MACHAKOS UNIVERSITY

University Examinations for 2021/2022 Academic Year

SCHOOL OF BUSINESS AND ECONOMICS

DEPARTMENT OF ECONOMICS

FOURTH YEAR …… SEMESTER EXAMINATION FOR

BACHELOR OF ECONOMICS & STATISTICS

BACHELOR OF ECONOMICS

EES 301: STATISTICS FOR ECONOMISTS II

DATE: **TIME:**

INSTRUCTIONS:

Answer Question **ONE** and any other **TWO** questions

QUESTION ONE (COMPULSORY) (30 MARKS)

a) Distinguish between consistency and efficiency in reference to properties of a statistical estimator. (4 marks)

b) Highlight conditions necessary for the students – t distribution to be used in statistical analysis. (3 marks)

c) The makers of Duracell batteries want to demonstrate that their size AA battery lasts on an average of at least 45 minutes longer than Duracell's main competitor, the Energizer. Two independent random samples of 100 batteries of each kind are selected. The sample average lives for Duracell and Energizer batteries are found to be $\bar{X}_1 = 308$ minutes and $\bar{X}_2 = 254$ minutes respectively. Assume $\sigma_1 = 84$ minutes and $\sigma_2 = 67$ minutes. Is there any evidence to substantiate Duracell's claim that its batteries last, on average, at least 45 minutes longer than Energizer of the same size? (6 marks)

d) Find the regression equation of X and Y and the coefficient of correlation from the following data:
$\sum X_i = 60, \quad \sum Y_i = 60, \quad \sum X_i Y_i = 1{,}150, \quad \sum X_i^2 = 4{,}160, \quad \sum Y_i^2 = 1{,}720$ and N = 10. (8 marks)

e) The following 3x3 contingency table contains observed frequencies for a sample of 240.

f) Test the independence of the row and column variables with $\alpha = 0.05$ (6 marks)

	Column variable		
Row variable	A	B	C
P	20	30	20
Q	30	60	25
R	10	15	30

g) Distinguish between Central limit theorem and finite population correction factor (3 marks)

QUESTION TWO (20 MARKS)

a) Ndiku Stockbrokers, an investment firm, employs 20 investment analysts. Every morning each analyst is assigned up to five stocks to evaluate. The assignments that were made on a particular day are:

Outcome (X_i) (No. of Stock)	1	2	3	4	5
Frequency of X_i (No. of Analyst)	4	2	3	5	6

i) Generate a probability distribution (3 marks)

ii) Determine the mean and variance of the probability distribution (4 marks)

b) A firm conducts a study to determine if absenteeism of day workers is different from those who work the night shift. A comparison is made of 150 workers for each shift. The results show that 37 day workers have been absent at least five times over the past year, while 52 night workers have missed at least five times. What does this reveal about the tendency for absenteeism among the workers? Calculate a 90% Confidence Interval for the difference in the proportion of workers on the two shifts who missed at least five times. (6 marks)

c) The time it takes an international telephone operator to place an overseas phone call is normally distributed with a mean of 45 seconds and a standard deviation of 10 seconds.

i) What is the probability that my call will go through in less than 1 minute? (2 marks)

ii) What is the probability that my call will get through in less than 40 seconds? (2 marks)

iii) What is the probability that I will have to wait more than 70 seconds for my call to go through? (3 marks)

QUESTION THREE (20 MARKS)

a) An institution wants its entire staff to be computer literate. Three institutions A, B, and C are available. From past experience, 14 people who underwent training from the institutions were evaluated as follows: -

A	3	4	2	4	2
B	4	5	4	6	
C	6	6	6	4	3

i) State the appropriate null and alternative hypothesis for determining whether a difference exists in training offered. (4 marks)

ii) Construct the relevant ANOVA table and test the null hypothesis at $a = 0.05$ (8 marks)

b) Use relevant examples to distinguish between Poisson and Exponential distributions and explain the situations where each distribution may be used. (6 marks)

c) Explain the meaning of P-Value in tests of hypothesis (2 marks)

QUESTION FOUR (20 MARKS)

a) Use a relevant example to distinguish between an estimate and an estimator (4 marks)

b) The following data shows the number of shares covered by option grants and the number of shares outstanding for 13 companies *(Bloomberg Personal Finance January .2000)*

Companies	*Share of option Grants outstanding(millions)*	*Common shares Outstanding (Millions)*
Adobe systems	*20.3*	*61.8*
Apple compute	*52.7*	*160.9*
Applied materials	*109.1*	*375.5*
Auto Desk	*15.7*	*58.9*
Best Buy	*44.2*	*203.8*
Fruit of the loom	*14.2*	*66.9*
ITT industries	*18.0*	*87.9*
Merril lynch	*89.9*	*365.5*
Novel	*120.2*	*335.0*
Parametric technology	*78.3*	*269.3*
Reebok international	*12.8*	*56.1*
Silicon graphics	*52.6*	*188.8*
Toy world	*54.8*	*247.6*

Required

i) Develop the estimated regression equation that could be used to estimate the number of shares of option grants outstanding given the number of common shares outstanding (8 marks)

ii) Use the estimated regression equation to estimate the number of shares of option grants outstanding for a company that has 150 million shares of common stock outstanding (2 marks)

iii) Find a 95% confidence interval for the slope of the regression line (4 marks)

c) Distinguish between One-tailed test and Two-tailed test (2 marks)

QUESTION FIVE (20 MARKS)

a) A manufacturer of screws has noticed that, on average, 0.02 proportion of screws produced are defective. A random sample of 400 screws is examined for the proportion of faulty screws. Find the probability that the proportion of the defective screws in the sample is between 0.01 and 0.03 (6 marks)

b) In an opinion survey regarding a particular political issue, there was some question as to whether or not the eligible voters under 25 years of age might view the issue differently from those over 25 years. 1,500 individuals of those over 25 years were interviewed, and 1,000 of those under 25 years were interviewed with the following results

	Opposed	Undecided	Favour	Total
Under 25	400	100	500	1,000
Over 25	600	400	500	1,500
Total	1,000	500	1.000	2,500

Test the null hypothesis that there is no evidence of a difference of opinion due to the different age grouping; take $\alpha = 0.05$ (10 marks)

c) State four properties of the standard normal distribution (4 marks)

Useful Formulae and Statistical Tables

Normal distribution

Z-test

To find the probability with which a score X comes from a normal distribution with a mean of μ and a standard deviation of σ.

1. Calculate:

$$z = \frac{X - \mu}{\sigma}$$

2. Ignore the sign of z. The table below gives the one-tailed probability. Double the p values in the table for two-tailed probability.

Z-table

z	p	z	p	z	p	z	p
0.00	0.5000	1.52	0.0643	2.16	0.0154	2.76	0.0029
0.05	0.4801	1.54	0.0618	2.18	0.0146	2.78	0.0027
0.10	0.4602	1.56	0.0594	2.20	0.0139	2.80	0.0026
0.15	0.4404	1.58	0.0571	2.22	0.0132	2.82	0.0024
0.20	0.4207	1.60	0.0548	2.24	0.0125	2.84	0.0023
0.25	0.4013	1.62	0.0526	2.26	0.0119	2.86	0.0021
0.30	0.3821	1.64	0.0505	2.28	0.0113	2.88	0.0020
0.35	0.3632	1.66	0.0485	2.30	0.0107	2.90	0.0019
0.40	0.3446	1.68	0.0465	2.32	0.0102	2.92	0.0018
0.45	0.3264	1.70	0.0446	2.34	0.0096	2.94	0.0016
0.50	0.3085	1.72	0.0427	2.36	0.0091	2.96	0.0015
0.55	0.2912	1.74	0.0409	2.38	0.0087	2.98	0.0014
0.60	0.2743	1.76	0.0392	2.40	0.0082	3.00	0.0013
0.65	0.2578	1.78	0.0375	2.42	0.0078	3.02	0.0013
0.70	0.2420	1.80	0.0359	2.44	0.0073	3.04	0.0012
0.75	0.2266	1.82	0.0344	2.46	0.0069	3.06	0.0011
0.80	0.2119	1.84	0.0329	2.48	0.0066	3.08	0.0010
0.85	0.1977	1.86	0.0314	2.50	0.0062	3.10	0.0010
0.90	0.1841	1.90	0.0287	2.52	0.0059	3.12	0.0009
1.00	0.1587	1.92	0.0274	2.54	0.0055	3.14	0.0008
1.05	0.1469	1.94	0.0262	2.56	0.0052	3.16	0.0008
1.10	0.1357	1.96	0.0250	2.58	0.0049	3.18	0.0007
1.15	0.1251	2.00	0.0228	2.60	0.0047	3.20	0.0007
1.20	0.1151	2.02	0.0217	2.62	0.0044	3.22	0.0006
1.25	0.1056	2.04	0.0207	2.64	0.0041	3.24	0.0006
1.30	0.0968	2.06	0.0197	2.66	0.0039	3.26	0.0006
1.35	0.0885	2.08	0.0188	2.68	0.0037	3.30	0.0005
1.40	0.0808	2.10	0.0179	2.70	0.0035	3.50	0.0002
1.45	0.0735	2.12	0.0170	2.72	0.0033	3.75	0.0001
1.50	0.0668	2.14	0.0162	2.74	0.0031	4.00	0.0000

t-table

	Significance Level			
df	*0.1 two-tailed* (0.05 one-tailed)	*0.05 two-tailed* (0.025 one-tailed)	*0.025 two-tailed* (0.0125 one-tailed)	*0.01two-tailed* (0.005 one-tailed)
1	6.314	12.706	25.452	63.656
2	2.920	4.303	6.205	9.925
3	2.353	3.182	4.177	5.841
4	2.132	2.776	3.495	4.604
5	2.015	2.571	3.163	4.032
6	1.943	2.447	2.969	3.707
7	1.895	2.365	2.841	3.499
8	1.860	2.306	2.752	3.355
9	1.833	2.262	2.685	3.250
10	1.812	2.228	2.634	3.169
11	1.796	2.201	2.593	3.106
12	1.782	2.179	2.560	3.055
13	1.771	2.160	2.533	3.012
14	1.761	2.145	2.510	2.977
15	1.753	2.131	2.490	2.947
16	1.746	2.120	2.473	2.921
17	1.740	2.110	2.458	2.898
18	1.734	2.101	2.445	2.878
19	1.729	2.093	2.433	2.861
20	1.725	2.086	2.423	2.845

Unrelated samples t-test

***Equal* N**

1. Let N be the sample size of each group.
2. Calculate the mean for each group, $\overline{X}_1$ and $\overline{X}_2$
3. Calculate the variance for each group, s^2_1 and s^2_2.

$$s^2 = \frac{\sum\left(X - \overline{X}\right)^2}{N-1}$$

4. Calculate the t statistic:

$$t = \frac{\left(\overline{X}_1 - \overline{X}_2\right)}{\sqrt{\frac{s_1^2 + s_2^2}{N}}}$$

5. Ignore the sign of *t*. The result is significant if *t* is greater than the appropriate value in the *t*-table.

- *df = 2N-2.*

***Unequal* N**

1. Let N_1 and N_2 be the sample sizes of the two groups.
2. Calculate the mean for each group, $\overline{X}_1$ and $\overline{X}_2$
3. Calculate the variance for each group, s^2_1 and s^2_2.

$$s^2 = \frac{\sum (X - \overline{X})^2}{N - 1}$$

4. Calculate the pooled variance estimate:

$$s_p^2 = \frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}$$

5. Calculate the *t*-statistic:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_p^2 \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}}$$

6. Ignore the sign of *t*. The result is significant if *t* is greater than the appropriate value in the *t*-table.

- $df = N_1 + N_2 - 2$

F Table

Degrees of Freedom for Denominator	Significance Level: 0.025 one-tailed, 0.05 two-tailed Degrees of Freedom for Numerator											
	1	2	3	4	5	6	7	8	9	10	15	20
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.6	963.3	968.6	984.9	993.1
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.43	39.45
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.25	14.17
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.56
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.33
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.17
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.47
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	4.00
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.67
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.42
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.33	3.23
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.18	3.07
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.05	2.95
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	2.95	2.84
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.86	2.76
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.79	2.68
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.72	2.62
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.67	2.56
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.62	2.51
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.57	2.46

Chi-Square Test

- The formula for chi-square is:

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

where O = observed frequency and E = expected frequency.

Chi-Square Table

	Significance Level			
df	*0.1*	*0.05*	*0.025*	*0.01*
1	2.706	3.841	5.024	6.635
2	4.605	5.991	7.378	9.210
3	6.251	7.815	9.348	11.345
4	7.779	9.488	11.143	13.277
5	9.236	11.070	12.832	15.086
6	10.645	12.592	14.449	16.812
7	12.017	14.067	16.013	18.475
8	13.362	15.507	17.535	20.090
9	14.684	16.919	19.023	21.666
10	15.987	18.307	20.483	23.209
11	17.275	19.675	21.920	24.725
12	18.549	21.026	23.337	26.217
13	19.812	22.362	24.736	27.688
14	21.064	23.685	26.119	29.141
15	22.307	24.996	27.488	30.578
16	23.542	26.296	28.845	32.000
17	24.769	27.587	30.191	33.409
18	25.989	28.869	31.526	34.805
19	27.204	30.144	32.852	36.191
20	28.412	31.410	34.170	37.566