CHUKA UNIVERSITY

THIRD YEAR EXAMINATION FOR THE AWARD OF BACHELOR OF SCIENCE, ART AND EDUCATION

MATH 343: APPLIED STATISTICS

STREAM: BSC, BED, BA   **TIME: 2 HOURS**

INSTRUCTION: Answer Question One and any other TWO Questions

QUESTION ONE (30 MARKS)

(a) Explain the following terms as used in statistical hypothesis
    (i)   Null and alternative hypothesis
    (ii)   Simple and composite hypothesis   [5marks]
    (iii)   Level of significance

(b) Alfafa (kind of plants grown as fodder for animal) yields of 6 test plots are 1.5, 1.9, 1.2, 1.4, 2.3 and 1.3 tons respectively per hectare. Use a critical region of $\alpha = 0.05$ to test the hypothesis $H_0: \mu = 1.8$ vs $H_1: \mu \neq 1.8$. Assume that the yields have a normal distribution   [6marks]

(c) The following data represent the change (in ml) in the amount of carbon monoxide transfer in smokers with chickenpox over a one week period:

<div align="center">

33   2   24   17   4   1   -6

</div>

Is there evidence of significant improvement in lung function
    (i)   If the data are normally distributed with alpha=10?   [4marks]
    (ii)   If the data are normally distributed with alpha unknown?   [3marks]

(d) The efficacy of a treatment for hypertension is to studied using a small clinical trial. 38 hypertensive patients were randomly allocated to either group 0 (Placebo control) or group 1(treatment) and a tree month follow-up study was carried out. At the end of the study the difference in blood pressure was measured for patients in each group and recorded. The summary of the results is presented below.

Group	Sample size	Mean	Variance
0	21	-0.208	4.101^2
1	17	3.953	4.630^2

Is there evidence of significant improvement in the treatment group? At 5% significance level   [6marks]

(e) A departmental store A has for competitors; B,C,D & E. Store A hires a consultant to determine if the percentage of shoppers who prefer each of the five stores is the same. A survey of 1100 randomly selected shoppers is conducted and the results about which one of the stores shoppers prefer are as shown below.

Store	A	B	C	D	E
No. of shoppers	262	234	204	190	210

Is there enough using a significant level of 5% to conclude that the proportions are really the same?   [6marks]

QUESTION TWO (20 MARKS)

(a) The following data is a sample of 11 loan applicants in a certain bank in Kenya. The applied loan is either approved or declined (status) against five risk variables namely age of applicants(X_1), applicants gender (X_2), amount of loan applied (X_3), applicant's salary (X_4) and proposed repayment period (X_5) as shown below.

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	.793	.246		3.225	.002
X_1	-.040	.066	-.043	-.607	.546
X_2	.186	.074	-.201	-2.511	.014
X_3	.200	.074	.203	2.691	.009
X_4	.685	.053	.883	12.992	.000
X_5	.345	.049	.526	17.523	.000
R Square=0.740:		Adjusted R Square=0.727;		F-Statistic=59.687(**.000**)	

Required

Write a report on Multiple regression function and interpret the results   [10 Marks]

(b) A study investigating the association between size of cars and country found the following frequency counts

	USA	JAPAN	UK	FRANCE
ECONOMY	21	24	33	55
COMPACT	27	35	37	40
FULL SIZE	36	11	12	4
LUXURY	15	3	7	8

Is there sufficient evidence of a significant relationship between size of car and country?
[10marks]

QUESTION THREE (20 MARKS)

(a) Patel is the manager of a bakery in Chuka town. He believes that the smell of fresh baking will encourage customers to purchase goods form his bakery. To investigate the this belief , he recorded the daily sales for ten weeks when all bakery windows are open and the daily sales for another ten weeks when all the windows are closed as shown below.

Windows Open	190.8	215.5	207.0	204.5	202.0
	185.7	204.1	187.8	208.8	215.6
Windows closed	205.4	177.6	199.4	192.2	193.5
	192.8	172.2	169.2	181.8	200.6

Assuming that these data may be deemed to be random samples from normal populations with same variance, investigate the bakers belief at 5% significance level. [10marks]

(b) The data below represents a sample of mathematics achievement test scores and calculate grades for 10 independently selected Chuka University students.

Math test score (X)	72	82	93	65	76	89	81	58	95	91
Final calculate grade (Y)	75	79	84	71	82	91	85	68	90	92

Required:

Test whether the achievement test scores and calculated grades are independent at 5% significance level. [10marks]

QUESTION FOUR (20 MARKS)

(a) The table of unit of fertilizer used and the units of yield in a science laboratory experience is as shown below.

Fertilizer	23	27	28	29	30	31	33	35	36	39
Yield	18	22	23	24	25	26	28	29	30	32

Required:

(i)   Determine the Pearson correlation coefficient between fertilizer(X) and Yield (Y)
[4marks]

(ii)   Using the results in b(i), test for the significance of the correlation coefficients at 5% significance level [6marks]

(b) Two random samples taken from two normal populations are as follows

sample I	20	16	26	27	23	22	18	24	25	19	
Sample II	17	23	32	25	22	24	28	18	31	33	20	27

Estimate the variances of the populations and test whether the two populations have equal variance at 5% level of significance. [10marks]

QUESTION FIVE (20 MARKS)

The data in the accompanying table relate mean yields of soybean plant obtained in response to the indicated levels of ozone exposure over the growing season.

X	Y
10	5
14	3
7	5
12	2
5	7
6	8

Required

i. Fit a simple linear regression model
ii. Compute the ANOVA
iii. Compute coefficient of determination and make comment

[20marks]

****

THIRD YEAR EXAMINATION FOR THE AWARD OF BACHELOR OF SCIENCE, ART AND EDUCATION

MATH 454: NON PARAMETRIC METHODS

STREAM: BSC, BED, BA   **TIME: 2 HOURS**

INSTRUCTION: Answer Question One and any other TWO Questions

QUESTION ONE (30 MARKS)

(a) The following is an arrangement of Boys (B) and Girls (G) selected haphazardly at a co-educational school

        GBBGGBGBBBGGGGBGBGBBGBGGGBBGBGGBGGBGBGBBG

Test for randomness at 5% level of significance   [7marks]

(b) The Sociology marks (X) and Statistics marks (Y) of 10 students are given below

X	64	67	71	85	90	55	85	63	74	80
Y	62	58	53	74	97	53	94	49	86	78

Test whether the rank correlation coefficient is significant at 5% level of significance
[8marks]

(c) The following data represents the grammar score made by an entering group of students at a private school

22 12	36	34	25	51	46	38	32	41	17	32
44 31	21	33	29	11	18	23	29	18	26	30

Test $H_o: \mu=35$ against $H_1: \mu<35$ at 5% level of significance using normal approximation
[6marks]

(d) Given three A's and Five B's: Find the probability of getting the highest number of runs
[4marks]

(e) Let $Y_1<Y_2<\ldots<Y_{20}$ denote the order statistics of a random sample of size 20 from a distribution of a continuous type.

    Compute $P(Y_{12}<\pi_{0.7}<Y_{15})$ [5marks]

QUESTION TWO (20 MARKS)

(a) In an evaluation of various methods of teaching, the following marks were awarded to 3 groups of students taught by different methods.

Group X	26	19	16	28	24	23	33	24	31	20
Group Y	23	30	33	34	21	20	19	29	18	29
Group W	30	25	28	28	33	31	26	16	31	19

Required

Use Kruskal-Wallis test at alpha=5% to test whether the methods are equally effective

[10marks]

(b) Let $Y_1 < Y_2 < \ldots < Y_{36}$ denote the order statistics of a random sample of size 36 from a distribution of a continuous type.

Compute $P(Y_{11} < \pi_{0.4} < Y_{19})$

[10marks]

QUESTION THREE (20 MARKS)

(a) The following table shows the number of tins of vanilla and chocolate sold over a 12 day period

Vanilla	4	11	5	7	10	13	12	5	9	6	2	1
Chocolate	19	4	6	8	18	17	15	3	3	16	14	0

Using the Mann-Whitney test, can we conclude that on the average the same number of tins is sold at 5% level of significance

(b) The weights of eighty cartons of cooking fat are given below.

19	50	57	25	61	42	26	33	46	45
63	31	80	36	29	56	38	69	83	40
52	17	35	65	28	63	72	29	56	57
22	45	53	44	78	47	86	55	66	48
41	64	38	43	13	58	55	32	52	46

Test the null hypothesis that the sample is random at 5% significance level [10marks]

QUESTION FOUR (20 MARKS)

(a) A movie producer is bringing out a new movie. In order to map out her advertising, she wants to determine whether the movie will appeal most to a particular age group or whether it will appeal equally to all age groups. The producer takes a random sample from persons attending a pre-reviewing show of the new movie and obtained the result in the table below.

Persons	Age-groups (in years)				
	Under 20	20-39	40– 59	60& over	Total
Liked the movie	320	80	110	200	710
Disliked the movie	50	15	70	60	195
Indifferent	30	5	20	40	95
Total	400	100	200	300	1,000

Required

Determine whether the movie will appeal most to a particular age group or whether it will appeal equally to all age groups at 5% significance level   [10marks]

(b) A quality control chart has been maintained for the weight of paint cans tabbed from a conveyor belt at fixed point in a production time. 16 weight obtained in order of time are as follows:

| 68.2 | 71.6 | 69.3 | 71.6 | 70.4 | 65.0 | 63.6 | 64.7 |
| 65.3 | 64.2 | 67.6 | 68.6 | 66.8 | 68.9 | 66.8 | 70.1 |

Required

Use the run test to determine whether the weight of the point cans on the conveyor belt deviated from randomness   [10marks]

QUESTION FIVE (20 MARKS)

(a) Using the Kolmogorov-Smirnov methods, test the hypothesis that the following values form a random sample from a normal distribution with mean of 2 and variance of 4. Use alpha=5%   [10marks]

| 2.72 | 3.84 | 0.88 | 5.72 | 5.48 | 3.12 | 0.10 | 2.48 | 1.70 | 0.52 |
| 2.64 | 3.64 | 3.40 | 1.80 | -0.52 | -0.12 | 2.30 | 3.10 | 1.04 | 1.02 |

(b) The following data represent the breaking strength of a certain kind of elastic material in pounds

| 163 | 165 | 160 | 189 | 161 | 171 | 158 | 151 | 169 | 163 |
| 139 | 162 | 172 | 165 | 148 | 166 | 172 | 163 | 187 | 173 |

Use the sign-test to test whether the average breaking strength is greater than 160.Use alpha=5%   [10marks]

****

CHUKA UNIVERSITY

THIRD YEAR EXAMINATION FOR THE AWARD OF BACHELOR OF SCIENCE, ART AND EDUCATION

MATH 342: QUALITY CONTROL METHODS

STREAM: BSC, BED, BA   **TIME: 2 HOURS**

INSTRUCTION: Answer Question One and any other TWO Questions

QUESTION ONE (30 MARKS)

(a) Briefly outline two approach to the management of quality in statistics   [4marks]

(b) Outline 3 approaches to lot sentencing in statistics   [3marks]

(c) Samples of size n=5 are taken from a manufacturing process at regular intervals. A quality characteristic is measured and X and S values calculated for each sample. After 30 subgroups, we have

$$\sum_{i=1}^{30} \bar{X}_i = 58395 \quad and \quad \sum_{i=1}^{30} S_i = 1516$$

Required
  i. Compute the control limits for $\bar{X}$ and S charts   [5marks]
  ii. Estimate the value of sigma assuming the process is operating in statistical control
                                                                                  [1mark]
  iii. Assuming that the distribution generated by process is approximately normal, what percentage of the product meets specifications of 2000 ±150   [4marks]

(d) Find the probability of acceptance in a single sampling plan with n=10 and c=5. Assuming the lot fraction defective is 5%.   [5marks]

(e) Summarize the acceptance samples procedure on a flow chart   [8marks]

QUESTION TWO (20 MARKS)

(a) A sample of 5 bars from each of the last 10 days is sent for a chemical analysis of the calorie content. The results are shown below. Does it appear that there are any days where the calorie count is out of control?

Sample	Calorie Count				
	1	2	3	4	5
1	426	406	418	431	432
2	421	422	415	412	411
3	425	420	406	409	414
4	424	419	402	400	417
5	421	408	423	410	421
6	427	417	408	418	422
7	422	417	426	435	426
8	419	417	412	415	417
9	417	432	417	416	422
10	420	422	421	415	422

Required

Develop an appropriate control chart ($\bar{X}$- and R- Chart) and analyze your findings.

[10marks]

(b) The Early Morning Delivery Service guarantees delivery of small packages by 10:30 A.M. Of course, some of the packages are not delivered by 10:30 A.M. For a sample of 200 packages delivered each of the last 15 working days, the following number of packages were delivered after the deadline:

4	9	14	2	13	9	5	9
3	4	3	3	8	4	3	

Determine
(i) the mean proportion of packages delivered after 10:30 A.M.
(ii) the control limits for the proportion of packages delivered after 10:30 A.M. Were any of the sampled days out of control?
(iii) If 10 packages out of 200 in the sample were delivered after 10:30 A.M. today, is this sample within the control limits? [10marks]

QUESTION THREE (20 MARKS)

(a) Outline Advantages of quality control [5marks]

(b) An X chart is used to control the mean of a quality characteristic. It is known that sigma=6.0 and n=4. The center line=200, UCL=209 and LCL=191. If the process mean shift to 188, find the probability that shift is detected on the first subsequent sample. [5marks]

(c) The data below represents the number of nonconformities per a 1000 meters in telephone cable. Assuming that assignable causes can be found for point that plot out of control.

Sample No	No. of Nonconformities	Sample No	No. of Nonconformities	Sample No	No. of Nonconformities
1	1	9	0	16	8
2	1	10	19	17	3
3	3	11	24	18	6
4	7	12	6	19	7
5	8	13	9	20	4
6	10	14	11	21	9
7	5	15	15	22	20
8	13				

Required
Obtain the control limits for 2500 meters of cable   [10marks]

QUESTION FOUR (20 MARKS)

(a) Outline the merits of acceptance sampling   [5marks]

(b) A double sampling plan, has parameters $n_1=50$, $c_1=2$, $n_2=100$ and $c_2=8$. Consider a lot with exactly 5% defective. Compute;
   (i) the probability of acceptance on the 1st sample   [5marks]
   (ii) the probability of acceptance on the 2nd sample   [8marks]
   (iii) the probability of acceptance   [2marks]

QUESTION FIVE (20 MARKS)
(a) Lahey Motors specializes in selling cars to buyers with a poor credit history. Listed below is the number of cars that were repossessed from Lahey customers because they did not meet the payment obligations over the last 36 months.

6	5	8	20	11	10	9	3	9
15	12	4	11	9	9	6	18	6
9	7	13	7	11	8	11	13	6
13	5	5	8	10	11	9	8	14

Required
Develop a c-bar chart for the number repossessed. Were there any months when the number was out of control? Write a brief report summarizing your findings   [10marks]

(b) Compute the C_{pk} measure of process capability for the following machine and interpret the findings. What value would you have obtained with the C_p measure? Machine Data:
USL = 80
LSL =50
Process δ= 5
Process μ=60
                    [10marks]
****