THE PAPUA NEW GUINEA UNIVERSITY OF TECHNOLOGY

DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE

FIRST SEMESTER SUPPLEMENTARY EXAMINATIONS – 2022

SECOND YEAR BACHELOR IN COMPUTER SCIENCE

CS221 – APPLIED STATISTICS

TIME ALLOWED: 3 HOURS

INFORMATION FOR CANDIDATES

1. Write your name and student number clearly on the front of the examination answer booklet.
2. You have 10 minutes to read this paper. You must not begin writing during this time.
3. This paper contains FIVE (5) questions. You are to **answer ALL** the questions.
4. All answers must be written in examination answer booklets provided. No other written materials will be accepted.
5. Start the answer for each question on a **new** page. Do **not** use red ink.
6. Notes, textbooks, mobile phones and other recording devices are not allowed in the examination room.
7. Scientific and business calculators are allowed in the examination room.
8. A formula sheet and a normal distribution graph is attached.

MARKING SCHEME

Marks are indicated at the beginning of each question. The total is **100 marks.**

QUESTION 1 [(3 + 7 + 2) + 7 + 3 = 22 marks)]

(a) The data supplied below is the diameter (in cm) of bacteria colonies as measured by a microbiologist 12 hours after seeding.

2.1	3.4	3.9	4.7	3.7	0.8	3.6	4.1	4.9	2.5	3.1	1.5
2.6	4.0	1.3	3.5	0.9	1.5	4.2	3.5	2.1	3.0	1.7	3.6
2.8	3.2	3.3	0.4								

(i) Find the median

(ii) Group the data in 5 groups and display it using a frequency histogram.

(iii) Comment on the skewness of the data.

(b) Prove that the variance of any set of five consecutive integers is 2.

(c) Find the mean of the data set $\{x_1, x_2, ..., x_{25}\}$ given that $\sum_{i=1}^{25} x_i^2 = 2568.25$ and the standard deviation is 5.2.

QUESTION 2 [5 + 8 + 7 = 20 Marks]

A sociologist collected the following data regarding the age of a wife her husband when they were married.

Wife's age	19	42	28	25	36
Husband's age	20	32	31	24	33

(a) Plot a scatter diagram for these data with the wife's age on the horizontal axis.

(b) Determine an appropriate regression line for this distribution and fit in (a) above.

(c) Compute and interpret the correlation coefficient.

QUESTION 3 [3 + (5 + 6 + 5) + 8 = 27 Marks]

(a) Given that Z is the standard normal random variable, find Z if the area to left of Z is 0.67

(b) The lifetime of a type of bearing is normally distributed with average life of 1,600 hours and standard deviation of 35 hours.

    (i)   What percentage of the bearings are likely to fail before 1,550 hours?

    (ii)   Determine the number of bearings in a batch of 1,000 that are likely to last more than 1,675 hours.

    (iii)   What lifetime is the 67th percentile distribution?

(c) Find the mean and standard deviation of a normally distributed random variable X if $P(X \leq 20) = 0.1$ and $P(X \geq 29) = 0.15$

QUESTION 4 [3 + 3 + 4 + 4 = 14 Marks]

In a group of 40 boys, 23 have dark hair, 18 have brown eyes, and 26 have dark hair, brown eyes or both. One of the boys is selected at random. With an aid of a Venn diagram determine the probability that he has:

(a) Dark hair and brown eyes

(b) Neither dark hair nor brown eyes

(c) Dark hair but no brown eyes

(d) Brown eyes given that he has dark hair.

QUESTION 5 [5 + (4 + 4 + 4) = 17 Marks]

(a) 5% of a batch of batteries are defective. A random sample of 80 batteries is taken with replacement. Find the mean and standard deviation of the defective batteries in the sample.

(b) During a season, a basketball player has a 94% success rate in shooting from the free throw line. In one match the basketballer has 20 shots from the free throw line.

  (i)   Find the probability that the basket baller is successful on all 20 throws.

  (ii) Find the probability that the basket baller is successful on at least 18 throws.

  (iii) Find the expected number of successful throws for the match.

FORMULA SHEET

Grouped Data

1. $Mean = \frac{\sum FX}{\sum f} = \frac{\sum fx}{N}$

2. $Median\ Estimate = LL + W'(\frac{N}{2} - CB)/F$

3. $Mode\ Estimate = LL + Wd_1/(d_1 + d_2)$

4. $Quartile\ 1 = LL + W\left(\frac{N}{4} - CB\right)/F$

5. $Quartile\ 3 = LL + W\left(\frac{3N}{4} - CB\right)/F$

6. $Variance = \sigma^2 = \frac{(\sum FX^2)}{N} - \left(\frac{\sum FX}{N}\right)^2$

Ungrouped Data

7. $Variance = \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \mu^2$

8. $Standard\ deviation = \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$

Regression and Correlation

9. $r = \dfrac{N\sum XY - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{\left[N\sum X^2 - \left(\sum X\right)^2\right] * \left[N\sum Y^2 - \left(\sum Y\right)^2\right]}}$

10. $b = \dfrac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2}$

11. $a = \bar{y} - b\bar{x}$

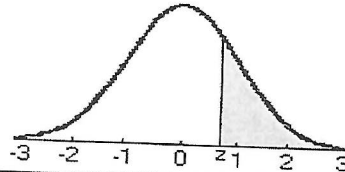## Normal Distribution

12. $Confidence\ Interval = \bar{x} \pm Z_{\frac{\alpha}{2}} \times \delta$

NORMAL DISTRIBUTION GRAPH

AREAS UNDER THE STANDARD NORMAL CURVE

The following table gives the area
indicated in the sketch on the right

z	0	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1921	.1894	.1867
.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	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0014	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0005	.0005
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000