BACHELOR OF INFORMATION TECHNOLOGY (HONS)

FINAL EXAMINATION
MARCH 2021

Course : EB3125 (Statistics)

Time: 8.00am – 11.00am
(3 hours)

Lecturer : Nur Suaidah Rosli

Date : 27 July 2021

Instructions:

Answer **ALL** questions.

This examination paper is confidential. The questions must be answered individually. Students are NOT PERMITTED to discuss or consult with other students or individuals.

Using Google is not allowed at all. Thus, the answer from the Internet will be considered plagiarism. Plagiarism is an offence. University guidelines on plagiarism will apply.

All exams submitted are final. Students will NOT BE PERMITTED to submit any additional work or alternative version, even if time is remaining. Only the initial submission will be forwarded for grading.

Your answer MUST be submitted within the stipulated time. Failure to submit your answers within the deadline given may result in the award of zero marks. You will be given an additional 30 minutes at the end of the specified exam duration. This extra time is for you to submit your completed exam. It's not intended as extra working time. If you experience technical difficulties, you can use this time at your own discretion, but you must leave sufficient time to submit your completed exam.

Answer Format:

i. Do not put your name on any materials related to the exam. Use only your Student ID Number for identification.
ii. All answers must be handwritten, scanned (using Camscanner) and converted to PDF file.
iii. Save your answers in the following format: **STUDENT ID_COURSE CODE_COURSE TITLE**
iv. You are required to cc your email to BIT_PIC1@nilai.edu.my.

Honor Pledge for Exams

"I affirm that I have not given or received any unauthorised help on this exam, and that all work is my own."

Name and Signature: _____

This question paper consists of 7 pages.
(excluding front cover)

1. a. Explain whether each of the following constitutes a population or a sample.

   i. Number of personal fouls committed by all NBA players during the 2008–2009 seasons. **(1 MARK)**

   ii Yield of potatoes per acre for 10 pieces of land. **(1 MARK)**

   iii. Weekly salaries of all employees of a company. **(1 MARK)**

   iv. Number of computers sold during the past week at all computer stores in Mesa mall. **(1 MARK)**

   b. Cytosol levels of 40 randomly selected people attended to a particular clinic are given below

172	131	175	163	176	142	163	127	131	132
130	151	155	127	152	152	162	124	137	137
162	151	174	142	150	157	152	154	150	155
147	160	156	145	147	144	171	162	165	140

   i. Classify the data above using the class intervals 120 – 129, 130 – 139, .. . **(3 MARKS)**

   ii. Draw a histogram for the data. **(3 MARKS)**

   iii. Find mode and mean of the data. **(4 MARKS)**

   iv. Find the sample variance and standard deviation of the data. **(6 MARKS)**

2. a. Define mutually exclusive events and give an example. **(2 MARKS)**

b. Each salesperson at Stiles-Compton is rated either below average, average or above average with respect to sales ability. Each salesperson is also rated with respect to his or her potential for advancement – either fair, good or excellent. The traits for 500 sales people were cross-classified into the following table

	Fair	Good	Excellent
Below average	16	12	22
Average	45	60	45
Above average	93	72	135

i. What is the probability a salesperson selected at random will have above average and excellent potential for advancement? **(3 MARKS)**

ii. What is the probability a salesperson selected at random will have above average or excellent potential for advancement? **(3 MARKS)**

iii. Find the probability a salesperson selected at random will have below average given that he or she from good potential for advancement. **(4 MARKS)**

b. A manufacturer of window frames knows from long experience that 5% of the production will have some type of minor defect that require an adjustment. What is the probability that in a sample of 10 window frames:

(2 MARKS)

i. none will need adjustment.

ii. at least one will need adjustment. **(2 MARKS)**

iii. compute the mean and standard deviation of the distribution. **(4 MARKS)**

3. a. The following table lists the probability distribution for cash prizes in a lottery conducted at Lawson's Department Store.

Prizes RM	Probability
0	0.45
10	0.30
100	0.20
500	0.05

If James buys a single ticket, what is the probability that James win:

i. exactly RM100. **(1 MARK)**

ii. at least RM 10 **(1 MARK)**

iii. no more than RM100. **(1 MARK)**

iv. compute the mean and standard deviation **(3 MARKS)**

b. Recent crime reports indicate that 3.1 motor vehicle thefts occur each minute in Malaysia. Assume that the distribution of the thefts per minute can be approximated by the Poisson probability distribution.

i. Calculate the probability exactly four thefts occur in a minute. **(2 MARKS)**

ii. What is the probability there are no thefts in a minute. **(2 MARKS)**

iii. Find the probability there is at least one theft in a minute. **(3 MARKS)**

c. WNAE, an all –news FM station, finds that the distribution of the lengths of time listeners are tuned to the station follows the normal distribution. The mean of the distribution is 15.0 minutes and standard deviation is 3.5 minutes. What is the probability that a particular listener will tune in:

i. more than 20 minutes. **(2 MARKS)**

ii. for less than 10 minutes. **(2 MARKS)**

iii. between 10 to 12 minutes. **(3 MARKS)**

4. a. Suppose a small finite population consists of only $N = 4$ numbers:

4 5 8 9

i. What is the population mean? **(2 MARKS)**

ii. List all the possible samples of size five (without replacement) that can be selected from this population. **(4 MARKS)**

iii. Determine the sampling distribution of the mean for random samples of size 2. Hence calculate the mean of the sampling distribution. **(4 MARKS)**

b. AC Nielsen reported that children between age 2 and 5 watch an average of 25 hours television per week. Assume the variable is normally distributed and the standard deviation is 3 hours. If 20 children between the ages of 2 and 5 are randomly selected, find the probability that the mean number of hours they watch television will be greater than 27 hours? **(5 MARKS)**

$$\left[\text{Hint}: z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}\right]$$

c. Samples of Alzheimer's patients are tested to access the amount of time in stage IV sleep. It has been hypothesized that individual suffering from Alzheimer may spend less time per night in the deeper stages of sleep. Number of minutes spent in stage IV sleep is recorded for fifty patients. The sample produced the mean of 45 minutes with standard deviation of 12 minutes of stage IV sleep over a 24 hour period of time. Compute a 95% confidence interval for this data. **(5 MARKS)**

$$\left[\text{Hint}: \overline{X} \pm z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right),\ 95\%\,\text{Confidence Interval} = 1.96\right]$$

5. a. Define Type I and Type II errors related to a hypothesis test. **(2 MARKS)**

b. It is believed that Arsenic (As) content in a particular product is higher than the maximum allowable level (k) specified by the World Health Organization. Assume that you want to test the validity of this belief. Write down the null and alternative hypothesis related to this test. **(2 MARKS)**

c. The following data have been collected from 10 basketball players. It is believed that taller players are better basketball players. Scatter plot and some summary statistics of the data are also given.

Player	Height (cm)	Goals (y)
1	175	11
2	178	15
3	173	12
4	183	17
5	180	15
6	188	19
7	180	16
8	185	18
9	175	11
10	178	15

i. Draw and comment the scatter plot for the data given. **(4 MARKS)**

ii. Calculate the correlation coefficient and interpret. **(5 MARKS)**

iii. Estimate a regression line to predict the number of goals scored based on height of the player. **(7 MARKS)**

-END OF QUESTION PAPER-

FORMULAE

UNGROUPED DATA	GROUPED DATA
Mean, $\bar{x} = \dfrac{\sum f}{n}$ Sample Variance, $s^2 = \dfrac{\sum x^2 - \dfrac{\left(\sum x\right)^2}{n}}{n-1}$ Standard deviation, $s = \sqrt{s^2}$	Mean, $\overline{x} = \frac{\sum f x_m}{n}$ Median, MD $= L + \left(\dfrac{\frac{n}{2} - C_f}{f} \right) C$ Sample Variance , $s^2 = \dfrac{\sum X_m{}^2 f - \dfrac{\left(\sum X_m f\right)^2}{n}}{n-1}$ Standard deviation, $s = \sqrt{s^2}$

PROBABILITY	CONDITIONAL PROBABILITY

Complement Rule
$P\left(\overline{A}\right) = 1 - P(A)$
 	
A and B are mutually exclusive events
$P(A \cup B) = P(A) + P(B)$
 	
A and B are mutually non exclusive events
$P\left(A \cup B\right) = P(A) + P(B) - P(A \cap B)$
	P (A | B) = $\dfrac{P\left(A \cap B\right)}{P(B)}$
	
	If A and B are independent, then
	$P(A \cap B) = P(A) \cdot P(B)$
	
	If A and B are dependent, then
	$P(A \cap B) = P(B) \cdot P(B/A)$

BINOMIAL DISTRIBUTION	DISCRETE RANDOM VARIABLE
$P(X = x) = {}^{n}C_{x}\,(p)^{x}\,(q)^{n-x} = \dfrac{n!}{x!(n-x)!}\,p^{x}(1-p)^{n-x}$ $\mu = np\,;\qquad \sigma^2 = np(1-p)$	$E(X) = \mu = \sum X \cdot P(X)$ $V(X) = \sigma^2 = \sum\left[X^2 \cdot P(X)\right] - \mu^2$

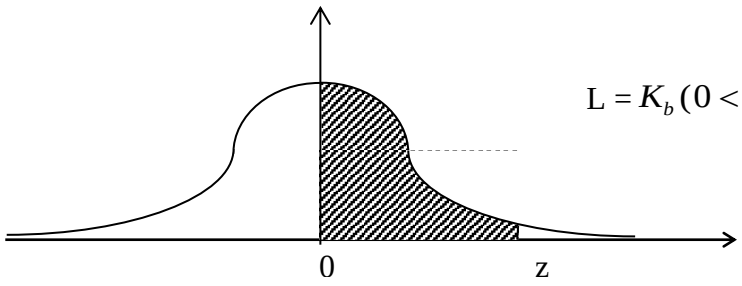
STANDARD NORMAL DISTRIBUTION	POISSON DISTRIBUTION
$z = \dfrac{x - \mu}{\sigma}$	P (X = x) = $\dfrac{e^{-\lambda}\,\lambda^{x}}{x!}$

EQUATION OF THE REGRESSION LINE	CORRELATION COEFFICIENT
$y' = a + bx$ $a = \dfrac{\left(\sum y\right)\left(\sum x^2\right) - \left(\sum x\right)\left(\sum xy\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2}$ $b = \dfrac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{n\left(\sum x^2\right) - \left(\sum x\right)^2}$	$r = \dfrac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right) - \left(\sum x\right)^2\right]\left[n\left(\sum y^2\right) - \left(\sum y\right)^2\right]}}$

POPULATION MEAN	SAMPLE MEAN
$\mu = \sum X \cdot P(X)$ $\sigma^2 = \sum\left[X^2 \cdot P(X)\right] - \mu^2$ $\sigma = \sqrt{\sigma^2}$	$\mu_{\bar{x}} = \sum \overline{X} \cdot P\left(\overline{X}\right)$ $\sigma_{\bar{x}}{}^2 = \sum\left[\overline{X}^2 \cdot P\left(\overline{X}\right)\right] - \mu_{\bar{x}}{}^2$ $\sigma_{\bar{x}} = \sqrt{\sigma_{\bar{x}}{}^2}$

THE STANDARD NORMAL DISTRIBUTION TABLE

Area under the standard normal curve from 0 to z



$$L = K_b(0 < Z \le z) = \Phi(z) = \int_0^z \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}\, dz$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	**0.0000**	**0.0040**	**0.0080**	**0.0120**	**0.0160**	**0.0199**	**0.0239**	**0.0279**	**0.0319**	**0.0359**
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0635	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	**0.3413**	**0.3438**	**0.3461**	**0.3485**	**0.3508**	**0.3531**	**0.3554**	**0.3577**	**0.3599**	**0.3621**
1.1	0.3643	0.3665	03686	0.3708	0.3729	0.3750	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	.0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4516	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	**0.4772**	**0.4778**	**0.4783**	**0.4788**	**0.4793**	**0.4798**	**0.4803**	**0.4808**	**0.4812**	**0.4817**
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3.	0.4893	0.4896	0.4998	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4974	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	**0.4987**	**0.4987**	**0.4987**	**0.4988**	**0.4988**	**0.4989**	**0.4989**	**0.4989**	**0.4990**	**0.4990**
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4996
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4988	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000