



RAJARATA UNIVERSITY OF SRI LANKA
FACULTY OF APPLIED SCIENCES

BSc in Applied Sciences

First Year – Semester I Examination – May 2022

MAA 1302 – PROBABILITY AND STATISTICS I

Time: Three (03) hours

Answer all questions.

01. a) State whether each of the following statements is True/False. Justify your answer in one sentence. (2.5 * 2 = 5 marks)
- (i) The power of a study is the probability that the null hypothesis will not be correctly rejected when it is true. (True / False)
 - (ii) If the data used in a hypothesis test produce a p -value less than 0.001, the null hypothesis must be false. (True / False)
- b) In a chi-square distribution with 17 degrees of freedom, what value marks the upper 5% of the distribution? (5 marks)
- c) Suppose for two events, A and B , $P(A) = 0.5$, $P(B) = 0.2$ and A and B are disjoint. Given that, $P(A \cap B) = 0.01$. Justify the claim that A and B are disjoint. (5 marks)
- d) A 95% confidence interval for the difference between two means is reported as (0.8, 2.3). What can you conclude from the test results of the null hypothesis that the two-population means are equal? (5 marks)
02. All human blood can be “ABO-type”, i.e., can be classified as O, A, B or AB. However, the distribution of blood types varies across racial and ethnic backgrounds. The table below shows the probability of each blood type in Western and the Northern provinces.

	O	A	B	AB
Western	0.35	0.27	0.26	0.12
Northern	0.45	0.40	0.11	0.04

Suppose that a couple, made up of a female from the Northern province and a male from the Western province, decided to marry and that the individual blood type has not influenced their decision.

- a) What is the probability that they both have type AB blood? (5 marks)
- b) What is the probability that they both have the same blood type? (5 marks)

03. In June, the daily average high temperature in a particular city is 77°F with a standard deviation of 5°F . Suppose that the temperatures in June closely follow a normal distribution.
- What is the probability of observing an 83°F temperature or higher in the city on a randomly chosen day in June? (5 marks)
 - How cool are the coldest 10% of days (days with the lowest average high temperature) during June in the city? (5 marks)
04. A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked.
- Write the hypotheses in words? (5 marks)
 - What is the type I error in this context? (5 marks)
 - What is the type II error in this context? (5 marks)
 - Which error is more problematic for the restaurant owner? Why? (5 marks)
05. A simple random sample of 826 payday loan borrowers was surveyed to better understand their interests in regulation and costs. 70% of the responses supported new regulations on payday lenders.
- Is it reasonable to model $\hat{p} = 0.70$ using a normal distribution? (2.5 marks)
 - Estimate the standard error of $\hat{p} = 0.70$. (5 marks)
 - Construct a 95% confidence interval for p , the proportion of payday borrowers who support increased regulation for payday lenders. (5 marks)
 - Set up hypotheses to evaluate whether borrowers have majority support or majority opposition for this type of regulation. (2.5 marks)
06. A university newspaper is conducting a survey to determine what fraction of students support a \$200 per year increase in fees to pay for a new football stadium. How big of a sample is required to ensure that the margin of error is smaller than 0.04 using a 95% confidence level? (10 marks)
07. New York is known as “the city that never sleeps”. A random sample of 25 New Yorkers was asked how much sleep they get per night. Statistical summaries of these data are shown below. The point estimate suggests New Yorkers sleep less than 8 hours a night on average.

n	\bar{x}	s	\min	\max
25	7.73	0.77	6.17	9.78

Where \bar{x} – sample mean, s – sample variance, \min – minimum value, \max – maximum value

- a) Write the hypotheses in symbols and in words. (2 marks)
- b) Check conditions, then calculate the test statistic, T , and the associated degrees of freedom. (6 marks)
- c) Find and interpret the p -value in this context. (4 marks)
Hint: Drawing a picture may be helpful.
- d) What is the conclusion of the hypothesis test? (4 marks)
- e) If you were to construct a 90% confidence interval that corresponded to this hypothesis test, would you expect 8 hours to be in the interval? (4 marks)
08. The drug *diflunisal* is used to treat arthritis pain and was recently examined in a study of its effect on glaucoma. Ten patients were recruited for the study, and for each patient, the change in intraocular pressure (after-before drug administration) was measured in mmHg. For these 10 patients, the average change was -1.6 mmHg, with an estimated standard deviation of 1.5 mmHg.
- a) Using a two-sided hypothesis test, assess the evidence for the claim that this drug changes intraocular pressure in the eye. (15 marks)
Note: Be sure that all five steps in testing are clearly indicated in your solution.
Use type I error $\alpha = 0.05$
- b) Calculate a 95% confidence interval for the change in pressure. (5 marks)
09. The distribution of the number of eggs laid by a certain species of the hen during their breeding period has a mean of 35 eggs with a standard deviation of 18.2. Suppose a group of researchers randomly samples 45 hens of this species, counts the number of eggs laid during their breeding period, and records the sample mean. They repeat this experiment for 1,000 times and build a distribution of sample means.
- a) What is this distribution called? (4 marks)
- b) Would you expect the shape of this distribution to be symmetric, right skewed, or left skewed? Explain your reasoning. (6 marks)
- c) Calculate the variability of this distribution and state the appropriate term used to refer to this value. (5 marks)
- d) Suppose the researchers' budget is reduced and they can only collect random samples of 10 hens. The sample mean of the number of eggs is recorded, and we repeat this 1,000 times, and build a new distribution of sample means. How will the variability of this new distribution compare to the variability of the original distribution? (5 marks)

Providing an answer for the following problem is not mandatory. You could use this problem to substitute quiz marks:

Measurements of prostate specific antigen (PSA) are often used as the initial screening test for prostate cancer. It is an appealing test, because it can be measured using urine, so it is non-invasive and inexpensive. It is controversial, though, and the National Institute of Health and other organizations have issued conflicting guidelines for its use in the last 25 years.

Recent studies by the American Cancer Society have estimated that the sensitivity of the test is approximately 0.51 or 51%, and the specificity is 0.91, or 91%. The estimated prevalence of undiagnosed prostate cancer in men between 60 and 70 years of age is approximately 6,000 cases per 100,000 males in the population.

- a) Suppose a 65-year-old male has been told that his PSA test is positive for prostate cancer. Use Bayes' rule and the formulas used in class to calculate the probability that the man has prostate cancer. (5 marks)
- b) Construct a table of a hypothetical population of 100,000 males between 60 - 70 years of age and use the table to confirm your calculation in part (a). (5 marks)

Hint: Organize your table carefully and clearly, indicating the meaning of columns and rows. You will need this table to answer the remaining questions below.

- c) Use your table to calculate the number of men in the hypothetical population of 100,000 men who will have a positive PSA test. (5 marks)
- d) The PSA test is typically followed by a biopsy for men who are tested positive. For men with a positive PSA test, assume that the biopsy has a sensitivity and specificity of 0.99. If a man with a positive PSA test subsequently has a positive biopsy, what is the probability that he has prostate cancer? (5 marks)
- e) Many studies have shown that prostate cancer is almost twice as prevalent in black males as it is in the general population of males. Would you expect the positive predictive value of PSA screening to be larger or smaller than what you calculated in part (a)? Justify your answer either numerically or by referring to properties of the positive predictive value of a test. (5 marks)

*** END ***