidence interval for the mean reaction time.
2. The contents of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 liters. Find a 95% confidence interval for the mean contents of all such containers.
3. A genetic theory says that a cross between two pink flowering plants will produce red flowering plants 25% of the time. To test the theory, 100 crosses are made and 31 of them produce a red flowering plant.
- (a) Write down the null and alternative hypotheses.
- (b) Find the test statistic and the p-value.
- (c) State your conclusions

————— *End of Examination* —————