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 hbox 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 $\alpha= 0.05$

{Ho: $\mu_A = \mu_B = \mu_C = \mu_D$ Ha: Not all means are equal 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? { Ho: $\mu_A = \mu_B = \mu_C$ Ha: Not all means are equal 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.

{ Ho: $\mu_1 = \mu_2 = \mu_3$ Ha: Not all means are equal Fcalc = 2.905< 4.459, Fail to RHo}

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

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 *(i.e.* 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.

{ Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$ Ha: Not all means are equal FBl = 1.5758<2.90, Fail to RHo depending on rounding}

(Ho: $\mu_A = \mu_B = \mu_C = \mu_D$ Ha: Not all means are equal FTr = 37.153>3.287, RHo}

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

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

{Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ Ha: Not all means are equal, FBl = 3.462 <4.7571, Fail to RHo}

{ Ho: $\mu_1 = \mu_2 = \mu_3$ Ha: Not all means are equal FTr = 3.2299 < 5.1433, Fail to RHo}

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
{ Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ Ha: Not all means are equal FBl = 5.31 > 3.837, RHo}

{ Ho: $\mu_1 = \mu_2 = \mu_3$ Ha: Not all means are equal FTr = 3.92 < 4.46, Fail to RHo}

(c) Find the p-value for the above F calculated value for Treatment. {.065}

Goodness of fit test.

Computer instruction
Enter observed values in column 1 and expected values in column 2.
Click on Calc<Calculator.
Type c3 in the box Store Results.
Type the formula (c1-c2)**2 /c2 in the large box. Hit enter.
Click on Calc<column statistics
Click on sum.
Type c3 in the input variable box.

1. An official of a plastics industry claimed that the industry employed 30% white women, 5% minority women, 50% white men, and 15% minority men. To test the claim, an affirmative action committee randomly sampled 150 employees and obtained the following information:

Category	observed
White females	40
Minority females	15
White males	80
Minority males	15

Test the official's claim at a 5% level of significance and find the p-value

[Ho: $p_{wf} = .3, p_{mf} = .05, p_{wm} = .5, p_{mm} = .15$ Ha: Ho is fales 10.89>7.81473 RHo, p-value = 0.0123]]

2. A computer science major claimed to have written a program that would randomly generate integers from 1 to 100. The program generated the following data. Use a 5% level of significance to test the claim. Find the p-value   [Ho: the computer randomly generates numbers, Ha: the computer does not randomly generate numbers, 6.40<16.919, Fail to RHo, p-value = .6993]

Integers	Observed
1-10	6
11-20	6

21-30	13
31-40	9
41-50	13
51-60	11
61-70	8
71-80	12
81-90	10
91-100	12

3. Given below are the frequencies observed from 310 tosses of a die. Do theses data cast doubt on the fairness of the die at the 5% significance level? Find the p-value   [Ho: the die is fair, Ha: the die is not fair, 13.225>11.0705, RHo, p-value = .0214]

Face No.	1	2	3	4	5	6
Frequency	38	61	54	65	55	37

4. A shipment of assorted nuts is labeled as having 45% walnuts, 20% hazelnuts, 20% almonds, and 15% pistachios. By randomly picking several scoops of nuts from this shipment, an inspector find the following counts.

	Walnuts	Hazelnuts	Almonds	Pistachios	Total
Counts	92	69	32	42	235

Could these findings be a strong basis for an accusation of mislabeling? Test at the 5% significance level. Find the p-value [Ho: $p_w = .45, p_H = .2, p_A = .2, p_p = .15$, Ha: Ho is false, 18.165>7.81473, RHo, p-value = .0004]

Chi-Square Tests of Independence

Minitab will perform all necessary calculations for chi-square tests on contingency tables, presenting the expected values, the value of the test statistic, degrees of freedom and the p-value.

1. Enter the observed frequencies into rows and columns just as they are given in the contingency table.
2. Go to the main header and click on **STAT>Tables>Chi-Square Test**
3. In **Columns containing the tables,** enter the columns at which your contingency table is contained.
4. Click on **OK.**

1. Over the years pollsters have found that the public's confidence in big business has been closely tied to the economic climate of the country. When businesses are growing and employment is increasing public confidence is high. When the opposite occurs, public confidence is low. In one study, Harvey Kahalas (1981) explored the relationship between confidence in big business and job satisfaction. He hypothesized that there is a relationship between the level of confidence and job satisfaction and that this relationship holds true for both union and nonunion workers. To test his hypotheses he used the sample data given in the tables below:

Union Members

Job Satisfaction

Confidence in Major Corporations	Very Satisfied	Moderately Satisfied	Little dissatisfied	Very Dissatisfied
A great deal	30	19	6	6
Only some	99	77	20	9

Hardly any	38	32	14	15

NonUnion Members

Job Satisfaction

Confidence in Major Corporations	Very Satisfied	Moderately Satisfied	Little dissatisfied	Very Dissatisfied
A great deal	111	52	13	5
Only some	246	142	37	18
Hardly any	73	51	19	9

Perform a hypothesis test on each of these data sets at the 5% significance level. Does the data support Kahalas's theory? [Ho: There is no relationship between job satisfaction and confidence in major corporations for union members, Ha: There is a relationship …., 13.359>12.5916, RHo]   [Ho: There is no relationship between job satisfaction and confidence in major corporations for non-union members, Ha: There is a relationship …., 8.298<12,5916, Fail to RHo]

2. A personnel administrator provided the following data as an example of hiring to fill 12 positions from among 40 male and 40 female applicants.

Applicant	Selected	Not Selected	Total
Male	7	33	40
Female	5	35	40

Does this sample indicate a selection bias in favour of males? Use a p-value in your conclusion. [Ho: there is no selection bias, Ha: There is a selection bias, 0.392, p-value =0.5312, Fail to RHo]

3. Applicants for public assistance are allowed an appeals process when they feel unfairly treated. At such a hearing, the applicant may choose self-representation or representation by an attorney. The appeal may result in an increase, decrease, or no change of the aid recommendation. Court records of 320 appeals cases provided the following data.

Amount of Aid

Type of Representation	Increased	Unchanged	Decreased
Self	59	108	17
Attorney	70	63	3

Are the patterns of the appeals decision significantly different between the two types of representation? Test at $\alpha = 0.05$ [Ho:There is no association between amount of aid and type of representation, Ha: There is an association…, 15.734> 5.99147, RHo]

4. A survey was conducted by sampling 400 persons who were questioned regarding union membership and attitude toward decreased national spending on social welfare programs. The cross-tabulated frequency counts are presented.

	Support	Indifferent	Opposed
Union	112	36	28
NonUnion	84	68	72
Total	196	104	100

Can these observed differences be explained by chance or are there real differences of attitude between the populations of members and non-members at the 5% significance level?
[Ho:There is no real difference of attitudes between union and nonunion, Ha: There is a real difference… 27.847>5.99147 RHo]

5. In a Study of possible genetic influence of parental hand preference, a sample of 400 children was classified according to each child's handedness and the handedness of the biological parents. Do these findings demonstrate an association between the handedness of parents and their biological offspring at

the 5% significance level?   [Ho:There is no relationship between handedness of biological offspring and parents' handedness, Ha: There is a relationship…10.653>5.99147,RHo]

Handedness of Biological Offspring

Parents' Handedness Father x Mother	Right	Left	**Total**
Right x Right	303	37	340
Right x Left	29	10	39
Left x Right	16	6	22
Total	348	52	401

6. In a genetic study of chromosome structures, 132 individuals are classified according to the type of structural chromosome aberration and carriers in their parents. The following counts are obtained.

Carrier

Type of Aberration	One Parent	Neither Parent	Total
Presumably innocuous	28	19	47
Substantially unbalanced	35	50	85
Total	63	69	132

Test the null hypothesis that type of aberration is independent of parental carrier. Use p-value [Ho: Carriers and type of aberration are independent, Ha: carriers and type of aberrations are not independent, χ^2 =4.106, p-value = .043, RHo (p-value is small)]