UNIRAZAK
UNIVERSITI TUN ABDUL RAZAK

FINAL EXAMINATION
MARCH 2024

COURSE TITLE	INTRODUCTION TO STATISTICAL ANALYSIS
COURSE CODE	EMAT3153
DATE/DAY	27 JUNE 2024 / THURSDAY
TIME/DURATION	02:00 PM - 04:00 PM / 02 Hour(s) 00 Minute(s)

INSTRUCTIONS TO CANDIDATES :

1. Please read the instruction under each section carefully.
2. Candidates are reminded not to bring into examination hall/room any form of written materials or electronic gadget except for stationery that is permitted by the Invigilator.
3. Students who are caught breaching the Examination Rules and Regulation will be charged with an academic dishonesty and if found guilty of the offence, the maximum penalty is expulsion from the University.

(This Question Paper consists of 4 Printed Pages including front page)

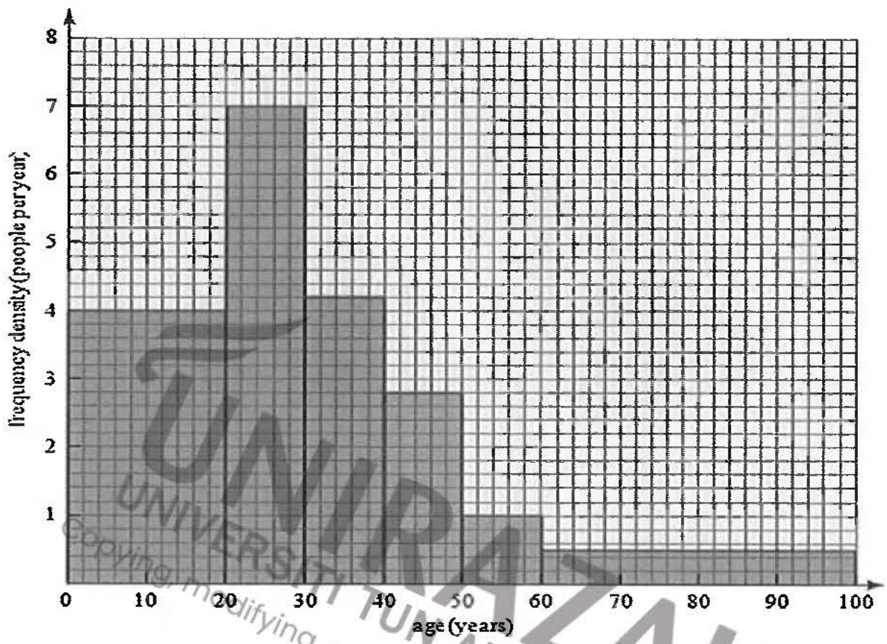
DO NOT OPEN THE QUESTION PAPER UNTIL YOU ARE TOLD TO DO SO

This question paper consists of ONE (1) section. Answer ALL questions in the answer booklet provided. **[50 MARKS]**

SECTION A

There are FIVE (5) questions only. Answer all in the provided answer sheets.

1. A random sample of people was asked how old they were when they first met their partner. The histogram represents this information.



(a) What is the modal age group?

(1 mark)

(b) How many people took part in the survey?

(2 marks)

(c) Find an estimate for the mean age that a person first met their partner.

(2 marks)

(d) Draw a cumulative frequency curve for the data and use the curve to provide an estimate for the median.

(5 marks)

2. Assuming the distribution of the heights of adult men is normal, with mean 174 cm and standard deviation 7 cm. Giving your answers to 2 significant figures, find the probability that a randomly selected adult man is:

(a) under 185 cm

(2 marks)

(b) over 185 cm

(2 marks)

(c) over 180 cm

(2 marks)

(d) between 180 cm and 185 cm

(2 marks)

(e) under 170 cm

(2 marks)

3. The mean breaking strength of a cables supplied by a manufacturer is 1800 with the standard deviation of 100. By a new technique in the manufacturing process, it is claimed that the breaking strength of the cable has increased. To test this claim a sample of 50 cables is tested and is found that the mean breaking strength is 1850.

Can we support the claim at 1% level of significance?

(10 marks)

4. The resistances (in ohms) of a random sample from a batch of resistors were

Resistors	2314	2456	2389	2361	2360	2332	2402

Assuming that the sample is from a normal distribution calculate

(a) a 95% confidence interval for the mean,

(5 marks)

(b) a 90% confidence interval for the mean.

(5 marks)

5. The manager of MZ System randomly selected 10 sales representatives and determined the number of sales calls each one made last month and the number of units of the product he or she sold last month.

Sales Representative	Number of Sales Calls	Number of Unit Sold
Ali	14	28
Budi	35	66
Chin	22	38
Fatimah	29	70
Henry	6	22
Cami	15	27
Resti	17	28
Rose	20	47
Seok	12	14
Siti	29	68

(a) Determine the coefficient of correlation for the table above.

(8 marks)

(b) Summarizes the strength and direction of the coefficient of correlation based on the result from (a).

(2 marks)

***** END OF QUESTION PAPER *****

Table & Formulae

Standard Normal Probabilities

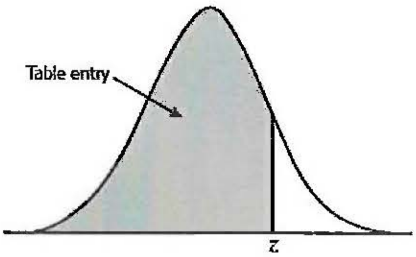
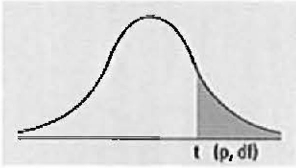


Table entry for z is the area under the standard normal curve to the left of z.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Numbers in each row of the table are values on a *t*-distribution with (*df*) degrees of freedom for selected right-tail (greater-than) probabilities (*p*).



df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%

$n =$ sample size

$N =$ population size

$f =$ frequency

$\Sigma = sum$

$w = weight$

Sample mean: $\bar{x} = \dfrac{\Sigma x}{n}$

Population mean: $\mu = \dfrac{\Sigma x}{N}$

Weighted mean: $\bar{x} = \dfrac{\Sigma(w \bullet x)}{\Sigma w}$

Mean for frequency table: $\bar{x} = \dfrac{\Sigma(f \bullet x)}{\Sigma f}$

Midrange $= \dfrac{\text{highest value + lowest value}}{2}$

Range = Highest value - Lowest value

Sample standard deviation: $s = \sqrt{\dfrac{\Sigma(x-\bar{x})^2}{n-1}}$

Population standard deviation: $\sigma = \sqrt{\dfrac{\Sigma(x-\mu)^2}{N}}$

Sample variance: s^2

Population variance: σ^2

Limits for Unusual Data

Below : $\mu - 2\sigma$

Above: $\mu + 2\sigma$

Empirical Rule

About 68%: μ-σ to $\mu+\sigma$

About 95%: μ-2σ to $\mu+2\sigma$

About 99.7%: μ-3σ to $\mu+3\sigma$

Sample coefficient of variation: $CV = \dfrac{s}{\bar{x}} \bullet 100\%$

Population coefficient of variation: $CV = \dfrac{\sigma}{\mu} \bullet 100\%$

Sample standard deviation for frequency table:

$$s = \sqrt{\frac{n\,[\,\Sigma(f \bullet x^2)\,] - [\,\Sigma(f \bullet x)\,]^2}{n\,(n-1)}}$$

Sample z-score: $z = \dfrac{x-\bar{x}}{s}$

Population z-score: $z = \dfrac{x-\mu}{\sigma}$

Interquartile Range: (IQR) $= Q_3 - Q_1$

Modified Box Plot Outliers

lower limit: Q_1 - 1.5 (IQR)

upper limit: Q_3 + 1.5 (IQR)

Multiplication rule for independent events
$P(A \text{ and } B) = P(A) \bullet P(B)$

General multiplication rules
$P(A \text{ and } B) = P(A) \bullet P(B, \text{ given } A)$
$P(A \text{ and } B) = P(A) \bullet P(A, \text{ given } B)$

Addition rule for mutually exclusive events
$P(A \text{ or } B) = P(A) + P(B)$

General addition rule
$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

Permutation rule: ${}_nP_r = \dfrac{n!}{(n-r)!}$

Combination rule: ${}_nC_r = \dfrac{n!}{r!(n-r)!}$

One Sample Confidence Interval

for proportions (p): ($np > 5$ and $nq > 5$)

$\hat{p} - E < p < \hat{p} + E$
where $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

$\hat{p} = \dfrac{r}{n}$

for means (μ) when σ is known:

$\bar{x} - E < \mu < \bar{x} + E$
where $E = z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

for means (μ) when σ is unknown:

$\bar{x} - E < \mu < \bar{x} + E$
where $E = t_{\alpha/2}\dfrac{s}{\sqrt{n}}$
with $d.f. = n-1$

for variance (σ^2): $\dfrac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \dfrac{(n-1)}{\chi_L^2}$

with $d.f. = n-1$

Mean of a discrete probability distribution:
$\mu = \sum[x \bullet P(x)]$

Standard deviation of a probability distribution:

$\sigma = \sqrt{\sum[x^2 \bullet P(x)] - \mu^2}$

Binomial Distributions
r = number of successes (or x)
p = probability of success
q = probability of failure
$q = 1 - p$ $\quad$ $p + q = 1$

Binomial probability distribution
$P(r) = {}_nC_r p^r q^{n-r}$

Mean: $\mu = np$

Standard deviation: $\sigma = \sqrt{npq}$

Normal Distributions

Raw score: $x = z\sigma + \mu$

Standard score: $z = \dfrac{x-\mu}{\sigma}$

Mean of $\bar{x}$ distribution: $\mu_{\bar{x}} = \mu$

Standard deviation of $\bar{x}$ distribtuion: $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$
(standard error)

Standard score for $\bar{x}$: $z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

One Sample Hypothesis Testing

for p $(np > 5$ and $nq > 5)$: $z = \dfrac{\hat{p} - p}{\sqrt{pq/n}}$

where $q = 1 - p$; $\hat{p} = r/n$

for μ (σ known): $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

for μ (σ unknown): $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ with $d.f. = n-1$

for σ^2: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$ with $d.f. = n-1$

Regression and Correlation

Linear Correlation Coefficient (r)

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

OR

$r = \dfrac{\Sigma(z_x z_y)}{n-1}$ where $z_x =$ z score for x and $z_y =$ z score for y

Coefficient of Determination: $r^2 = \dfrac{\text{explained variation}}{\text{total variation}}$

Standard Error of Estimate: $s_e = \sqrt{\dfrac{\Sigma(y - \hat{y})^2}{n-2}}$

or $s_e = \sqrt{\dfrac{\sum y^2 - b_0 \sum y - b_1 \sum xy}{n-2}}$

Prediction Interval: $\hat{y} - E < y < \hat{y} + E$

where $E = t_{\alpha/2}\, s_e \sqrt{1 + \dfrac{1}{n} + \dfrac{n(x_0 - \bar{x})^2}{n(\Sigma x^2) - (\Sigma x)^2}}$

Sample test statistic for r

$t = \dfrac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}$ with $d.f. = n-2$