CONFIDENTIAL

UNIVERSITY COLLEGE TATI

UNIVERSITY COLLEGE TATI (UC TATI)

FINAL EXAMINATION QUESTION BOOKLET	
COURSE CODE	: BGE 2123
COURSE	: STATISTICS
SEMESTER/SESSION	: 1-2022/2023
DURATION	: 3 HOURS

Instructions:

1. This booklet contains **11** questions. Answer **ALL** questions.
2. All answers should be written in answer booklet.
3. Write legibly and draw sketches wherever required.
4. If in doubt, raise your hands and ask the invigilator.

DO NOT OPEN THIS BOOKLET UNTIL YOU ARE TOLD TO DO SO

THIS BOOKLET CONTAINS 13 PRINTED PAGES INCLUDING COVER PAGE

SECTION A (50 MARKS)

INSTRUCTION: ANSWER ALL QUESTIONS.

QUESTION 1

A manager of a company in Kuantan, Pahang wishes to estimate the average number of days of medical leave taken by the workers in his company last year. The company has 6 departments. The following table shows the number of workers in each department.

Department	Number of workers
A	20
B	15
C	10
D	25
E	30
F	15

The manager randomly selects a number of workers from each department. He decides to select a sample of 40 from 115 workers.

a) State the population and sample of interest? (2 marks)

b) Identify the variable of interest? (1 mark)

c) What is the sampling technique used for this study? (1 mark)

QUESTION 2

Based on the data from USA Today, the ages of a random sample of 300 adults who shop by catalogue are given below.

Age	18-22	23-27	28-32	33-37	38-42	43-47
Number	78	75	48	33	33	33

a) Calculate the mean age of the adults. (3 marks)

b) Calculate the median age of the adults. (4 marks)

c) Calculate the standard deviation of the age. (4 marks)

QUESTION 3

Suppose that during periods of transcendental meditation the reduction of a person's oxygen consumption is a random variable having a normal distribution with $\mu = 17.6cc$ per minute and $\sigma = 4.6cc$ per minute. Find the probabilities that the time it takes to develop one of the prints will be:

a) At least 1minutes. (3 marks)

b) Anywhere from 15 to 15.80 minutes. (4 marks)

QUESTION 4

There are 90 applicants for a job with the news department of a television station. Some of them are college graduates and some are not, some of them have at least three years' experience and some have not, with the exact breakdown being.

	College graduates	Not college graduates
At least three years' experience	18	9
Less than three years' experience	36	27

If the order in which the applicants are interviewed by the station manager is random, G is the event that the first applicants interviewed is a college graduates, and T is the event that the first applicant interviewed has at least three years' experience, determine each of the following probabilities directly from the entries and the row and column totals of the table:

a) $P(G)$ (2 marks)

b) $P(G \cap T)$ (2 marks)

c) $P(G' | T')$ (2 marks)

d) Are the events G and T independent? (3 marks)

QUESTION 5

The discrete random variable X has a probability distribution function given by $f(x)=k(1-x)^2$ for $x=-1,0,1,2$. Find:

a) The value of k. (2 marks)

b) $P(X\geq 0)$ (3 marks)

c) $E(X)$ (3 marks)

d) $V(X)$ (4 marks)

QUESTION 6

A research team at Cornell University conducted a study showing that approximately 10% of all businessmen who wear ties wear them so tight that they actually reduce blood flow to the brain, diminishing cerebral functions. At a board meeting of 20 businessman, all of them wear ties. Using binomial model, find:

a) The probability that at least one tie is too tight? (2 marks)

b) The probability that two ties are too tight? (2 marks)

c) The mean and variance. (3 marks)

SECTION B (30 MARKS)

INSTRUCTION: ANSWER ALL QUESTIONS.

QUESTION 1

The number of houses sold by an estate agent follows a Poisson distribution, with a mean of 2 per week.

a) Find the probability that in the next 4 weeks the estate agent sells exactly 3 houses. (3 marks)

b) The estate agent will receive a bonus if he sells more than 25 houses in the next 10 weeks. Use a suitable approximation to estimate the probability that the agent receives. (4 marks)

QUESTION 2

A mail-order house employs three stock clerks, U, V and W, who pull items from shelves and assemble them for subsequent verification and packaging. U makes a mistake in an order (gets a wrong item or the wrong quantity) one time in a hundred, V makes a mistake in an order five times in a hundred, and W makes a mistake in an order three times in a hundred. If U, V and W fill, respectively, 30,40 and 30 percent of all orders, what are the probabilities that:

a) A mistake will be made in an order. (3 marks)

b) If a mistake is made in an order, the order was filled by U? (2 marks)

c) If a mistake is made in an order, the order was filled by V? (2 marks)

QUESTION 3

A child psychiatrist is studying the mental development of children. A random sample of nine children were given a standard of questions appropriate to the age of each child. The number of irrelevant responses to the questions was recorded for each child. In the following data, x is an age of child in years and y is a number of irrelevant responses.

x	5	7	9	10	11	12
y	13	11	10	8	6	5

a) Find the Pearson's correlation between x and y. Explain the value.

(7 marks)

b) Find the linear regression equation of y on x using the least squares method.

(7 marks)

c) If a child is 8 years old, what does the least squares line predict for the number of irrelevant responses?

(2 marks)

SECTION C (20 MARKS)

INSTRUCTION: ANSWER ALL QUESTIONS.

QUESTION 1

An importer has ordered a large consignment of tomatoes. When it arrives, he examines a randomly chosen sample of 30 boxes and finds that 10 boxes contain bad tomatoes. Assuming that these boxes may be regarded as being a random sample from the boxes in the consignment, obtain an approximate 99% confidence interval for the proportion of boxes containing bad tomatoes.

(5 marks)

QUESTION 2

An Air Force base mess hall has received a shipment of 10,000 gallon size cans of cherries. The supplier claims that the average amount of liquid is 0.25 gallon per can. A government inspector took a random sample of 100 cans and found the average liquid content to be 0.28 gallon per can with a standard deviation of 0.10.

a) Does this indicate that the supplier's claim is difference? Test the claim at 5% level of significance. Use 5 step of hypothesis testing. (10 marks)

b) Set up a 99% confidence interval for the average amount of liquid. (5 marks)

---End of question---

FORMULA

$$\bar{x} = \frac{\sum x}{n} \text{ or } \frac{\sum fx}{\sum f}$$

$$\tilde{x} = L_m + \left(\frac{\frac{n}{2} - F}{f_m} \right) \cdot i$$

$$\hat{x} = L_{mo} + \left(\frac{f_0 - f_1}{(f_0 - f_1) + (f_0 - f_2)} \right) \cdot i$$

$$s^2 = \frac{1}{n-1} \left[\sum fx^2 - \frac{(\sum fx)^2}{n} \right]$$

Where

n	:	total frequency
L_m	:	lower median class boundary
L_{mo}	:	lower modal class boundary
F	:	cumulative frequencies for the classes before the median class
f_m	:	median class frequency
f_0	:	frequency of the class containing mode
f_1	:	frequency of the class before the class containing mode
f_2	:	frequency of the class after the class containing mode
i	:	class size

Probability

1. $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

2. a) $P(A \cap B) = P(B) \cdot P(A \mid B)$

   b) $P(A \cap B) = P(A) \cdot P(B)$ if and only if A and B are independent events.

3. $P(A \mid B) = \dfrac{P(A) \cdot P(B \mid A)}{P(A) \cdot P(B \mid A) + P(A') \cdot P(B \mid A')}$

4. a) $P(X = x) = \binom{n}{x} p^x q^{n-x} \quad x = 0, 1, .., n$ b) $P(X = x) = \dfrac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, ..$

5. a) $Z = \dfrac{X - np}{\sqrt{npq}}$ b) $Z = \dfrac{X - \lambda}{\sqrt{\lambda}}$

6. $E(X) = \sum_{-\infty}^{\infty} x \cdot f(x) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$

7. $E(X^2) = \sum_{-\infty}^{\infty} x^2 \cdot f(x) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx$

8. $Var(X) = E(X^2) - [E(X)]^2$

Regression and Correlation

1. $r = \dfrac{(\Sigma XY) - \dfrac{(\Sigma X)(\Sigma Y)}{n}}{\sqrt{\left[(\Sigma X^2) - \dfrac{(\Sigma X)^2}{n}\right]\left[(\Sigma Y^2) - \dfrac{(\Sigma Y)^2}{n}\right]}}$

2. $b = \dfrac{(\Sigma XY) - \left(\dfrac{(\Sigma X)(\Sigma Y)}{n}\right)}{\left((\Sigma X^2) - \dfrac{(\Sigma X)^2}{n}\right)}$

3. $a = \dfrac{(\Sigma y)}{n} - b\dfrac{(\Sigma x)}{n}$

Confidence Interval and Hypothesis Testing

Estimator	Confidence Interval	Test Statistics
$\mu_1 - \mu_2$	$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$Z_0 = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
	$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ $v = n_1 + n_2 - 2$	$T_0 = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
	$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$
	$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2,v} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2,v} s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ $v = n_1 + n_2 - 2$ $s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$	$T_0 = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

Confidence Interval and Hypothesis Testing

Estimator	Confidence Interval	Test Statistics
μ	$(\overline{x}) - z_{\alpha/2}\dfrac{s}{\sqrt{n}} < \mu < (\overline{x}) + z_{\alpha/2}\dfrac{s}{\sqrt{n}}$	$Z_o = \dfrac{\overline{X} - \mu_o}{s / \sqrt{n}} \sim N(0,1)$
	$(\overline{x}) - t_{\alpha/2, n-1}\dfrac{s}{\sqrt{n}} < \mu < (\overline{x}) + t_{\alpha/2, n-1}\dfrac{s}{\sqrt{n}}$	$T_o = \dfrac{\overline{X} - \mu_o}{s / \sqrt{n}} \sim t_{n-1}$

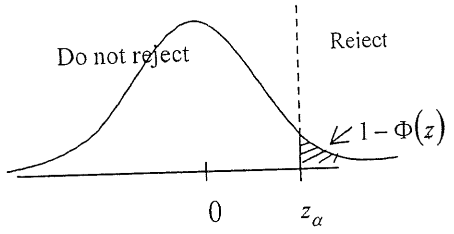
Estimator	Confidence Interval	Sample size
$\hat{p}$	$\hat{p} - z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$	$n = \dfrac{z^2_{\alpha/2} \cdot \hat{p}(1-\hat{p})}{E^2}$

APPENDIX I

Table I Standard Normal Distribution

$$1-\Phi(z)=P(Z>z)=\frac{1}{\sqrt{2\pi}}\int_{z}^{\infty} e^{-z^2/2}\,dz$$

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



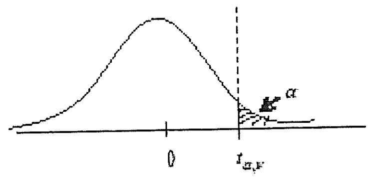
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
3.7	.000108	.000104	.000100	.000096	.000092	.000088	.000085	.000082	.000078	.000075
3.8	.000072	.000069	.000067	.000064	.000062	.000059	.000057	.000054	.000052	.000050
3.9	.000048	.000046	.000044	.000042	.000041	.000039	.000037	.000036	.000034	.000033
4.0	.000032									

5.0 → 0.0000002867 5.5 → 0.0000000190 6.0 → 0.0000000010

APPENDIX II

Table II Student's t Distribution

The table gives the value of $t_{\alpha,\nu}$ - the 100 α percentage point of the t distribution for ν degress of freedom.



ν \ α	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.706	31.821	63.657	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291