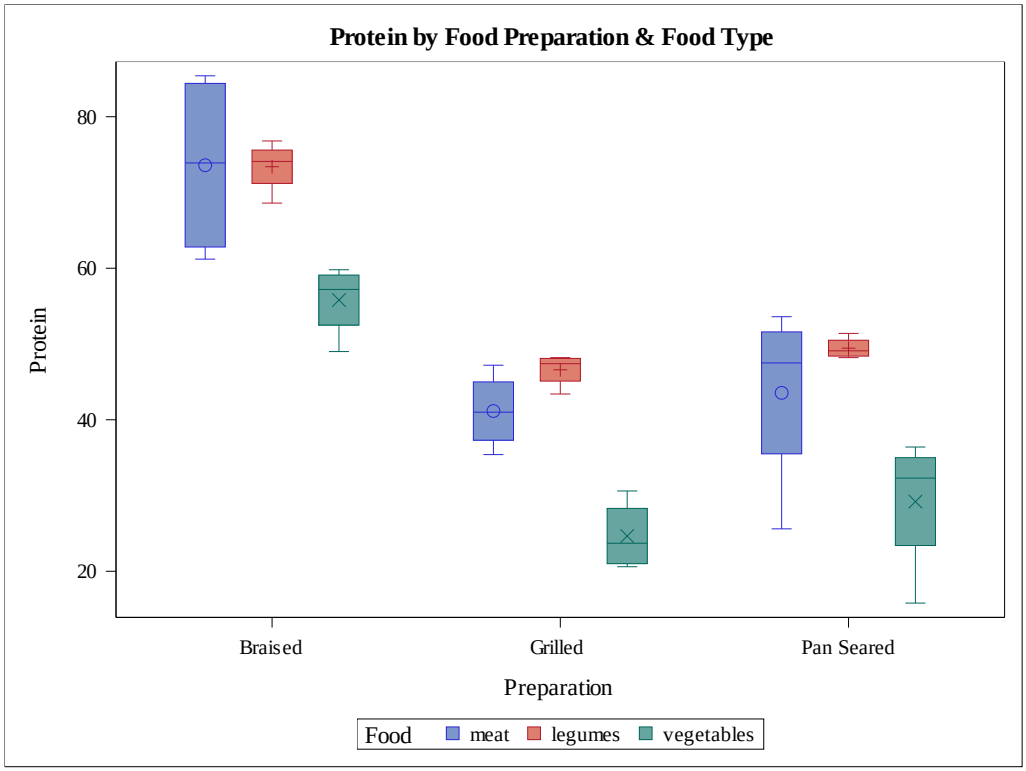
STA 9797 – HW 3

Erik Carrion

21 October 2021

Question1

The distribution of protein content for each of the three food types depends on the method of preparation. While the three different types of foods used – meat, legumes, and vegetables each have different levels of protein if we look at the side-by-side box plots by food type and preparation we see that braising results in foods with higher protein contents than pan searing and grilling, with grilling resulting in the lowest protein content out of all three methods.



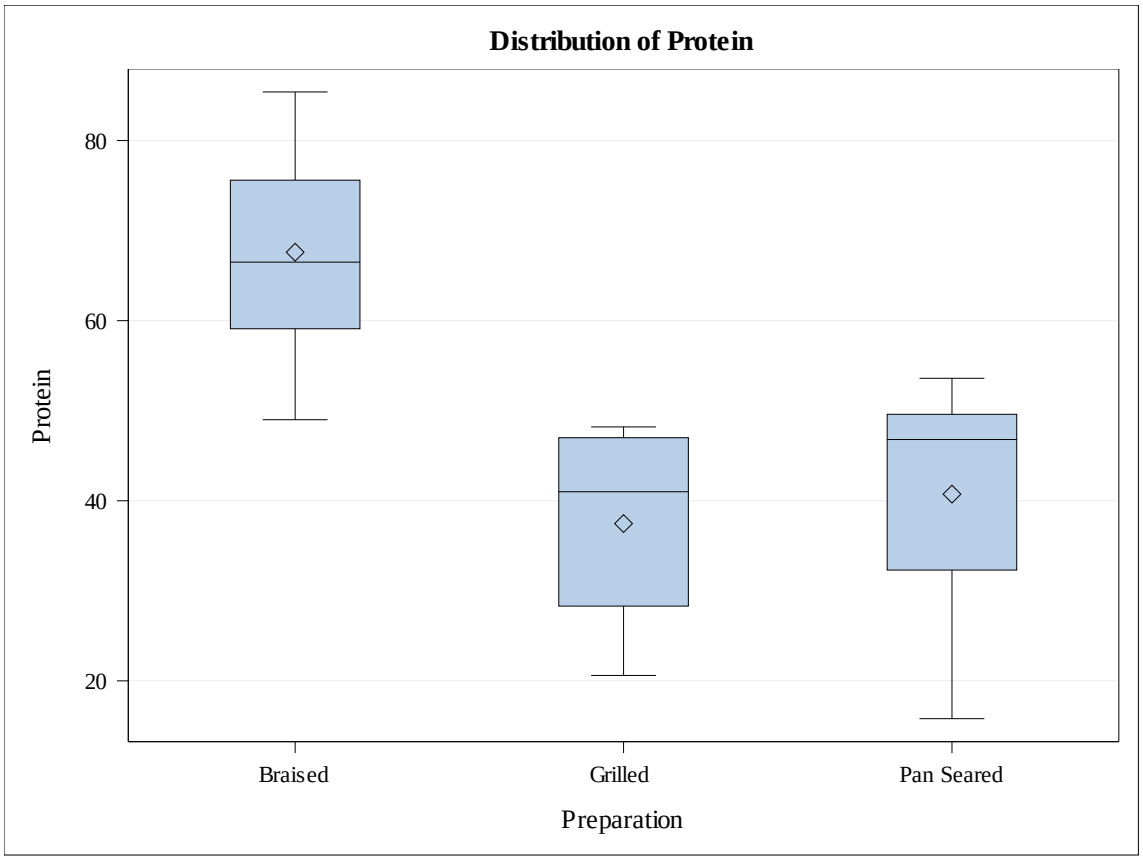
The first thing we notice is that Braising does a much better job than Grilling or Pan Searing in terms of protein retention and does so with a much lower amount of variability as we can see from its Coefficient of Variation of .1678 which is 43.35% lower than for Pan Searing.

Method	Mean Protein	Standard Deviation	Coeff. Of Variation
Braising	67.60	11.34	.1678
Grilling	37.47	10.44	.2786
Pan Searing	40.73	12.06	.2961

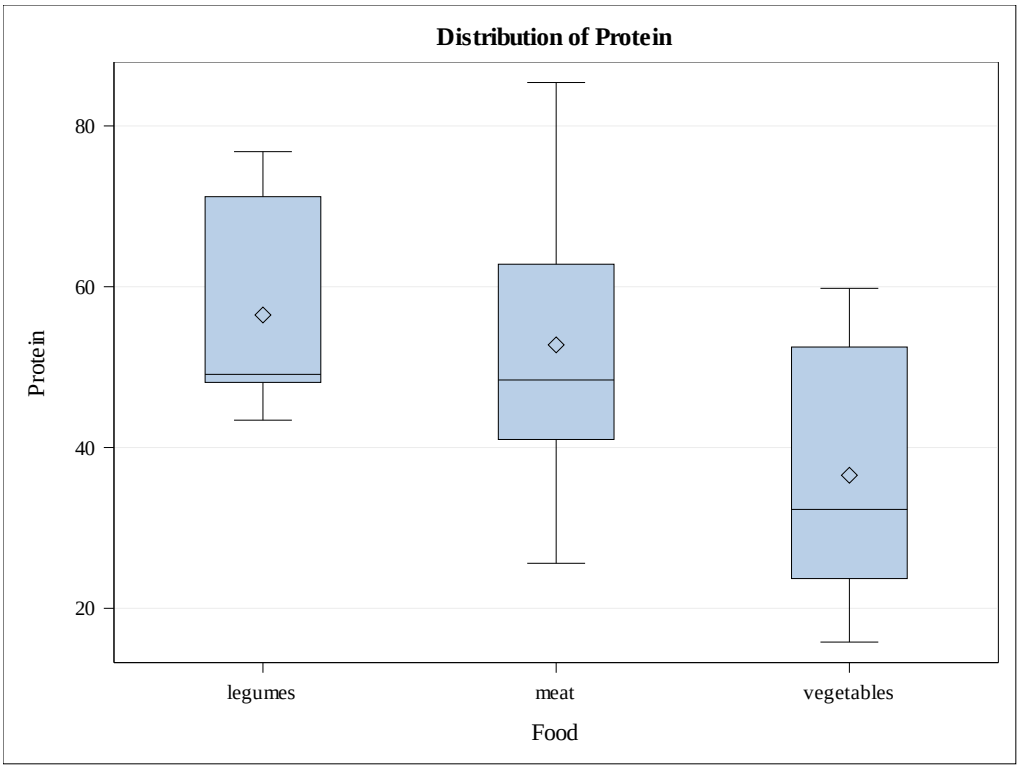
In terms of protein retention among the three food types, we see that regardless of method, vegetables have the lowest protein content. This is due to its naturally lower protein content by weight in general, so this is to be expected. Looking to the summary statistics we see that meat and legumes contain approximately the same amount of protein. However, legumes have significantly less variability with a Coefficient of Variation of .2259 compared to .4246 for vegetables.

Food	Mean Protein	Standard Deviation	Coeff. Of Variation
Meat	52.77	18.16	.3441
Legumes	56.48	12.76	.2259
Vegetables	36.55	15.52	.4246

Turning to the boxplots we start with the cooking method and immediately we see that the distribution of protein content across the three food types is symmetrically distributed, while for Pan Searing and Grilling there is some left skew in the data.



Looking at the distribution of protein by food type we notice that for each, protein content is right skewed and, as our summary statistics showed earlier, meat and legumes have similar protein content



Question 2

The purpose of the study is to investigate how protein varies from food type to food type and according to the method of preparation. There are a myriad number of types of meats, legumes, and vegetables. Furthermore, there are many more methods of food preparation than the three chosen for this study so neither the methods of preparation, nor the food types chosen, represent the full populations for either factor. Since the purpose of the study is to draw inferences regarding protein content by food type and method of preparation, we will employ a random effects model.

Running the analysis in SAS yields the following output:

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	9310.28000	1163.78500	21.57	<.0001
Error	27	1457.00000	53.96296		
Corrected Total	35	10767.2800			

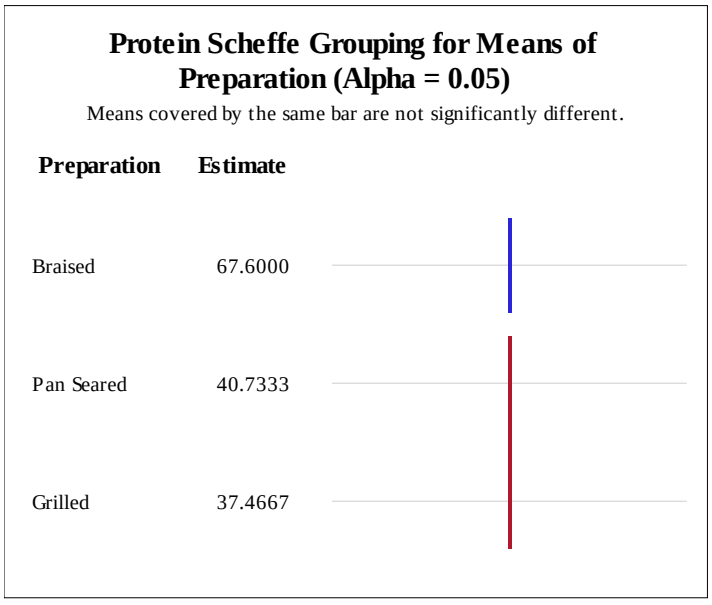
As we can see from the F-Statistic of 21.57 the overall model is highly significant with a p-value of less than 1/10,000. Furthermore, the model accounts for 86.47% of the variability in protein content.

R-Square	Coeff Var	Root MSE	Protein Mean
0.864683	15.11512	7.345949	48.60000

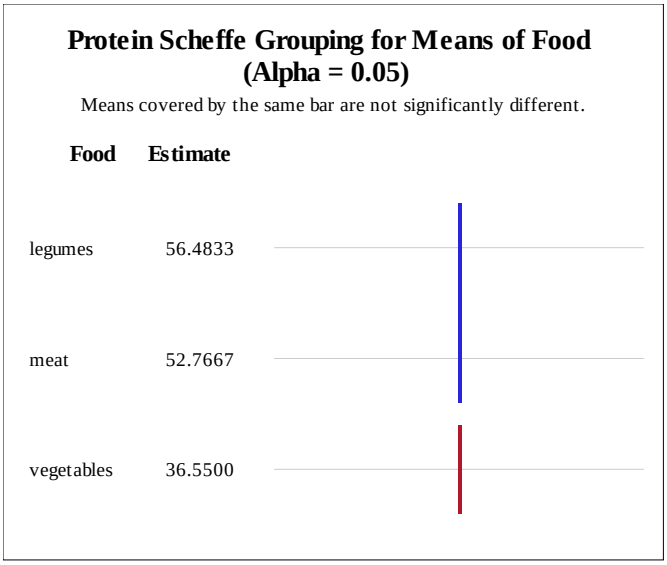
Looking at the individual factors of interest we see that both method of Preparation and the Food Type are significant, but the interaction is not at all significant with a p-value of .9134 >>>.05.

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Preparation	2	6562.026667	3281.013333	60.80	<.0001
Food	2	2696.526667	1348.263333	24.98	<.0001
Preparation*Food	4	51.726667	12.931667	0.24	0.9134

Looking at which of the means are different using Scheffe's method, we see that there is no significant difference in protein content for Pan Searing and Grilling but Braising is significantly different from the other two.



Looking at the food types, as our previous analysis of the boxplots and summary statistics suggested, there is no significant difference in protein content for meat and legumes while vegetables are significantly different from the other two.



We conclude that braising your food will retain more protein than the other two methods of preparation, suggesting that a slow, wet method of cooking at a relatively lower temperature will lead to greater preservation of protein compared to the high & dry heat methods that pan searing and grilling represent. Furthermore, regardless of the method of preparation, legumes will retain, on average, more protein than meat or vegetables. Therefore, if you are interested in a high protein diet, then a diet composed primarily of legumes and meat that have been cooked with a slow, wet method will yield the best results.

We note that without further information about which exact meats, legumes, and vegetables were chosen our analysis is limited in what it can infer. Given that our variable of interest is protein content, had the foods been equated by protein such that each portion of meat, legumes, and vegetables in the stuyd had the same protein content prior to being cooked, we would have been able to better gauge the extent to which protein is preserved (or not) by the cooking method. As such, our inferential ability is hampered in that regard.

Question 3

Had we chosen a fixed effects model we would have been even more limited in our inferences as we can only say that braising the specific legumes, using the specific liquids and containers used in the study will lead to higher protein content in the final product. We would not be able to say, for example, that cooking any type of legume sous vide, for example, would lead to higher protein content. Cooking Sous vide shares many similarities with braising, so making that inference in a random effects model is legitimate while in a fixed effects model would be inappropriate.

```
proc import file = "protein.xlsx" dbms = xlsx out = protein;
   run;

ods rtf file = "hw3 summary.rtf";
proc summary data = protein;
      class Preparation Food;
      var Protein;
      output out = pdata;
run;
proc print data = pdata;
run;
ods rtf close;

ods rtf file = "boxplot.rtf";
proc sgplot data = protein;
      title "Protein by Food Preparation & Food Type";
      vbox protein / category = Preparation group = Food;
run;
ods rtf close;
ods rtf file="9797 HW3 2 Way ANOVA.rtf";
ods graphics on;
proc glm data = protein;
      title "Protein 2-Way ANOVA - Random Effects";
      class Preparation Food;
      model Protein = Preparation Food Preparation*Food;
      random Preparation Food Preparation*Food;
      means Preparation Food / Scheffe;
run;
ods graphics off;
ods rtf close;
```