Homework 4
(K-Sample Inference)

1. (*Kruskal-Wallis Test*) (**Source**: http://faculty.vassar.edu/lowry/ch14a.html) To assess the effects of expectation on the perception of aesthetic quality, an investigator randomly sorts 24 amateur wine aficionados into three groups, A, B, and C, of 8 subjects each. Each subject is scheduled for an individual interview. Unfortunately, one of the subjects of group B and two of group C fail to show up for their interviews, so the investigator must make do with samples of unequal size: n_a=8, n_b=7, and n_c=6, for a total of N=21. The subjects who do show up for their interviews are each asked to rate the overall quality of each of three wines on a 10-point scale, with "1" standing at the bottom of the scale and "10" at the top.

As it happens, the three wines are the same for all subjects. The only difference is in the texture of the interview, which is designed to induce a relatively high expectation of quality in the members of group A; a relatively low expectation in the members of group C; and a merely neutral state, tending in neither the one direction nor the other, for the members of group B. At the end of the study, each subject's ratings are averaged across all three wines, and this average is then taken as the raw measure for that particular subject. The following table shows these measures for each subject in each of the three groups.

	Group		
	A	B	C
	6.4	2.5	1.3
	6.8	3.7	4.1
	7.2	4.9	4.9
	8.3	5.4	5.2
	8.4	5.9	5.5
	9.1	8.1	8.2
	9.4	8.2	
	9.7		
mean	8.2	5.5	4.9

<u>Solution</u>: As we shall demonstrate, the Kruskal-Wallis test can be computed either by using the formula given in class or by applying one-way ANOVA technique to the resultant ranks – the between-group sum of squares (SSBW).

Raw Measures				Ranked Measures			
A	B	C		A	B	C	
6.4	2.5	1.3		11	2	1	A, B, C Combined
6.8	3.7	4.1		12	3	4	
7.2	4.9	4.9		13	5.5	5.5	

8.3 8.4 9.1 9.4 9.7	5.4 5.9 8.1 8.2	5.2 5.5 8.2		17 18 19 20 21	8 10 14 15.5	7 9 15.5	
sum of ranks				131	58	42	231
average of ranks				16.4	8.3	7.0	11

Thus it can be summarized as

	A	**B**	**C**	**All**
counts	8	7	6	**21**
sums	131	58	42	**231**
means	**16.4**	**8.3**	**7.0**	**11.0**

Following one-way ANOVA technique, SSBW can be found as

A: $8(16.4—11.0)^2 = 233.3$
B: $7(8.3—11.0)^2 = 51.0$
C: $6(7.0—11.0)^2 = 96.0$

SSBW(R) = 378.7

We can then compute the Kruskal-Wallis test **T** as

$$\mathbf{T} = \frac{\mathbf{SSBW_{(R)}}}{n(n+1)/12} = \frac{378.7}{21(21+1)/12} = 9.84$$

You may check that it is exactly the same as what the formula in class provides:

$$\begin{aligned} T &= \frac{12}{n(n+1)} \sum_{k=1}^{K} \frac{R_k^2}{n_k} - 3 \cdot (n+1) \\ &= \frac{12}{21 \cdot (21+1)} \left\{ \frac{131^2}{8} + \frac{58^2}{7} + \frac{42^2}{6} \right\} - 3 \times (21+1) \\ &= 9.84 \end{aligned}$$

Applying the chi-squared approximation at this moment, we have $T > \chi^2_{.95}(3-1) = 5.99$. Thus we reject the null at sig level 0.05.

2. (*Friedman Test and Quade Test for Block Designs*) The venerable auction house of Snootly & Snobs will soon be putting three fine 17th-and 18th-century violins, A, B, and C, up for bidding. A certain musical arts foundation, wishing to determine which of these instruments to add to its collection, arranges to have them played by each of 10 concert violinists. The players are blindfolded, so that they cannot tell which violin is which; and each plays the violins in a randomly determined sequence (BCA, ACB, etc.).

   They are not informed that the instruments are classic masterworks; all they know is that they are playing three different violins. After each violin is played, the player rates the instrument on a 10-point scale of overall excellence (1=lowest, 10=highest). The players are told that they can also give fractional ratings, such as 6.2 or 4.5, if they wish. The results are shown in the adjacent table. For the sake of consistency, the n=10 players are listed as "subjects."

	Violin		
subjects	**A**	**B**	**C**
1	9.0	7.0	6.0
2	9.5	6.5	8.0
3	5.0	7.0	4.0
4	7.5	7.5	6.0
5	9.5	5.0	7.0
6	7.5	8.0	6.5
7	8.0	6.0	6.0
8	7.0	6.5	4.0
9	8.5	7.0	6.5
10	6.0	7.0	3.0

a) Apply Friedman's test to compare the three violins.

Solution: We conduct the Friedman's test using the following work sheet. Again, the formula given in class for computing the test statistic can be equivalently obtained by applying ANOVA technique directly to the assigned ranks, as illustrated below.

First, the number of blocks is $B=10$, the number of treatments is $K=3$, and the total sample size is $n=K\times B=10\times 3=30$. Using the worksheet below, we have

$$\begin{aligned}\text{SSTrt} &= \sum_{k=1}^{K} b\left(\bar{y}_{\cdot k}-\bar{y}_{\cdot\cdot}\right)^2 \\ &= 10\times(2.65-2.0)^2+10\times(2.10-2.0)^2+10\times(1.25-2.0)^2 \\ &= 9.95\end{aligned}$$

And then $T_1=\dfrac{SSTrt}{K(K+1)/12}=\dfrac{9.95}{3\cdot(3+1)/12}=9.95$. Using the formula for the case of <u>no/few ties</u>, we have the same results:

$$T_1 = \frac{12}{B \cdot K \cdot (K+1)} \sum_{k=1}^{K} \left\{ R_k - \frac{B(K+1)}{2} \right\}^2, \text{where } \frac{B(K+1)}{2} = 20$$
$$= \frac{12}{10 \cdot 3 \cdot (3+1)} \left\{ (26.6-20)^2 + (21-20)^2 + (12.5-20)^2 \right\}$$
$$= 9.95$$

To make decision, we can compare $T_1 = 9.95 > \chi^2_{.95}(2) = 5.991$. Thus we reject the null at $\alpha = 0.05$..

Table 1: An Illustration of Friedman's Test for Complete Randomized Block Design

	Original Measures				Ranks			
Subject	**A**	**B**	**C**		**A**	**B**	**C**	
1	9	7	6		3	2	1	
2	9.5	6.5	8		3	1	2	
3	5	7	4		2	3	1	
4	7.5	7.5	6		2.5	2.5	1	
5	9.5	5	7		3	1	2	
6	7.5	8	6.5		2	3	1	
7	8	6	6		3	1.5	1.5	
8	7	6.5	4		3	2	1	
9	8.5	7	6.5		3	2	1	
10	6	7	3		2	3	1	**All**
				Size	10	10	10	30
				Sum	26.5	21	12.5	60
				Mean	2.65	2.1	1.25	2.0

Alternatively, we may the T_2 test statistic that follows F distribution. In this case,
$T_1 = \frac{(B-1) \cdot T_1}{B \cdot (K-1) - T_1} = \frac{(10-1) \times 9.95}{10 \times (3-1) - 9.95} = 8.910$, which is greater than $F_{0.95}^{\{K-1,(K-1)\cdot(B-1)\}} = F_{0.95}^{(2,18)} = 3.555$. Thus again the null is rejected at $\alpha = 0.05$.

b. Apply Quade test to compare the three violins.

Solution: Using the following worksheet

	Original Data				Ranks					S Values		
	A	B	C		A	B	C	range.b	Q.b	A	B	C
1	9	7	6		3	2	1	3	6.5	6.5	0	-6.5
2	9.5	6.5	8		3	1	2	3	6.5	6.5	-6.5	0
3	5	7	4		2	3	1	3	6.5	0	6.5	-6.5
4	7.5	7.5	6		2.5	2.5	1	1.5	1.5	0.75	0.75	-1.5
5	9.5	5	7		3	1	2	4.5	10	10	-10	0
6	7.5	8	6.5		2	3	1	1.5	1.5	0	1.5	-1.5
7	8	6	6		3	1.5	1.5	2	3.5	3.5	-1.75	-1.75
8	7	6.5	4		3	2	1	3	6.5	6.5	0	-6.5
9	8.5	7	6.5		3	2	1	2	3.5	3.5	0	-3.5
10	6	7	3		2	3	1	4	9	0	9	-9

Quade Test yields F = 5.1677, num df = 2, denom df = 18, which is greater than the critical value 3.55. Using R, it can be found that p-value = 0.01685.