

MATH 4044 – Statistics for Data Science

Practical Week 11 Solutions

Question 1

The data file for this practical is called `cereals.sas7bdat` located in `mydata` library on the SAS OnDemand server. This data file contains nutritional information, rating (whether healthy or not) and supermarket shelf location for 77 breakfast cereals. Variables in that file are as follows:

Variable	Description
<i>name</i>	Name of cereal
<i>mfr</i>	Manufacturer of cereal where A = American Home Food Products; G = General Mills; K = Kelloggs; N = Nabisco; P = Post; Q = Quaker Oats; R = Ralston Purina
<i>type</i>	C = cold, H = hot
<i>calories</i>	Calories per serve
<i>protein</i>	Grams of protein
<i>fat</i>	Grams of fat
<i>sodium</i>	Milligrams of sodium
<i>fiber</i>	Grams of dietary fibre
<i>carbo</i>	Grams of complex carbohydrates
<i>sugars</i>	Grams of sugar
<i>potass</i>	Milligrams of potassium
<i>vitamins</i>	Vitamins and minerals, 0, 25, or 100, indicating the typical percentage of FDA recommended
<i>shelf</i>	Display shelf (1 = bottom, 2 = middle, or 3 = top, counting from the floor)
<i>weight</i>	Weight in ounces of one serving
<i>cups</i>	Number of cups in one serving
<i>rating</i>	Rating of the cereals calculated from Consumer Reports, out of 100. The higher the score, the healthier the cereal

- (a) Is there a significant difference in ratings of cereals displayed on different shelves? Carry out a Kruskal-Wallis test and perform post-hoc tests if appropriate. Discuss the results.

We know from previous analysis that the distributions of cereal ratings per shelf are not Normal. A Kruskal-Wallis test may therefore be more appropriate despite the loss of statistical power.

The results of the Kruskal-Wallis test in Table 1 indicate that there is a significant difference in the medians ($H = 11.11$ has chi-square distribution with 2 df, $P\text{-value} = 0.0039$).

Because the overall test is significant, pairwise comparisons among the three groups should be completed.

Wilcoxon Scores (Rank Sums) for Variable rating Classified by Variable shelf					
shelf	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
3	36	1576.0	1404.0	97.948966	43.777778
1	20	899.0	780.0	86.081357	44.950000
2	21	528.0	819.0	87.429972	25.142857

Kruskal-Wallis Test	
Chi-Square	11.1134
DF	2
Pr > Chi-Square	0.0039

Table 1. Results of a Kruskal-Wallis test

Tables 2 to 5 show results of pairwise comparisons between shelves performed using Wilcoxon tests. P-values from these tests should be compared to $\alpha = 0.05/3 = 0.0167$, to control for the experimentwise error rate.

Wilcoxon Scores (Rank Sums) for Variable rating Classified by Variable shelf					
shelf	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
3	36	1009.0	1026.0	58.480766	28.027778
1	20	587.0	570.0	58.480766	29.350000

Wilcoxon Two-Sample Test	
Statistic	587.0000
Normal Approximation	
Z	0.2821
One-Sided Pr > Z	0.3889
Two-Sided Pr > Z	0.7778
t Approximation	
One-Sided Pr > Z	0.3894
Two-Sided Pr > Z	0.7789
Z includes a continuity correction of 0.5.	

Table 2. Results of Wilcoxon test comparing shelf 1 to shelf 3

From Table 2, there is no statistically significant difference between ratings of cereals displayed on shelves 1 and 3 (P-value = 0.7778 > 0.0167).

Wilcoxon Scores (Rank Sums) for Variable rating Classified by Variable shelf					
shelf	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	20	522.0	420.0	38.340579	26.100000
2	21	339.0	441.0	38.340579	16.142857

Wilcoxon Two-Sample Test	
Statistic	522.0000
Normal Approximation	
Z	2.6473
One-Sided Pr > Z	0.0041
Two-Sided Pr > Z	0.0081
t Approximation	
One-Sided Pr > Z	0.0058
Two-Sided Pr > Z	0.0115
Z includes a continuity correction of 0.5.	

Table 3. Results of Wilcoxon test comparing shelf 1 to shelf 2

Results in Tables 3 and 4 indicate that differences between ratings of cereals placed on shelf 2 relative to the other shelves are statistically significant. For the comparison between shelf 2 and shelf 1 (Table 3), the P-value is $0.0081 < 0.0167$. For the comparison between shelf 2 and shelf 3 (Table 4), the P-value is $0.0018 < 0.0167$.

Wilcoxon Scores (Rank Sums) for Variable rating Classified by Variable shelf					
shelf	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
3	36	1233.0	1044.0	60.448325	34.250
2	21	420.0	609.0	60.448325	20.000

Wilcoxon Two-Sample Test	
Statistic	420.0000
Normal Approximation	
Z	-3.1184
One-Sided Pr < Z	0.0009
Two-Sided Pr > Z	0.0018
t Approximation	
One-Sided Pr < Z	0.0014
Two-Sided Pr > Z	0.0029
Z includes a continuity correction of 0.5.	

Table 4. Results of Wilcoxon test comparing shelf 2 to shelf 3

- (b) Convert the data to ranks and perform a one-way analysis of variance on ranks. Include post-hoc tests if appropriate. Discuss the results and compare to part (a).

Since the P-value = $0.2019 > 0.05$ for the Levine's test for homogeneity of variance in Table 6, the assumption of equal variances cannot be rejected.

We therefore do not need to consider Welch's correction to the F-ratio. From Table 5 we conclude that there is a significant difference in mean ranks of cereals displayed on different shelves, $F(2,74) = 6.34$, P-value = $0.0029 < 0.05$.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	5562.25635	2781.12817	6.34	0.0029
Error	74	32475.74365	438.86140		
Corrected Total	76	38038.00000			

R-Square	Coeff Var	Root MSE	Rank_of_rating Mean
0.146229	53.71543	20.94902	39.00000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
shelf	2	5562.256349	2781.128175	6.34	0.0029

Table 5. Results of one-way ANOVA applied to ranks of cereal ratings

Levene's Test for Homogeneity of Rank_of_rating Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
shelf	2	676518	338259	1.64	0.2019
Error	74	15308430	206871		

Welch's ANOVA for Rank_of_rating			
Source	DF	F Value	Pr > F
shelf	2.0000	5.08	0.0109
Error	39.2955		

Table 6. Results of Levine's test for homogeneity of variance

Results of follow-up tests are shown in Table 7. Using simultaneous confidence limits we find that the difference between ratings of cereals places on shelves 1 and 3 are not statistically significant, however ratings of cereals from shelf 2 and significantly different from rating of cereals displayed on the other shelves.

Tukey's Studentized Range (HSD) Test for Rank_of_rating			
Note: This test controls the Type I experimentwise error rate.			
Alpha		0.05	
Error Degrees of Freedom		74	
Error Mean Square		438.8614	
Critical Value of Studentized Range		3.38247	
Comparisons significant at the 0.05 level are indicated by ***.			
shelf Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
1 - 3	1.172	-12.801	15.146
1 - 2	19.807	4.152	35.462 ***
3 - 1	-1.172	-15.146	12.801
3 - 2	18.635	4.877	32.393 ***
2 - 1	-19.807	-35.462	-4.152 ***
2 - 3	-18.635	-32.393	-4.877 ***

Table 7. Results of post-hoc comparisons using Tukey's method

Therefore, results of one-way ANOVA applied to ranks of ratings agree with the results from part (a) of a non-parametric Kruskal-Wallis test applied to ratings directly.

APPENDIX - SAS code

```
ods graphics on;

proc npar1way data=work.cereals wilcoxon;
  class shelf;
  var rating;
run;

data work.shelf13;
  set work.cereals;
  if shelf = 2 then delete;
run;

proc npar1way data=work.shelf13 wilcoxon;
  class shelf;
  var rating;
run;

data work.shelf12;
  set work.cereals;
  if shelf=3 then delete;
run;

proc npar1way data=work.shelf12 wilcoxon;
  class shelf;
  var rating;
run;

data work.shelf23;
  set work.cereals;
  if shelf=1 then delete;
run;

proc npar1way data=work.shelf23 wilcoxon;
  class shelf;
  var rating;
run;

proc rank data=work.cereals out=rank_cereals;
  var rating;
  ranks Rank_of_rating;
run;

proc glm data=work.rank_cereals;
  class shelf;
  model Rank_of_rating=shelf / ssl;
  means shelf / hovtest Welch Tukey;
run;

ods graphics off;
```