

MORE THAN TWO INDEPENDENT SAMPLES

Case C-Q

Now we will look quickly at an example of ANOVA.

Data: Hot Dogs

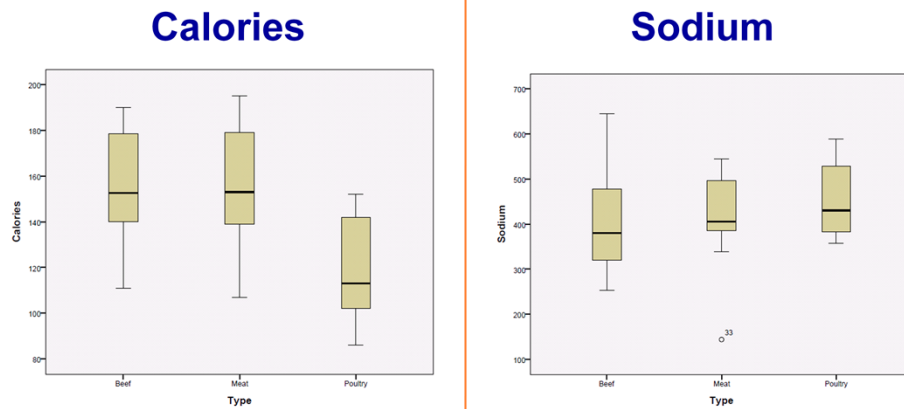
Type	Calories	Sodium
Beef	186	495
Beef	181	477
Meat	190	545
Meat	147	360
Poultry	87	359
Poultry	144	545

Data: <http://lib.stat.cmu.edu/DASL/Stories/Hotdogs.html>

In our previous section on Case C-Q, we discussed an example regarding the calories and sodium content of hot dogs.

We have two different response variables (calories and sodium) and we wish to compare beef, poultry, and other “meat” hot dogs.

Boxplots



Here are the SPSS boxplots for calories by hot dog type on the left and sodium by hot dog type on the right.

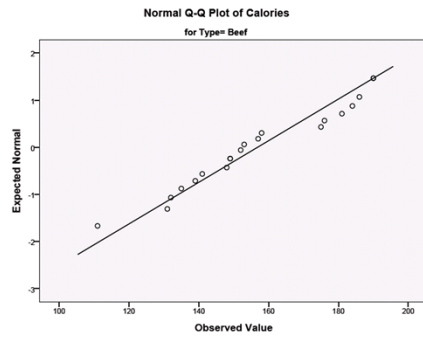
For calories, it seems clear that poultry hot dogs tend to be lower in calories but there is not much difference between Beef and Meat for calories. The variation in calories is similar for all types.

For sodium, there is no clear difference. There is one low outlier for type = meat.

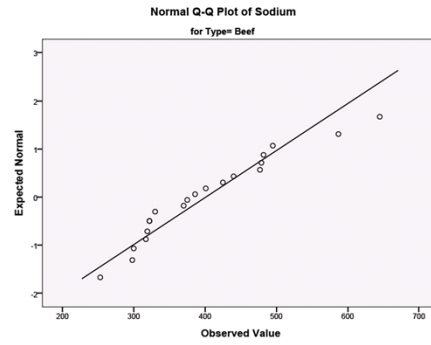
The variation in sodium content is not as consistent between the three types as for calories – however, neither is there a clear indication of a large difference in variation between these groups.

QQ-Plots: Beef

Calories



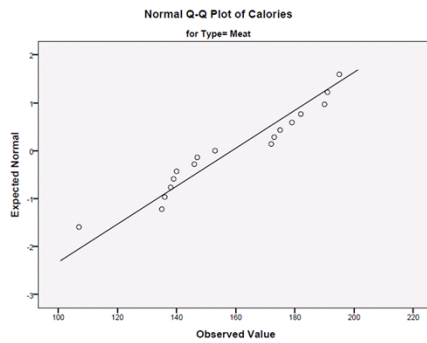
Sodium



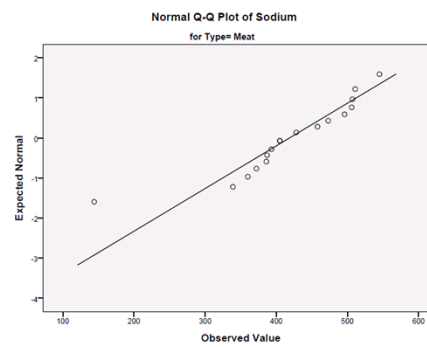
The QQ-plots for Beef show no major problems.

QQ-Plots: Meat

Calories



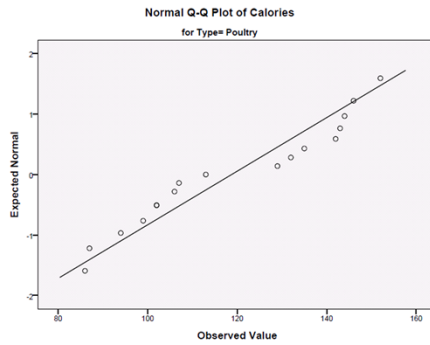
Sodium



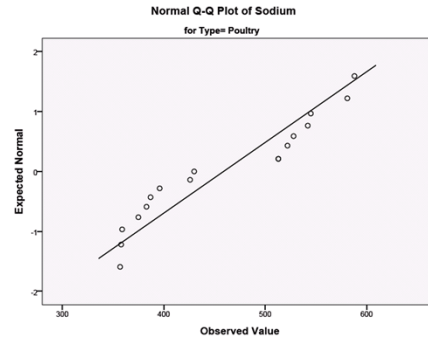
The QQ-plots for meat for both variables are not too bad but the one for sodium does show a fairly unusual outlier – as we saw in the boxplots.

QQ-Plots: Poultry

Calories



Sodium



The QQ-plots for poultry show no major problems.

Overall, the normality assumption seems reasonable for these responses within our hot dog type groups.

ANOVA - SPSS

		Sum of Squares	df	Mean Square	F	Sig.
Calories	Between Groups	17692.195	2	8846.098	16.074	.000
	Within Groups	28067.138	51	550.336		
	Total	45759.333	53			
Sodium	Between Groups	31738.715	2	15869.357	1.778	.179
	Within Groups	455248.785	51	8926.447		
	Total	486987.500	53			

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Calories is the same across categories of TypeCode.	Independent-Samples Kruskal-Wallis Test	.000	Reject the null hypothesis.
2	The distribution of Sodium is the same across categories of TypeCode.	Independent-Samples Kruskal-Wallis Test	.095	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

The results from SPSS for both the standard ANOVA and the Kruskal-Wallis test are provided here for both variables.

For calories both the ANOVA and Kruskal-Wallis test have p-values of 0.000 and thus we reject the null hypothesis.

There are statistically significant differences in mean calorie count between these three hot dog types. Although no formal test was conducted, based upon the boxplots, it seems clear that the mean calories for poultry hot dogs is different from (and in fact lower than) both beef and other meat hot dogs. However, the boxplots for the other two groups are extremely similar and thus are not likely to be found to have different means.

For sodium, however, both the ANOVA and Kruskal-Wallis test have p-values over 0.05. For the ANOVA we have a p-value of 0.179. For the Kruskal-Wallis test, we have a p-value of 0.095. In either case, we fail to reject the null hypothesis and find no evidence of any differences in the population mean sodium content between these three hot dog types.

SAS - ANOVA

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	17692.19510	8846.09755	16.07	<.0001
Error	51	28067.13824	550.33604		
Corrected Total	53	45759.33333			

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	31738.7147	15869.3574	1.78	0.1793
Error	51	455248.7853	8926.4468		
Corrected Total	53	486987.5000			

In SAS we have the same results for the standard ANOVA.

Boxplots

Calories

Wilcoxon Scores (Rank Sums) for Variable Calories Classified by Variable Type					
Type	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Beef	20	675.50	550.00	55.809023	33.775000
Meat	17	577.50	467.50	53.675400	33.970588
Poultry	17	232.00	467.50	53.675400	13.647059

Average scores were used for ties.

Kruskal-Wallis Test	
Chi-Square	19.2514
DF	2
Pr > Chi-Square	<.0001

Sodium

Wilcoxon Scores (Rank Sums) for Variable Sodium Classified by Variable Type					
Type	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
Beef	20	441.00	550.00	55.819666	22.050000
Meat	17	478.50	467.50	53.685635	28.147059
Poultry	17	565.50	467.50	53.685635	33.264706

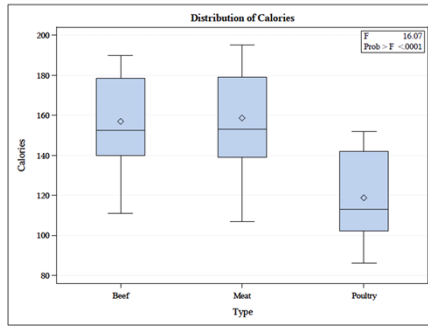
Average scores were used for ties.

Kruskal-Wallis Test	
Chi-Square	4.7128
DF	2
Pr > Chi-Square	0.0948

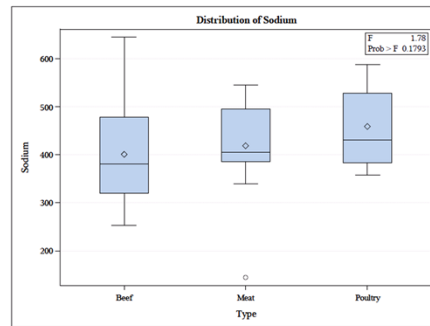
And for the Kruskal-Wallis test.

Boxplots

Calories



Sodium



These boxplots are provided by SAS with the ANOVA analysis. They provide the p-value for the F-test and illustrate the mean as well as the median for comparison.