THE OPEN UNIVERSITY OF SRI LANKA
Faculty of Engineering Technology
Department of Mathematics & Philosophy of Engineering

Bachelor of Industrial Studies Honors / Bachelor of Software Engineering Honors

171

Final Examination (2020/2021)
MHZ4357: Applied Statistics

Index No: .....

Date: 28th January 2022 (Friday) **Time: 0930 hrs. – 1230 hrs.**

Instructions:

- *Part A is Compulsory*
  - *Provide short answers in given space.*
  - *Do not need to show any workings.*
- *Answer **five (05)** questions only from **Part B**.*
  - *Provide answers in separate sheets (answer booklet) which will be given in the examination.*
  - *Show all your workings.*
- *Number of pages in the paper is **Seven (07).***
- *All the symbols are in standard notation unless they are defined.*
- *All the relevant statistical tables are attached with this paper.*
- *Do not need to use graph sheets in this paper.*
- ***Attach the Part A to the Answer script of the Part B***

Index No: .....

Part A

Provide short answers in the given space. Do not need to show any workings or works. Please attach this part to your answer script.

01. A researcher wishes to estimate the average monthly living cost per person in a particular city. Then the average monthly living cost per person of a sample of 30 people in the same city is an example of a ..... (10%)

02. A null hypothesis of a test was accepted at 5% significant level. What will be the result of the same hypothesis test at 1% significant level? ..... (10%)

03. What is the objective of inferential statistics?

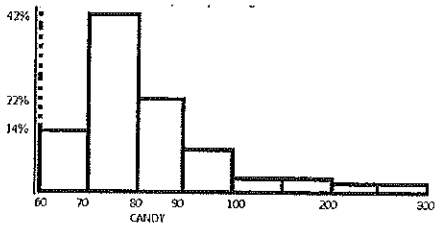
..... (10%)

04. Let the random variable X follow a normal distribution with mean μ and standard deviation σ. Then what will be the standard deviation of the sampling distribution of sample mean$(\bar{X})$ when the sample size is 35? ..... (10%)

05. What is the difference between a Nominal scale and an Ordinal scale?

..... (10%)

06. What is the best suited central tendency measure for the data, which represented by the following histogram? ..... (10%)



07. Let A and B be two independent events such that $P(A) = \frac{1}{5}$ while $P(A \cup B) = \frac{1}{2}$. Then what is the value of $P(B)$? ..... (10%)

08. Suppose that a certain hypothesis testing results a p-value of 0.084. At the 5% significant level what will be the conclusion of this test? ..... (10%)

09. Define the term 'sample space'.

..... (10%)

10. Refer to the discrete probability distribution provided in the table below.

$X = x$	0	1	2	3	4
$P(X = x)$	0.04	0.11	0.45	0.23	0.17

Find the probability that X is greater than to 2. (Round to 2 decimal places.)

..... (10%)

Part B

Q1. a) A plantation of apples is consisting with two sites, Site A and Site B. The following table gives the number of apples in a sample of 1000, according to the size of them as Small, Medium, and Large from each site.

Category	Number of apples from Site A	Number of apples from Site B
Small	360	310
Medium	420	380
Large	220	310

I. Draw an appropriate graph to present the production from Site A and interpret your graph in appropriate way. (15%)

II. Draw an appropriate graph to compare the productions from Site A and Site B. Interpret your graph in meaningful way. (20%)

b) The thickness of a sample of 20 steel plates were measured and the results (in millimeters) were as follows.

6.2 6.5 6.5 6.6 6.7 6.8 6.8 6.9 6.9 7.0
7.0 7.1 7.2 7.2 7.3 7.3 7.4 7.5 7.6 7.8

I. Compile a table showing the frequency distribution and the relative frequency distribution by choosing class size as 0.3. (25%)

II. Draw the histogram for the above frequency distribution. (20%)

III. Draw the frequency polygon for above frequency distribution on the same diagram of the part b) II. (15%)

IV. Comment about the shape of the distribution of the sample data. (05%)

Q2 The following table gives the time taken by employees to complete an operation was recorded on 70 occasions.

Time (Hours)	10	10.5	11	11.5	12	12.5	13	135
Frequency	4	8	14	10	20	10	3	1

a) Determine the central tendency measures (Mean, Median and Mode) of the above sample. (20%)

b) Find the range, standard deviation, and inter-quartile range for the sample data. (30%)

c) What are the best suited measures for the central tendency and the dispersion respectively? Explain your answers. (20%)

d) Draw the box-whisker plot for the above data. (10%)

e) Are there any outliers in the given data set? Give reason for your answer (20%)

Q3 a) A packet contains 100 washers, which 24 are made of brass, 36 are made of copper and the remainders are made of steel. One washer is taken at random and retained to the packet and second washer similarly drawn. Determine the probability that,

I. both washers are steel (10%)

II. the first is brass and the second is copper (10%)

III. at least one of the washers is brass (20%)

b) Let X and Y be two events from a same sample space with $P(X) = 0.5$, $P(Y) = 0.4$, and $P(X \cup Y) = 0.8$. Then what is the value of $P(Y|\bar{X})$? (20%)

c) A software company develops mobile applications to satisfy client requirements. The following table shows job allocation rates and the success rates for three teams of the company: Team A, Team B, and Team C.

Team	Job allocation % out of Total Jobs	Probability of success
A	35%	0.95
B	35%	0.80
C	30%	0.98

I. Find the probability that any job done by this company is success. (20%)

II. If you are given that a certain job is success, find the probability that the job was done by Team C. (20%)

Q4 a) Discuss the characteristics of the Geometric distribution and Binomial Distribution (20%)

b) An office was appointed to check each item produce by a certain machine one by one. The probability that any item produced by the said machine is a defective item is 0.02.

I. Find the probability that the officer detects the first defective item at the 4^{th} item that he checks on a specific day. (15%)

II. Find the probability that the officer has to check more than 3 items to detect first defective item. (25%)

c) The length of pregnancy of a cow from conception to birth approximates a normal distribution with a mean of 283days and standard deviation of 12 days.

I. What proportion of pregnancy of a cow will last within 285 days? (20%)

II. Find the length of pregnancy period that makes the 70% of the full pregnancy period? (20%)

Q5 a) Define the following terms.

I. Point estimate (10%)

II. Interval estimate (10%)

III. Sampling distribution of a statistic (15%)

b) Suppose a researcher measures the pH level of soil samples of a site to check whether there is an effect of nitrogen fertilizer. He collected 40 samples of soil in different areas of the site, and he found that sample average level of pH is 6.48 and the sample standard deviation is 0.712.

I. State the point estimate for the population mean of the pH level of the soil in the given site. (10%)

II. Find the standard error of the population mean of the pH level of the soil in the given site. (10%)

III. Find the interval estimate for the population mean of the pH level of the soil in the given site at 1% significant level. (25%)

IV. Interpret the findings in Q5 part b) III. (10%)

V. State the assumptions you made for the calculations in Q5part b) III clearly. (10%)

Q6 The volume of the juice in a cup which filled by a machine is set to be 250ml. However, customers have complained to the manufacturer that the volume of the juice in a cup is less than 250ml. So, the machine owner collected juice volume data from a sample of 20 cups of juice and identified that the sample mean of the volume of juice is 248ml and sample standard deviation is 25ml. Then owner of the juice machine was testing a hypothesis to see the validity of the customers compline.

a) State the corresponding hypothesis (Null and alternative) that the owner of the juice machine test to see the validity of the customers compline. (20%)

b) Find the test statistics associated with this hypothesis testing (20%)

c) If the owner wants to test this hypothesis at 5% significant level, then what is the critical value/s associates with this test? (15%)

d) Test the hypothesis at 5% significant level and state your decision regarding the hypothesis. (20%)

e) Interpret your findings in question Q6 d). (10%)

f) State the assumptions that you made for this hypothesis testing. (15%)

Q7 A study was conducted to determine whether there is an impact of the number of years of experience to the annual salary for a particular profession. The following table provide the experience in years and the annual salary of 10 number of employees in the particular profession.

Number of years of Experience	13	16	22	2	8	26	19	20	1	4
Current annual salary (Rs. 00000)	26.1	33.2	36.5	16.5	26.4	37.1	33.8	36.1	16.9	19.8

a) Identify the independent variable and dependent variable. (10%)

b) Calculate the correlation coefficient for above data. Interpret your answer of the value of correlation coefficient. (20%)

c) Draw the scatter plot for dependent variable against the independent variable. What can you say about the relationship between independent variable and dependent variable according to the plot? (20%)

d) Find the equation of the best fitted simple linear regression model of the above data. (30%)

e) Interpret the found regression coefficient in part d). (10%)

f) Use the estimated line of regression to annual salary of an employee who has 5 years' experience. (10%)

End.

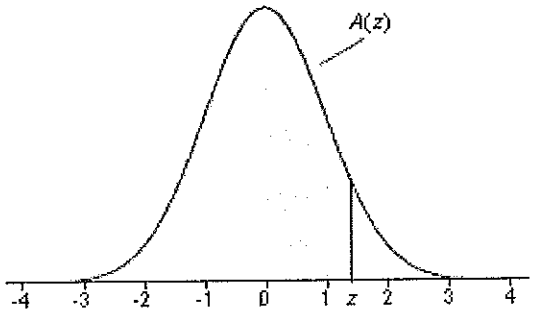
Useful Formulars

Correlation Coefficient: $r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$

Regression Coefficient: $\widehat{\beta_1} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n\sum x^2 - (\sum x)^2}}$

Z-table

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							

t Distribution: Critical Values of *t*

Degrees of freedom	*Two-tailed test:* / *One-tailed test:*	*Significance level* 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