Name:

Enrolment No:



UNIVERSITY WITH A PURPOSE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES End Semester Examination, July 2020

SECTION A

Course: Applied Statistical Analysis Program: BTech – CS (BAO) Course Code: CSBA 1002

Semester: II Time 03 hrs. Max. Marks: 60

Instructions: All questions are compulsory.

	SECTION A	1	
S. No.		Marks	CO
Q 1	 You are being asked to collect and plot the net value (wealth/money) of each individual in this country (India). As you plot the data what kind of distribution should you expect? a) Exponential Distribution b) Normal Distribution c) Fat-tailed Distribution d) Poisson Distribution 	1	C01
Q 2	Arithmetic mean is affected by both change of scale and origin. a) True b) False	1	CO1
Q 3	Find the mean deviation from the mean for the following data: x 1 2 3 4 5 6 7 f 3 5 8 12 10 7 5 a) 1.348 b) 1.897 a) 2.098 b) 2.678	1	CO1
Q 4	Suppose a series of 100 observations has mean 50 and variance 20. Another series of 200 observations has mean 80 and variance 40. What is the combined variance of the given series? a) 233.33 b) 179.45 a) 398.56 b) 258.98	1	C01
Q 5	 Suppose batsman A has mean 50 with SD 10. Batsman B has mean 30 with SD 3. What do you infer about their performance? a) A is a better run maker. b) B is a better run maker. a) A is more consistent. 	1	C01

	b) B is more consistent.		
Q 6	 First four moments about mean of a distribution are 0, 2.5, 0.7 and 18.75. Find coefficient of skewness. a) 0.031 b) 0.458 c) 0.167 d) 0.246 	1	CO1
Q 7	 Karl Pearson's coefficient of skewness is 1.28, its mean is 164 and mode 100, find the standard deviation. a) 50 b) 25 c) 75 d) 63 	1	CO1
Q 8	If the value of mean, median and mode are same in any distribution, then the skewness does not exist in that distribution. Larger the difference in these values, larger the skewness. a) True b) False	1	CO1
Q 9	Fit a straight line $Y = a + bX$ to the following data: x 6 7 8 9 11 y 5 4 3 2 1 a) $a = 9.6486$ b) $a = 2.5891$ c) $a = 10.986$ d) $a = 7.9165$	1	CO1
Q 10	Fit a straight line $Y = a + bX$ to the following data: x 6 7 8 9 11 y 5 4 3 2 1 a) b = -0.8108 b) b = 0.5312 c) b = -2.915 d) b = 1.789	1	CO1
Q 11	The regression line of y on x is: $(y - \overline{y}) = \frac{r\sigma_y}{\sigma_x} (x - \overline{x})$ a) True b) False	1	CO1
Q 12	An unbiased coin is tossed six times. Find the probability of obtaining at least 3 heads a) 21/32	1	CO1

	b) 16/32		
	c) 23/32		
	d) 19/32		
Q 13	The chances of catching cold by workers working in an ice		
X 10	factory during winter are 25%. What is the probability that out of 5 workers 4		
	or more will catch cold?		
	a) 0.015627	1	CO1
	b) 0.146789		
	c) 0.176549		
	d) 0.0089643		
Q 14	Let X be a random variable having Bernoulli distribution with mean $p = 0.4$. Find its		
	variance.		
	a) 0.24	1	CO1
	b) 0.4	1	CO1
	c) 0.17		
	d) 0.48		
Q 15	A policeman fires 6 bullets on a dacoit. The probability that the dacoit will		
	be killed by a bullet is 0.6. What is the probability that the dacoit is still		
	alive?		
	a) 0.0041	1	CO1
	b) 0.0098		
	c) 0.0037		
	d) 0.0056		
Q 16	Out of 800 families with 4 children each, how many families would you		
	expect to have 3 boys and 1 girl, assuming equal probability of boys and		
	girls?		
	a) 32	1	CO2
	b) 16		
	c) 28		
0.15	d) 31		
Q 17	It is known that the number of heavy trucks arriving at a railway		
	station follows the Poisson distribution. If the average number of truck arrivals		
	during a specified period of an hour is 2, find the probabilities that during a		
	given hour no heavy truck arrive.	1	CO2
	a) 0.1353		
	b) 0.16753		
	c) 0.1176		
0.19	d) 0.12576		
Q 18	If the probability that an individual suffers a bad reaction from an injection of a given serum is 0.001, determine the probability that out of 500		
	individuals, 3 individuals suffer from bad reaction.		
	a) 0.1255	1	CO2
	a) 0.1255 b) 0.1567	T	
	c) 0.1587		
	d) 0.1599		
	uj 0.1377		

Q 26	In a particular branch of a bank, it is noted that the duration/waiting time of the customers for being served by the teller is normally distributed	1	CO2
Q 25	The mean and standard deviation of a Poisson distribution are 6 and 2 respectively. Test the validity of this statement. a) Valid b) Invalid	1	CO2
Q 24	In a university the mean weight of 1000 male students is 60 kg and standard deviation is 16 kg. What is the lowest weight of the 100 heaviest male students? (Assuming that the weights are normally distributed) a) 80.48 b) 60.45 c) 74.76 d) 59.78	1	CO2
Q 23	In a university the mean weight of 1000 male students is 60 kg and standard deviation is 16 kg. Find the number of male students having their weights more than 70. (Assuming that the weights are normally distributed) a) 264 b) 235 c) 247 d) 214	1	CO2
Q 22	If the r.v. X is normally distributed with mean 80 and standard deviation 5, then find P[64 < X <76] a) 0.2112 b) 0.3541 c) 0.2222 d) 0.2562	1	CO2
Q 21	Moment coefficient of skewness is zero and the curve is always platykurtic. a) True b) False	1	CO2
Q 20	b) 23 d) 50 Below, there is given the p.d.f. of a normally distributed random variable. Obtain the parameters (mean and variance) of the variable. $f(x) = \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{32}(x-60)^2}, -\infty < x < \infty$ a) Mean = 60, Variance = 16 b) Mean = 60, Variance = 4 c) Mean = 16, Variance = 60 d) Mean = 4, Variance = 60	1	CO2
Q 19	If an unbiased die is thrown 120 times, find the expected frequency of appearing 5 on the die. a) 20 b) 15 c) 25	1	CO2

	with mean 5.5 minutes and standard deviation 0.6 minutes. Find the		
	probability that a customer has to wait between 4.2 and 4.5 minutes.		
	a) 0.0325		
	b) 0.0461		
	c) 0.2567		
	d) 0.3975		
Q 27	Suppose that temperature of a particular city in the month of March is normally distributed with mean 24 degree C and standard deviation 6 degree C. Find the probability that temperature of the city on a day of the month of March is more than 26 degree C. a) 0.3707 b) 0.3156 c) 0.2875 d) 0.2931	1	CO2
Q 28	We want to estimate the average coffee intake of UPES students, measured in cups of		
-	coffee. A survey of 1,000 students yields an average of 0.55 cups per day, with a		
	standard deviation of 1 cup per day. Which of the following is not necessarily true ?		
	a) Population mean = 0.55, Population S.D. = 1	1	CO2
	b) Sample mean = 0.55, Sample S.D. = 1		
	c) The sample distribution is right skewed.		
	d) 0.55 is the point estimate for the population mean.		
	d) 0.00 is the point estimate for the population mean.		
Q 29	Which of the following is false?a) In order to reduce the standard error by half, sample size should be		
	doubled.		
	b) Standard error measures the variability in means of samples of the same size		
	taken from the same population.		
	c) As the sample size increases, the variability of the sampling distribution		
	decreases.		
	d) No false statement.		
Q 30	Suppose you took a large number of random samples of size n from a large population and calculated the mean of each sample. Then suppose you plotted the distribution of your sample means in a histogram. Now consider the following possible attributes of your collected data and the population from which they were sampled. For which of the following sets of attributes would you not expect your histogram of your sample means to follow a nearly normal distribution?		
	a) N=10. The population distribution is unknown, but the distribution of		
	data in each sample is heavily skewed.		
	b) N=120. The population distribution is unknown, but the distribution of data		
	in each sample is slightly skewed.		
	c) $N=20$. The population distribution is nearly normal.		
	d) N=120. The population distribution is slightly skewed.		
Q 31	Researchers investigating characteristics of gifted children collected data from schools		
	ד הפשבמיטווביש ווועבשנוקמנוווע טומומטנבוושנוש טו עוונבע טווועובוו טטוובטנבע עמנמ ווטווו שנווטטוא		
	in a large city on a random sample of thirty-six children who were identified as gifted		

	children soon after they reached the age of four. The following histogram shows the	
	distribution of the ages (in months) at which these children first counted to 10	
	successfully. Also provided are some sample statistics.	
	Calculate a 90% confidence interval for the average age at which gifted children first	
	count to 10 successfully. Choose the closest answer.	
	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	
	Age when the child first counted to 10 successfully (in months)	
	 a) (29.50, 31.88) b) (30.12, 31.26) c) (30.49, 30.89) 	
	d) (29.28, 32.10)	
Q 32	Suppose we wanted to compare the rates of return for two stocks: the technology company Intel and the U.S. airline Southwest Airlines. To compare the rates of return, we take a random sample of 50 days of Intel's stock returns and another random sample of 50 days for Southwest's stock returns (not necessarily the same days). These data should not be treated as paired. Why would these data not be considered paired data?	
	a) The dates for both the stock should be different.	
	b) Companies are in different industries.	
	 c) 50 observations is not enough. d) When rendem sampling is involved, data can't be treated as pairs. 	
Q 33	 d) When random sampling is involved, data can't be treated as pairs. How does the shape of the t-distribution change as the sample size increases? 	
	 a) It becomes more normal looking. b) It becomes wider. 	
	c) It becomes skewed.	
Q 34	d) It becomes fatter.Air quality measurements were collected in a random sample of 25 country capitals in	
	2013, and then again in the same cities in 2014. We would like to use these data to	
	compare average air quality between the two years. Which of the following tests is the most appropriate?	
	a) Paired t-test with two-sided alternative hypothesis	
	b) Paired t-test with one-sided alternative hypothesisc) Independent samples t-test with one-sided alternative hypothesis.	
	d) Independent samples t-test with two-sided alternative hypothesis.	

Q 35	• 1	othesis is a hypot tion that it is true	hesis which is tes	sted for poss	ible reject	ion under		
	a) True							
0.26	b) Fals		ng smoking and lu		ompared p	ationts		
Q 36	hospitalized other reason testing when than for pati	l with lung cancer ns), and recorded ther the proportio tents without lung	to similar patients whether each pa n of smokers is hi cancer, the p-val that smoking caus	s without lung tient smoked gher for the p ue is less that	g cancer (h l. For a hyp patients wi an 0.00000	nospitalized for pothesis test th lung cancer		
			smoking a	status				
				n-smoker	total			
		lung cancer	647	2	649			
	cancer	not sure	622	27	649			
		total	1269	29	1298			
Q 37	c) No, d) Non 80% of Ame	king causes lung with the given p-v e of the above ericans start the c	-value we would r cancer. value we would fa lay with a cereal b tement is true or f	il to reject Ho preakfast. Ba	o in favour	of Ha.	hat	
	-	-	the proportions of amples of size 40			he day with a		
	a) Fals b) True							
Q 38	The chi-squ a) Tru e b) Fals	e	lways non-negativ	/e.				
Q 39	communicat to collect yo your dorm a	ting with others o our own sample fo and find that on av	erage college stud nline. You believe or a hypothesis tes verage they spent ollowing is the corr	that this is a st. You rando 3.5 hours a	in underes omly sampl week com	timate and dec e 60 students f municating with	rom 1	
	b) Ho : c) Ho : d) Ho :	population mean population mean sample mean =	an = 2; Ha : popu n = 2; Ha : populat n = 3.5; Ha : popul 2; Ha : sample me 2; Ha : sample me	tion mean < 2 lation mean < ean > 2	2			

Q 40	A researcher found a 2006 - 2010 survey showing that the average age of women at first marriage is 23.44. Suppose a researcher believes that this value may have increased more recently, but as a good scientist he also wants to consider the possibility that the average age may have decreased. The researcher has set up his hypothesis test; which of the following states the appropriate Ha correctly? a) Ha not equals 23.44 years old b) Ha equals 23.44 years old c) Ha less than 23.44 years old d) Ha ergetates the appropriate Ha correctly?	
Q 41	 d) Ha greater than 23.44 years old A manufacturer of ball point pens claims that a certain pen manufactured by him has a mean writing-life at least 460 A-4 size pages. A purchasing agent selects a sample of 100 pens and put them on the test. The mean writing-life of the sample found 453 A-4 size pages with standard deviation 25 A-4 size pages. Should the purchasing agent reject the manufacturer's claim at 1% level of significance? a) Yes b) No 	
Q 42	A researcher wants to measure physical height in as much detail as possible. Which level of measurement does s/he employ?	
Q 43	You conduct a study on guitar color and you question 550 people. 110 of them have brown guitar and 44% of them have blue guitar. What percentage of the people you questioned has blue or brown guitar? a) 64 b) 45 c) 67 d) 82	
Q 44	Ten students completed an exam. Their scores were: 5, 7, 2, 1, 3, 4, 8, 8, 6, 6. What is the interquartile range (IQR)? a) 8 b) 5 c) 4 d) 5.5	
Q 45	A researcher wants to know what people in Amsterdam think of football. He asks ten people to rate their attitude towards football on a scale from 0 (don't like football at all) to 10 (like football a lot). The ratings are as follows: 1, 10, 6, 9, 2, 5, 6, 6, 5, 10. What is the standard deviation? a) 3,1 b) 9,2 c) 6,0 d) 9,3	

Q 46	Which of the following statement(s) is/are correct?		
X IV	which of the following statement(s) is/are correct?		
	I. If you conduct a significance test you assume that the alternative hypothesis is true		
	unless the data provide strong evidence against it.		
	II. The null hypothesis and the alternative hypothesis are always mutally exclusive.		
	a) I correct, II incorrectb) Both correctc) Both false		
Q 47	 d) Il correct, I incorrect You are interested in the question how long Dutch scuba-divers can stay underwater. 		
Q +7	You heard that the average number of minutes Dutch scuba-divers can stay underwater. You heard that the average number of minutes Dutch scuba-divers can dive without coming to the surface equals 68 minutes. You don't believe that's true and decide to conduct a test. Your null hypothesis is that the population mean equals 68. Your alternative hypothesis is that the population mean differs from 68. You question 40 randomly selected Dutch scuba-divers and discover that they can stay underwater for only 64 minutes. The standard deviation is 3 minutes. Do you reject your null hypothesis?		
	a) No b) Yes		
Q 48	You want to know how many hours of sleep new parents lose after they had their first		
	baby. You know that the population mean equals 2.3 hours. Because you can't		
	investigate the whole population, you take a sample of 100 people. You find an average		
	sleep loss of 2.1 hours. What is, based on this sample, the point estimate of your		
	population mean?		
	A)2.1 b)0.2 c)5.4 d)2.3		
Q 49	The following statements are about confidence intervals for proportions. Place in order from smallest to largest z-score.		
	(1) 99% confidence interval		
	(2) z = 1.645		
	(3) point estimate ± 1.96 * SE		
	a) 2,3,1 b) 1,3,2 c) 3,2,1 d) 2,1,3		
Q 50	Imagine you want to know the length of the beard of every male student in India. You know that the population mean equals 2.2 millimeters and the population standard		
		1	

	deviation equals 0.9 millimeters. What will be the mean (in millimeters) of the sampling distribution of the sample mean?		
	 a) 2.2 b) 3.1 c) 1.3 d) 0.3 		
Q 51	You know that twenty percent of the people in Dehradun describe themselves as Hipsters. You ask 400 respondents if they identify as a Hipster or not. What is the standard deviation of the sampling distribution of the sample proportion?		
	 a) 0.02 b) 0.2 c) 0.04 d) 0.1 		
Q 52	The ice cream shop has problems with the delivery of the different flavours. As a consequence the shop doesn't have the same amount of flavours every day. In the following list you see the probability distribution of the different amounts of flavours.		
	Amount of flavours (probability)		
	4 (0.14)		
	5 (0.35)		
	6 (0.31)		
	7 (0.20)		
	What is the mean amount of flavours the ice cream shop sells?		
	 a) 5.57 b) 4.25 c) 8.51 d) 2.98 		
Q 53	You have a random variable X with variance 3. Now you multiply X with 2. What becomes the variance of X ?		
	 a) 12 b) 3 c) 7 d) 6 		
Q 54	Imagine you're investigating the time people wait at traffic lights, a variable which appears to be approximately normally distributed with a mean of 1.3 minutes and a standard deviation of 0.57 minutes. Which of the following intervals contains 95% of the waiting times?		
i		<u> </u>	L

	a) 0.16 and 2.44	
	c) 1.3 and 2.44	
	 d) 0.16 and 1.3 e) 0.73 and 1.87 	
Q 55	You investigate the earnings of the 2nd year students in your school. They earn on	
Q 33	average Rs240, with a standard deviation of Rs90, One person stands out, because	
	she's a snooker champion. She makes on average Rs420, a week. What is the	
	corresponding z-score of her earnings?	
	2	
	1	
	3	
	5	
Q 56	On average, a proportion of 0.48 newborns are girls. What are the chances that in a	
	family with 4 children there are exactly three daughters? a) 0.23	
	b) 0.46	
	c) 0.33	
Q 57	d) 0.48	
Q J/		
	0.25	
	<i>≩</i>	
	0.15 D	
	8. –	
	8	
	0 5 10 15 20	
	x	
	n	

	Looking at the binomial distribution above, what would be reasonable values for the parameters of this distribution?	
	a) Number of trails= 2, probability of success = 0.29	
	b) Number of trails= 20, probability of success = 0.29	
	c) Number of trails= 2, probability of success = 0.1	
	d) Number of trails= 20, probability of success = 0.1	
Q 58	A multiple choice exam consists of 12 questions, each having 5 possible answers. To pass you must answer at least 8 out of 12 correctly. What are your chances of passing if you go into the exam without knowing a thing and resort to pure guessing?	
	 a) 0.0005 b) 0.005 c) 0.00005 d) 0.0025 	
Q 59	For a normally distributed variable with a mean of 10 and standard deviation of 5, what is the proportion of the data with negative values?	
	a) 0.025 b) 0.5	
	c) 0.25	
	d) 0.075	
Q 60	You collect four shells from the beach. You know that there are only three types of	
	shells on the beach, and these shells occur in equal amounts. How many different events are possible?	
	a) 81	
	b) 16	
	c) 27	
	d) 105	