The usual parametric procedure is called the "one-way analysis of variance," or sometimes simply the one-way F test. The statistic used is given by

$$F = \frac{\left(\sum_{i=1}^{k} T_i^2/n_i - C\right)/(k-1)}{\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^2 - \sum_{i=1}^{k} T_i^2/n_i\right)/(N-k)} \tag{19}$$

where T_i is the sum of the observations in the ith sample, and $C = T^2/N$ where T is the total of all of the observations. If the assumptions of the Kruskal-Wallis test are valid and, in addition, if the populations have in common a normal distribution, then the quantiles of the F statistic are given in Table A22. Look in the column for $k_1 = k - 1$ and the row marked $k_2 = N - k$ for N and k given by the experiment. Violation of the normality assumption usually has little effect on the distribution of the F statistic when the null hypothesis is true. However, the power of the F test may be considerably less than the Kruskal-Wallis test for certain types of nonnormality when H_0 is false. Data containing outliers are better suited for the Kruskal-Wallis test, for example.

The A.R.E. of the Kruskal-Wallis test relative to the F test is never less than 0.864 but may be as high as infinity if the distribution functions have identical shapes but differ only in their means. If the populations are normal the A.R.E. is $3/\pi = 0.955$. For uniform distributions the A.R.E. relative to the F test is 1.0; for double exponential distributions it is 1.5. Compared with the median test the A.R.E. of the Kruskal-Wallis test is 1.5, 3.0, and 0.75, respectively, for the three distributions just mentioned.

Rank sum tests similar to the Kruskal-Wallis test have been adapted by Steel (1960), Sherman (1965), and McDonald and Thompson (1967) for making multiple comparisons. Some tables for making multiple comparisons are provided by Tobach, Smith, Rose, and Richter (1967). Procedures for selecting the best populations are described by Rizvi and Sobel (1967), Sobel (1967), Rizvi, Sobel, and Woodworth (1968), and Puri and Puri (1969). Rank tests are presented for censored data by Basu (1967b) and Breslow (1970); for testing against ordered alternatives by G.R. Shorack (1967), Odeh (1971, 1972), and Tryon and Hettmansperger (1973); and for analysis of covariance by Puri and Sen (1969a). For other work concerned with rank tests and several independent samples see Sen (1962, 1966), Matthes and Truax (1965), Quade (1966), Crouse (1966), Sen and Govindarajulu (1966), Odeh (1967), Deshpande (1970), and Bhapkar and Deshpande (1968). Analysis of covariance is discussed by Quade (1967). Brunden (1972) considers using ranks to analyze 2 × 3 contingency tables.

EXERCISES

B

1. Random samples from each of three different types of light bulbs were tested to see how long the light bulbs lasted, with the following results.

Brand		
A	B	C
73	84	82
64	80	79
67	81	71
62	77	75
70		

Do these results indicate a significant difference between brands? If so, which brands appear to differ?

2. Four job training programs were tried on 20 new employees, where 5 employees were randomly assigned to each training program. The 20 employees were then placed under the same supervisor and, at the end of a certain specified period, the supervisor ranked the employees according to job ability, with the lowest ranks being assigned to those employees with the lowest job ability.

Program	Ranks
1	4, 6, 7, 2, 10
2	1, 8, 12, 3, 11
3	20, 19, 16, 14, 5
4	18, 15, 17, 13, 9

Do these data indicate a difference in the effectiveness of the various training programs? If so, which ones seem to be different?

A 3. The amount of damage to the soil on a farm caused by water and wind is examined for many different farms. At the same time the type of farming practiced on each farm is noted, with the following results.

	Type of Farming			
	Minimum Tillage	Contour	Terrace	Other
Amount of Damage	Number of Farms			
No damage	17	19	4	21
Slight damage	3	10	4	42
Moderate damage	0	2	2	34
Severe damage	0	0	2	6

Does the type of farming affect the degree of damage? If so, which types of farming are significantly different?

B 4. Three different types of radios, manufactured by the same company, all carry 1-year guarantees. A record is kept of how many radios needed to be replaced, were repairable, or were not returned under warranty.

	Type		
	A	B	C
Replaced	12	3	6
Repaired	10	8	7
Not Returned	82	96	58

Does there seem to be a significant difference among the reliabilities of the different radio types? If so, which ones seem to be different?

A

5. The amount of iron present in the livers of white rats is measured after the animals had been fed one of five diets for a prescribed length of time. There were 10 animals randomly assigned to each of the five diets.

Diet A	Diet B	Diet C	Diet D	Diet E
2.23	5.59	4.50	1.35	1.40
1.14	0.96	3.92	1.06	1.51
2.63	6.96	10.33	0.74	2.49
1.00	1.23	8.23	0.96	1.74
1.35	1.61	2.07	1.16	1.59
2.01	2.94	4.90	2.08	1.36
1.64	1.96	6.84	0.69	3.00
1.13	3.68	6.42	0.68	4.81
1.01	1.54	3.72	0.84	5.21
1.70	2.59	6.00	1.34	5.12

Do the different diets appear to affect the amount of iron present in the livers?

6. Twelve volunteers were assigned to each of three weight-reducing plans. The assignment of the volunteers to the plans was at random, and it was assumed that the 36 volunteers in all would resemble a random sample of people who might try a weight-reducing program. Test the null hypothesis that there is no difference in the probability distributions of the amount of weight lost under the three programs against the alternative that there is a difference. The results are given as the number of pounds lost by each person.

Plan A		Plan B		Plan C	
2	17	17	5	29	5
12	4	15	6	3	25
5	25	3	19	25	32
4	6	19	4	28	24
26	21	5	9	11	36
8	6	14	7	7	20

PROBLEMS

1. Show that Equations 3 and 5 are equivalent when there are no ties.
2. Find the exact distribution of the Kruskal-Wallis test statistic when H_0 is true, $n_1 = 3$,