UNIVERSITY OF ESWATINI

MAIN EXAMINATION, 2020/2021

BSc.IV, BASS IV

Title of Paper : Mathematical Statistics II

Course Number : MAT441

Time Allowed : Three (3) Hours

Instructions

1. This paper consists of SIX (6) questions in TWO sections.
2. Section A is **COMPULSORY** and is worth 40%. Answer ALL questions in this section.
3. Section B consists of FIVE questions, each worth 20%. Answer ANY THREE (3) questions in this section.
4. Show all your working.
5. Start each new major question (A1, B2, B3 ,B4, B5, B6) on a new page and clearly indicate the question number at the top of the page.
6. You can answer questions in any order.
7. Indicate your program next to your student ID.

Special Requirements: NONE

THIS EXAMINATION PAPER SHOULD NOT BE OPENED UNTIL PERMISSION HAS BEEN GIVEN BY THE INVIGILATOR.

SECTION A [*40 Marks*]: ANSWER ALL QUESTIONS

QUESTION A1 [*40 Marks*]

A1 (a) State the central limit theorem .

[5 Marks]

(b) $X_1, X_2, \ldots, X_n$ is a random sample from the uniform distribution between θ and 1 (i.e. $f(x) = (1-\theta)^{-1}$ for $\theta < x < 1$), where $\theta(< 1)$ is an unknown parameter. Denote the sample mean by $\overline{X}$. Show that the method of moments estimator, $\hat{\theta}$, of θ is

$$2\overline{X} - 1.$$

[5 Marks]

(c) Let $X_1, \ldots, X_n$, $n > 4$, be a random sample from a population with a mean μ and variance σ^2. Consider the following two estimators of μ:

$$\hat{\theta}_1 = \frac{1}{9}(X_1 + 2X_2 + 5X_3 + X_4),$$

$$\hat{\theta}_2 = \overline{X}.$$

Calculate the relative efficiency $e(\hat{\theta}_2, \hat{\theta}_1)$. Interpret.

[5 Marks]

(d) A large-sample α−level test of hypothesis for $H_0 : \theta = \theta_0$ versus $H_a : \theta > \theta_0$ rejects the null hypothesis if

$$\frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} > z_\alpha.$$

Show that this is equivalent to rejecting H_0 if θ_0 is less than the large-sample $100(1-\alpha)\%$ lower confidence bound for θ.

[5 Marks]

(e) Write down the four elements of a statistical test.

[4 Marks]

(f) Auditors are often required to compare the audited (or current) value of an inventory item with the book (or listed) value. If a company is keeping its inventory and books up to date, there should be a strong linear relationship between the audited and book values. A company sampled ten inventory items and obtained the audited and book values given in the accompanying table.

Item	Audit Value (y_i)	Book Value (x_i)
1	9	10
2	14	12
3	7	9
4	29	27
5	45	47
6	109	112
7	40	36
8	238	241
9	60	59
10	170	167

Fit the model $Y = \beta_0 + \beta_1 x + \varepsilon$ to these data.

[10 Marks]

(g) Let X be a binomial random variable with parameters n and p. Assume that the prior distribution of p is uniform on $[0, 1]$. Find the posterior distribution, $f(p|x)$.

[6 Marks]

SECTION B: ANSWER ANY *THREE* QUESTIONS

QUESTION B2 [*20 Marks*]

B2 (a) A study of parallel interchange ramps revealed that many drivers do not use the entire length of parallel lanes for acceleration, but seek, as soon as possible, a gap in the major stream of traffic to merge. At one site on the highway, 46% of drivers used less than one third of the lane length available before merging. Suppose we monitor the merging pattern of a random sample of 250 drivers at this site.

(i) What is the probability that fewer than 120 of the drivers will use less than one third of the acceleration lane length before merging?

[3 Marks]

(ii) What is the probability that more than 225 of the drivers will use less than one third of the acceleration lane length before merging?

[3 Marks]

(b) Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with variance $\sigma^2 < \infty$. If

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

is the variance of a random sample from an infinite population, show that S^2 is an unbiased estimator for σ^2.

[6 Marks]

(c) A biologist has hypothesized that high concentrations of actinomycin D inhibit RNA synthesis in cells and thereby inhibit the production of proteins. An experiment conducted to test this theory compared the RNA synthesis in cells treated with two concentrations of actinomycin D: 0.6 and 0.7 micrograms per liter. Cells treated with the lower concentration (0.6) of actinomycin D yielded that 55 out of 70 developed normally whereas only 23 out of 70 appeared to develop normally for the higher concentration (0.7). Do these data indicate that the rate of normal RNA synthesis is lower for cells exposed to the higher concentrations of actinomycin D? Find the $p-$value for the test.

[8 Marks]

QUESTION B3 [*20 Marks*]

B3 (a) Let $X_1, \ldots, X_n$ be a random sample from a population with pdf

$$f(x) = \begin{cases} (\theta+1)x^{\theta}, & \text{for } 0 < x < 1; \theta > -1, \\ 0, & \text{otherwise.} \end{cases}$$

Find the maximum likelihood estimator for θ

[5 Marks]

(b) Now suppose $f(x,\theta) = \frac{1}{\theta}e^{-x/\theta}, x > 0$, and $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

i. Show that $\overline{X}$ is an unbiased estimator of θ. [5 Marks]

ii. State the factorisation criterion for sufficient statistics and use it to show that $\overline{X}$ is sufficient for θ. [5 Marks]

iii. State the Cramer-Rao inequality for unbiased estimators of θ. Show that $\overline{X}$ attains the lower bound for the distribution above. Explain what that means regarding the efficiency of $\overline{X}$ as an estimator of θ. [5 Marks]

QUESTION B4 [*20 Marks*]

B4 (a) In the data set below, W denotes the weight (in pounds) and l the length (in inches) for 15 alligators captured in central Florida. Because l is easier to observe than W for alligators in their natural habitat, a researcher would like to construct a model relating weight to length. Such a model can then be used to predict the weights of alligators of specified lengths.

Alligator	length(l)	Weight (W)
1	47.94	130.32
2	36.97	50.91
3	75.94	639.06
4	30.88	27.94
5	45.15	79.84
6	46.06	109.95
7	31.82	33.12
8	42.94	90.12
9	33.12	35.87
10	35.87	38.09
11	66.02	365.04
12	43.82	83.93
13	40.85	79.84
14	41.68	83.10
15	43.82	70.12

Fit the model $E(W) = \alpha_0 l^{\alpha_1}$.

[10 Marks]

(b) Given the data

x	y
-2	0
-1	0
0	1
1	1
2	3

and

$$(XY)^{-1} = \begin{bmatrix} 17/35 & 0 & -1/7 \\ 0 & 1/10 & 0 \\ -1/7 & 0 & 1/14 \end{bmatrix},$$

Fit the model $Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$.

[10 Marks]

QUESTION B5 [*20 Marks*]

B5 (a) A psychological study was conducted to compare the reaction times of men and women to a stimulus. Independent random samples of 50 men and 50 women were employed in the experiment.The results are shown in Table below.

Men	Women
$n_1 = 50$	$n_2 = 50$
$\overline{y}_1 = 3.6$ seconds	$\overline{y}_2 = 3.8$ seconds
$s_1^2 = 0.18$	$s_2 = 0.14$

(i) Do the data present sufficient evidence to suggest a difference between true mean reaction times for men and women? Use $\alpha = 0.05$.

8 Marks]

(ii) Find the p-value for the statistical test.

[4 Marks]

(b) A random sample of size 36 from a population with known variance, $\sigma^2 = 9$, yields a sample mean of $\overline{x} = 17$. Compute the type II error β for testing the hypothesis $H_0 : \mu = 15$ versus $H_a : \mu = 16$. Assume $\alpha = 0.05$.

[8 Marks]

QUESTION B6 [*20 Marks*]

B6 (a) In Bayesian inference define a conjugate prior distribution.

[3 Marks]

(b) Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Bernoulli distribution where

$$P(Y_i = 1) = p \text{ and } P(Y_i = 0) = 1 - p,$$

and assume that the prior distribution for p is $beta(\alpha, \beta)$, i.e.

$$f(y) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}, & 0 < y < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(i) Find the posterior distribution for p.

[12 Marks]

(ii) Find the Bayes estimators for p.

[5 Marks]

END OF EXAMINATION PAPER