UNIVERSITY OF ESWATINI

MAIN EXAMINATION, 2019/2020

BSc.IV, BASS IV, BEd 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) Suppose that $X_1, X_2, \ldots X_n$ and $Y_1, Y_2, \ldots Y_n$ are independent random samples from populations with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2, respectively. Show that the random variable

$$U_n = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2 + \sigma_2^2)/n}}$$

satisfies the conditions of the Central Limit theorem and thus that the distribution function of U_n converges to a standard normal distribution function as $n \longrightarrow \infty$.

[5 Marks]

(b) Let $X_1, \ldots, X_n$, $n > 3$ be a random sample from a population with a true mean μ and variance σ^2. Consider the following three estimators of μ:

$$\hat{\theta}_1 = \frac{1}{3}(X_1 + X_2 + X_3),$$
$$\hat{\theta}_2 = \frac{1}{8}X_1 + \frac{3}{4(n-2)}(X_2 + \ldots + X_{n-1}) + \frac{1}{8}X_n,$$
$$\hat{\theta}_3 = \overline{X}.$$

Find the relative efficiency $e(\hat{\theta}_3, \hat{\theta}_1)$, and interpret.

[5 Marks]

(c) 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, $f(p|x)$.

[5 Marks]

(d) A large-sample $\alpha-$level test of hypothesis for $H_0 : \theta = \theta_0$ versus $H_a : \theta > 0$ 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 condence bound for θ.

[5 Marks]

(e) Let $X_1, \ldots, X_n$ be a random sample of size n from the exponential distribution whose pdf is

$$f(x, \beta) = \begin{cases} \beta e^{-\beta x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

(i) Use the method of moments to find a point estimator for β.

[5 Marks]

(ii) The following data represent the time intervals between the emissions of beta particles: Assuming the data follow an exponential distribution, use the estimator you found in e(i) above to compute a moment estimate for the parameter β.

[5 Marks]

0.9	0.1	0.1	0.8	0.9	0.1	0.1	0.7	1.0	0.2
0.1	0.1	0.1	2.3	0.8	0.3	0.2	0.1	1.0	0.9
0.1	0.5	0.4	0.6	0.2	0.4	0.2	0.1	0.8	0.2
0.5	3.0	1.0	0.5	0.2	2.0	1.7	0.1	0.3	0.1
0.4	0.5	0.8	0.1	0.1	1.7	0.1	0.2	0.3	0.1

(f) 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]

SECTION B: ANSWER ANY *THREE* QUESTIONS

QUESTION B2 [*20 Marks*]

B2 (a) The service times for customers coming through a checkout counter in a retail store are independent random variables with mean 1.5 minutes and variance 1.0. Approximate the probability that 100 customers can be served in less than 2 hours of total service time.

[6 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.

[8 Marks]

(c) The reaction of an individual to a stimulus in a psychological experiment may take one of two forms, A or B. If an experimenter wishes to estimate the probability p that a person will react in manner A, how many people must be included in the experiment? Assume that the experimenter will be satisfied if the error of estimation is less than 0.04 with probability equal to 0.90. Assume also that he expects p to lie somewhere in the neighborhood of 0.6.

[6 Marks]

QUESTION B3 [*20 Marks*]

B3 (a) Let $Y_1, \ldots, Y_n$ be a random sample from a population with pdf

$$fy) = \begin{cases} \frac{1}{\alpha}y^{(1-\alpha)/\alpha}, & \text{for } 0 < y < 1; \alpha > 0 \\ 0, & \text{otherwise} \end{cases}$$

(i) Show that the maximum likelihood estimator of α is $\hat{\alpha} = -(1/n)\sum_{i=1}^{n}\ln(Y_i)$.

[5 Marks]

(ii) Is $\hat{\alpha}$ a consistent estimator of α?

[5 Marks]

(b) Now suppose $f(y, \theta) = \frac{1}{\theta}e^{-y/\theta}, x > 0$.

(i) State the factorisation criterion for sufficient statistics and use it to find a sufficient for θ.

[5 Marks]

(ii) Show that the sufficient estimator found in 3b(i) is an unbiased estimator of θ.

[5 Marks]

QUESTION B4 [*20 Marks*]

B4 (a) A student examined the effect of varying the water/cement ratio on the strength of concrete that had been aged 28 days. For concrete with a cement content of 200 pounds per cubic yard, the student obtained the data presented in the Table below.

Water/Cement ratio	Strength (100 ft/lb)
1.21	1.302
1.29	1.231
1.37	1.061
1.46	1.040
1.62	0.803
1.79	0.711

Let Y denote the strength and x denote the water/cement ratio.

(i) Fit the model $E(Y) = \beta_0 + \beta_1 x$.

[4 Marks]

(ii) Test $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 < 0$ with $\alpha = 0.05$. Identify the corresponding attained significance level.

[8 Marks]

(iii) Find a 90% condence interval for the expected strength of concrete when the water/cement ratio is 1.5 Explain what would happen to the confidence interval if we computed the interval around the water/cement ratio is 2.7

[8 Marks]

QUESTION B5 [*20 Marks*]

B5 (a) An experimenter has prepared a drug dosage level that she claims will induce sleep for 80% of people suffering from insomnia. After examining the dosage,we feel that her claims regarding the effectiveness of the dosage are inflated. In an attempt to disprove her claim, we administer her prescribed dosage to 20 insomniacs and we observe Y, the number for whom the drug dose induces sleep. We wish to test the hypothesis $H_0 : p = 0.8$ versus the alternative, $H_a : p < 0.8$. Assume that the rejection region $\{y \leq 1\}$ is used.

(i) In terms of this problem, what is a type I error? Find α.

[5 Marks]

(ii) In terms of this problem, what is a type II error? Find β when $p = 0.6$.

[5 Marks]

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

[10 Marks]

QUESTION B6 [*20 Marks*]

B6 (a) In Bayesian inference define what is meant by 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]

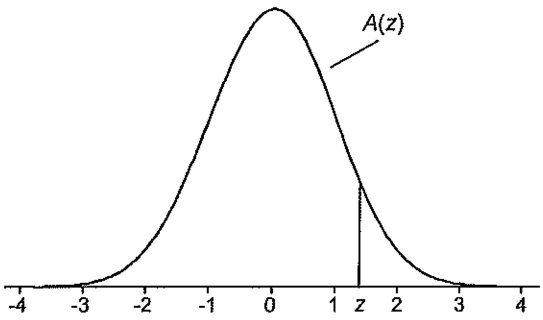
(ii) Find the Bayes estimators for p for $\alpha = 10, \beta = 30, n = 25$, and $\sum y_i = 10$.

[5 Marks]

END OF EXAMINATION PAPER

TABLE A.1

Cumulative Standardized Normal Distribution



$A(z)$ is the integral of the standardized normal distribution from $-\infty$ to z (in other words, the area under the curve to the left of z). It gives the probability of a normal random variable not being more than z standard deviations above its mean. Values of z of particular importance:

z	$A(z)$	
1.645	0.9500	Lower limit of right 5% tail
1.960	0.9750	Lower limit of right 2.5% tail
2.326	0.9900	Lower limit of right 1% tail
2.576	0.9950	Lower limit of right 0.5% tail
3.090	0.9990	Lower limit of right 0.1% tail
3.291	0.9995	Lower limit of right 0.05% tail

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999							

Table A.2

t Distribution: Critical Values of t

Degrees of freedom	Significance level — Two-tailed test: / One-tailed test:	10% / 5%	5% / 2.5%	2% / 1%	1% / 0.5%	0.2% / 0.1%	0.1% / 0.05%
1		6.314	12.706	31.821	63.657	318.309	636.619
2		2.920	4.303	6.965	9.925	22.327	31.599
3		2.353	3.182	4.541	5.841	10.215	12.924
4		2.132	2.776	3.747	4.604	7.173	8.610
5		2.015	2.571	3.365	4.032	5.893	6.869
6		1.943	2.447	3.143	3.707	5.208	5.959
7		1.894	2.365	2.998	3.499	4.785	5.408
8		1.860	2.306	2.896	3.355	4.501	5.041
9		1.833	2.262	2.821	3.250	4.297	4.781
10		1.812	2.228	2.764	3.169	4.144	4.587
11		1.796	2.201	2.718	3.106	4.025	4.437
12		1.782	2.179	2.681	3.055	3.930	4.318
13		1.771	2.160	2.650	3.012	3.852	4.221
14		1.761	2.145	2.624	2.977	3.787	4.140
15		1.753	2.131	2.602	2.947	3.733	4.073
16		1.746	2.120	2.583	2.921	3.686	4.015
17		1.740	2.110	2.567	2.898	3.646	3.965
18		1.734	2.101	2.552	2.878	3.610	3.922
19		1.729	2.093	2.539	2.861	3.579	3.883
20		1.725	2.086	2.528	2.845	3.552	3.850
21		1.721	2.080	2.518	2.831	3.527	3.819
22		1.717	2.074	2.508	2.819	3.505	3.792
23		1.714	2.069	2.500	2.807	3.485	3.768
24		1.711	2.064	2.492	2.797	3.467	3.745
25		1.708	2.060	2.485	2.787	3.450	3.725
26		1.706	2.056	2.479	2.779	3.435	3.707
27		1.703	2.052	2.473	2.771	3.421	3.690
28		1.701	2.048	2.467	2.763	3.408	3.674
29		1.699	2.045	2.462	2.756	3.396	3.659
30		1.697	2.042	2.457	2.750	3.385	3.646
32		1.694	2.037	2.449	2.738	3.365	3.622
34		1.691	2.032	2.441	2.728	3.348	3.601
36		1.688	2.028	2.434	2.719	3.333	3.582
38		1.686	2.024	2.429	2.712	3.319	3.566
40		1.684	2.021	2.423	2.704	3.307	3.551
42		1.682	2.018	2.418	2.698	3.296	3.538
44		1.680	2.015	2.414	2.692	3.286	3.526
46		1.679	2.013	2.410	2.687	3.277	3.515
48		1.677	2.011	2.407	2.682	3.269	3.505
50		1.676	2.009	2.403	2.678	3.261	3.496
60		1.671	2.000	2.390	2.660	3.232	3.460
70		1.667	1.994	2.381	2.648	3.211	3.435
80		1.664	1.990	2.374	2.639	3.195	3.416
90		1.662	1.987	2.368	2.632	3.183	3.402
100		1.660	1.984	2.364	2.626	3.174	3.390
120		1.658	1.980	2.358	2.617	3.160	3.373
150		1.655	1.976	2.351	2.609	3.145	3.357
200		1.653	1.972	2.345	2.601	3.131	3.340
300		1.650	1.968	2.339	2.592	3.118	3.323
400		1.649	1.966	2.336	2.588	3.111	3.315
500		1.648	1.965	2.334	2.586	3.107	3.310
600		1.647	1.964	2.333	2.584	3.104	3.307
∞		1.645	1.960	2.326	2.576	3.090	3.291