Stat 217
Assignment #4

One-Way ANOVA

A firm developing a new citrus-flavoured soft drink conducted an experiment to study customer preferences for the colour of the drink. Four colours were considered: colourless, pink, orange, and lime green. Twenty test localities, which were similar in sales potential and representative of the target market for this product, were selected. Each colour was randomly assigned to five localities for test marketing. The number of cases sold per 1000 population during the test period are recorded below:

Colourless	Pink	Orange	Lime Green
26.5	31.2	27.9	30.8
28.7	28.3	25.1	29.6
25.1	30.8	28.5	32.4
29.1	27.9	24.2	31.7
27.2	29.6	26.5	32.8

Are the sales the same for all colours of the drink? Test at the 5% significance level.

Minitab

Method 1

1. Enter the data into the MINITAB worksheet one sample per column- *ie.* The 5 values from the 'colourless' sample are entered in c1, the 5 values of the 'pink' sample in c2, and so forth.
2. From the MENU BAR, select STAT>ANOVA>ONEWAY(UNSTACKED)
   When the dialog box appears, list all the columns in which you entered data.
3. Note the table and other details are printed in the SESSION window.

Method 2

1. Enter all the sample values into one column of the worksheet. In a corresponding position in a second column, enter the number of the sample (1,2,3, etc.) from which each value came. For instance, if you entered all the 'colourless' observations, then pink, then orange, then lime green in column c1, c1 would contain 26.5, 28.7, 25.1, 29.1, 27.2, 31.2, 28.3, ... 32.8. You would enter into c2: 1, 1, 1, 1, 1, 2, 2,..., 4. NOTE: These entries in c2 are most efficiently accomplished with CALC>MAKE PATTERENED DATA>SIMPLE SET OF NUMBERS. In the dialog box, store patterned data in c2. FROM 1st VALUE box type **1**, TO LAST VALUE box type **4,** LIST EACH VALUE box type **5**, LIST THE WHOLE SEQUENCE box type **1.**
2. From the MENU BAR, select STAT>ANOVA>ONEWAY
   As the RESPONSE, specify the column in which you entered the observations, as the FACTOR specify the column in which you placed the sample number.
3. The output appears in the SESSION window.

1. A large consulting firm hires a West Coast university to provide an M.B.A program for its employees. The basic statistics course is taught at four locations of the firm. After completion of the course, standardized tests are given to the participating employees at each location. The results are:

	Location			
Observation	A	B	C	D
1	96	65	66	60
2	88	74	90	72
3	92	77	88	66
4	75	82	73	75
5	81	70	85	78
6	62	78	87	68
7	86	84	94	
8	86		74	
9	90			
10	85			

Can it be concluded that there is no difference in test results between location? Use $\alpha = 0.05$
{Fcalc = 4.130> 2.9680, Rho}

2. The sales research division of a large corporation is conducting research on sales methods for selling one of the products of the corporation. The division has designed a completely randomized one-factor analysis of variance model to investigate the efficiency of three sales methods. The responses are measured in units of $100 sales, and are listed below.

	Sales Method		
Response	A	B	C
1	20	13	31
2	23	18	28
3	22	16	48
4	22	29	28
5	36	14	30
6	45		25
7	21		
8	26		

Are there differences between the three different sales methods? Use $\alpha = 0.10$. What assumptions did you make? {Fcalc= 3.899> 2.668, Rho, Sales are normally distributed and share a common variance}

3. A county employs 3 assessors who are responsible for determining the values of residential property in the county. To see whether or not these assessors differ in their appraisal, 5 houses were selected and each assessor determined the market value of each house. The data (assessors are the treatment since your main concern is to see if there is difference in appraisals) was then analyzed using a two-way ANOVA routine giving the following (partially completed) ANOVA table.

(a) Complete the ANOVA table

Source	SS	DF	MS	F
Treatment	45.9			
Block	141.7			
Error				
Total	250.8			

(b) Is there any indication of a difference between appraisors? Use a 5% significance level for testing.
{Fcalc = 2.905< 4.459, Aho}

Two-Way ANOVA

The data in the following table represent the milliequivalents of sodium excreted by six subjects 2 hours after treatment with one of four diuretics assigned at random by a clinician over a 6-day period. Using

significance level of 0.05, analyze the data to determine whether or not there are any differences between patients and any differences in the effectiveness of diuretics.

	Treatments (Diuretics)			
Subjects	A	B	C	D
1	3.9	30.6	25.2	4.4
2	5.6	30.1	33.5	7.9
3	5.8	16.9	25.5	4.0
4	4.3	23.2	18.9	4.4
5	5.9	26.7	20.5	4.2
6	4.3	10.9	26.7	4.4

{F_{Bl} = 1.5758<2.90, Aho depending on rounding}
(F_{Tr} = 37.153>3.287, Rho}

Method

1. Enter all of the data into one column. You should first enter all the values for treatment A, then treatment B, and so forth.
2. In the second column, enter from which treatment each value came from *(ie.* You would enter A, A, A, A, A, A, B, B,..., D) These entries are most efficiently accomplished with
   CALC>MAKE PATTERNED DATA > TEXT VALUES
   Store patterned data in **C2**
   In the dialog box, type A B C D (Note: leave a space between each letter)
   In the list each value box, type **6**
   In the list the whole sequence box, type **1.**
3. In the third column, we want to enter the subject number (*ie.* 1,2,3,4,5,6,1,2,3,4,5,6,...
   Go to CALC>MAKE PATTERNED DATA>SIMPLE SET OF NUMBERS
   (a) Store patterned data in **C3**
   (b) From 1st value, type **1**
   (a) To last value, type **6**
   (b) In steps of , type **1**
   (c) List each value, type **1**
   (d) List the whole sequence, type **4**

   This will assign the numbers 1 through 6 four times.

4. From MENU BAR select STAT>ANOVA>TWO-WAY
   (a) Response, type **C1**
   (b) Row factor , type **C2** (MINITAB takes row factor to represent treatment effect))
   (c) Column factor, type **C3** (MINITAB takes column factor to represent block effect)
   Note: do not take "row factor" and "column factor" literally
5. Output appears in the SESSION window.

4. For the following data, present the ANOVA table. What conclusions can you draw from the two F tests? Use an $\alpha = 0.05$

	Block			
Treatment	1	2	3	4
1	14	8	10	4
2	7	7	4	2
3	12	6	16	6

{F_{Bl} = 3.462 <4.7571, Aho}
{ F_{Tr} = 3.2299 < 5.1433, Aho}

5. Three loaves of bread, each made according to a different recipe, are baked in one oven at the same time. Because of possible uncontrolled variations in oven performance, each baking is treated as a block. This procedure is repeated five times and the following measurements of density are obtained.

	Block				
Recipe	1	2	3	4	5

1	.95	.86	.71	.72	.74
2	.71	.85	.62	.72	.64
3	.69	.88	.51	.73	.44

(a) How should the three oven positions of the three loaves be selected for each trial?
{Randomize the position of the loaves in the oven}

(b) Perform an analysis of variance for these data using a 5% significance level
{$F_{Bl} = 5.31 > 3.837$, Rho}
{ $F_{Tr} = 3.92 < 4.46$, Aho}

(c) Find the p-value for the above Fcalculated value for Treatment.

Non-Parametric Tests

6. In 1995, the median age of Canadian residents was 34 years, as reported by the Census Bureau. A random sample taken this year of 10 Canadian residents yielded the following ages, in years,

40	60	12	55	34	43	47	37	9	34

At the 5% significance level, do the data provide sufficient evidence to conclude that the majority of today's Canadian residents are older than 34 years of age? {W+ =23, W r = 30, Aho}

7. Twenty years ago, the U.S. Bureau of Justice Statistics reported that the median educational attainment of jail inmates was 10 years. Ten current inmates are randomly selected and found to have the following educational attainments, in years,

14	10	5	6	8	10	10	8	9	9

Assume that the educational attainments of current jail inmates have a symmetric, non-normal distribution. At the 10% significance level, do the data provide sufficient evidence to conclude that the majority of this year's educational attainment has decreased from 10 years? {W+ = 5.5, Wl=6 Rho}

8. Several batches of fruit flies are exposed to each treatment, and the mortality percent is recorded as a measure of toxicity. The following data are obtained:

Treatment 1	40	28	31	38	43	46	29	18
Treatment 2	36	49	56	25	37	30	41	

Determine if the data strongly indicate different toxicity levels among the treatments at $\alpha = 0.05$ (assume non-normal distributions, but similar distributions). {M=62, Ml=39, Mr=73, Aho}

9. Sample data were collected to compare the ages of CEOs of top growth companies in western Canada and Quebec, shown in table below.

Western Canada	38	54	40	41	34	59	37	35	48	
Quebec	40	31	38	40	34	36	35	33	39	38

Test to see if the age for CEOs in Quebec is less than that for Western Canada at $\alpha = 0.05$.
(M=111, Mr = 111Rho)

10. Human beta-endorphin (HBE) is a hormone secreted by the pituitary gland under conditions of stress. An exercise physiologist measured the resting (unstressed) blood concentration of HBE in two groups of mean: Group 1 consisted of 11 me who had been jogging regularly for some time, and group 2 consisted of 15 men who had just entered a physical fitness program.

Joggers	Fitness Program Entrants
39 40 32 60	70 47 54 27 31
19 52 41 32	42 37 41 9 18
13 37 28	33 23 49 41 59

(a) He believes that the HBE should be higher in group 2. Test this claim using $\alpha = 0.05$. (z= -.5709, AHo)

(b) Find the p-value. (.2843)

11. Provincial governments have been re-examining their delivery of services, such as health care. The accompanying tables is based on one random sample of hospitals in New Brunswick and another in Nova Scotia, and show the total number of beds in each hospital.

New Brunswick		Nova Scotia	
23	47	15	13
153	47	27	136
397	15	8	26
12	500	85	311
15	56	12	132
398		64	

(a) At the 0.05 significance level, test the claim that the two provinces have the same distribution of hospital-bed numbers. (z=.85, Aho)

(b) What is the p-value? (.3954)

(c) At what levels of significance would you come to a different conclusion in (a)?

12. Two critics rate the service at seven award winning restaurants on a continuous 0 to 10 scale. Is there a difference between the critics' ratings at a 0.05 significance level? {W+ = 23, or W+ = 5, Aho}

Restaurant	1	2	3	4	5	6	7
Critic 1	6.1	5.2	8.9	7.4	4.3	9.7	5.5
Critic 2	7.3	5.5	9.1	7.0	5.1	9.8	5.7

13. 13 people were given a pill. Their blood pressure was measured before and after they took a pill. At the 5% significance level, determine if blood pressure has decreased after taking the pill. {W+ = 61.5, right tailed test or W+ = 4.5, left tailed test, Rho)

Before	70	80	72	76	76	76	72	78	82	64	74	92	74
After	68	72	62	70	58	66	68	52	64	72	74	60	74

14. Use the appropriate nonparametric method to perform the test in question 1. (KW=9.68>7.815, Rho)
15. Use the appropriate nonparametric method to perform the test in question 2. (KW=7.56>4.605, Rho)