Chapter 26 – Multifactor Analysis of Variance

Section 26.1

1. Student TV time.

$$TV_{ijk} = \mu + Sex_j + Athlete_k + \varepsilon_{ijk}$$

2. Cookies.

$$Score_{ijk} = \mu + Chips_j + Sugar_k + \varepsilon_{ijk}$$

Section 26.2

3. Student TV assumptions.

The boxplots show different distributions for Women and Men, so the Additivity Assumption is probably not satisfied. The plots show that, at least for men, the Equal Variance Assumption seems unlikely. And both plots show outliers, making the Normality Assumption questionable. If the survey was randomized, the responses may be mutually independent, but we don't have enough information to tell.

4. Cookie assumptions.

The boxplots show some differences, which may indicate a failure of additivity, but the differences are not very great. The ratings of the cookies should be independent. There is some difference in the variabilities, but it is not great. There are a few outliers, which may challenge the Normality Assumption.

Section 26.3

5. Student TV interactions.

The interaction term reflects the fact that there is very little difference in TV watching among women between athletes and non-athletes, but there is a difference for men with athletes watching more TV.

6. Have ANOVA cookie.

Milk chocolate chips combined with the middle amount of sugar made a low-scoring cookie. The other recipe combinations were additive.

Chapter Exercises.

7. Popcorn revisited.

a) H_0: The effect due to power level is the same for each level. $(\gamma_{Low} = \gamma_{Med} = \gamma_{High})$
H_A: Not all of the power levels have the same effect on popcorn popping.

H_0: The effect due to popping time is the same for each time. $(\tau_3 = \tau_4 = \tau_5)$
H_A: Not all of the popping times have the same effect on popcorn popping.

b) The *power* sum of squares has 3 – 1 = 2 degrees of freedom.
The *popping time* sum of squares has 3 – 1 = 2 degrees of freedom.
The error sum of squares has (9 – 1) – 2 – 2 = 4 degrees of freedom.

c) Because the experiment did not include replication, there are no degrees of freedom left for the interaction term. The interaction term would require 2(2) = 4 degrees of freedom, leaving none for the error term, making any tests impossible.

8. Gas mileage revisited.

a) H_0: The effect on gas mileage due to tire pressure is the same for each level. $(\gamma_{Low} = \gamma_{Med} = \gamma_{Full})$
H_A: Not all of the tire pressure levels have the same effect on gas mileage.

H_0: The effect on gas mileage due to acceleration is the same for each level. $(\tau_S = \tau_P)$
H_A: The two acceleration levels have different effects.

b) The *tire pressure* sum of squares has 3 – 1 = 2 degrees of freedom.
The *acceleration* sum of squares has 2 – 1 = 1 degrees of freedom.
The error sum of squares has (24 – 1) – 2 – 1 = 20 degrees of freedom.

8. (continued)

c) Yes, he should consider fitting an interaction term. Because this is a replicated experiment, there are enough degrees of freedom left for an interaction term. The effect of tire pressure may change, depending on the level of acceleration.

d) Interaction degrees of freedom equals (3 – 1) (2 – 1) = 2.

9. Popcorn again.

a) The *F*-statistic for *power* is 13.56 with 2 and 4 degrees of freedom, resulting in a P-value equal to 0.0165. The *F*-statistic for *time* is 9.36 with 2 and 4 degrees of freedom, resulting in a P-value equal to 0.0310.

b) With a P-value equal to 0.0165 we reject the null hypothesis that *power* has no effect and conclude that the mean number of uncooked kernels is not equal across all 3 *power* levels. With a P-value equal to 0.0310 we reject the null hypothesis that *time* has no effect and conclude that the mean number of uncooked kernels is not equal across all 3 *time* levels.

c) **Randomization condition:** The bags should be randomly assigned to treatments.
Similar Variance condition: The side-by-side boxplots should have similar spreads. The residuals plot should show no pattern, and no change in spread.
Nearly Normal condition: The Normal probability plot of the residuals should be straight, the histogram of the residuals should be unimodal and symmetric.

10. Gas mileage again.

a) With a P-value equal to 0.241 there is no interaction effect of *tire pressure* with *acceleration*. The *F*-statistic for *tire pressure* is 4.29 with 2 and 20 degrees of freedom, resulting in a P-value equal to 0.030, we reject the null hypothesis and conclude that *tire pressure* has a significant effect on gas mileage. The *F*-statistic for *acceleration* is 2.35 with 1 and 20 degrees of freedom, resulting in a P-value equal to 0.143, we fail to reject the null hypothesis, there is not enough evidence to conclude that acceleration has any effect on gas mileage.

b) **Randomization condition:** The trials are randomized.
Similar Variance condition: The side-by-side boxplots should have similar spreads. The residuals plot should show no pattern, and no systematic change in spread.
Nearly Normal condition: The Normal probability plot of the residuals should be straight, the histogram of the residuals should be unimodal and symmetric.

c) You have made a Type II error.

11. Crash analysis.

a) H_0: The effect on head injury severity is the same for both seats. $(\gamma_D = \gamma_P)$
H_A: The effects on head injury severity are different for the two seats.

H_0: The size of the vehicle has no effect on head injury severity. $(\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \tau_6)$
H_A: The size of the vehicle does have an effect on head injury severity.

b) **Randomization condition:** Assume that the cars are representative of all cars.
Additive Enough condition: The interaction plot is reasonably parallel.
Similar Variance condition: The side-by-side boxplots have similar spreads. The residuals plot shows no pattern, and no systematic change in spread.
Nearly Normal condition: The Normal probability plot of the residuals should be straight, the histogram of the residuals should be unimodal and symmetric.

c) With a P-value equal to 0.838 the interaction term between *seat* and *vehicle size* is not significant. With a P-value less than 0.0001 for both *seat* and *car size* we reject both null hypotheses and conclude that both *seat* and *car size* affect the severity of head injury. By looking at one of the partial boxplots we see that the mean head injury severity is higher for the driver's side. The effect of driver's *seat* seems to be roughly the same for all six cars.

12. Sprouts.

a) H_0: The temperature level has no effect on biomass. $(\gamma_{32} = \gamma_{34} = \gamma_{36})$
H_A: The effects of the temperature levels are not all the same.

H_0: The effect on biomass is the same for all salinity levels. $(\tau_0 = \tau_4 = \tau_8 = \tau_{12})$
H_A: The effects of salinity levels are not all the same.

b) **Randomization condition:** Assume that the mung beans are representative of all mung beans.
Additive Enough condition: The interaction plot is reasonably parallel.
Similar Variance condition: The side-by-side boxplots have similar spreads. The residuals plot shows no pattern, and no systematic change in spread.
Nearly Normal condition: The Normal probability plot of the residuals should be straight, the histogram of the residuals should be unimodal and symmetric.

c) With a P-value of 0.3244 there is no evidence of an interaction between *temperature* and *salinity* on biomass. The P-values for *temperature* and *salinity* are both less than 0.0001. We reject the null hypothesis and conclude that the mean biomass is different for different *temperature* levels and different *salinity* levels. We conclude that both *salinity* and *temperature* affect the amount of mung bean sprout biomass produced.

13. Baldness and Heart Disease.

a) A two-factor ANOVA must have a quantitative response variable. Here the response is whether they exhibited baldness or not, which is a categorical variable. A two-factor ANOVA is not appropriate.

b) We could use a chi-square analysis to test whether *baldness* and *heart disease* are independent.

14. Fish and prostate.

a) A two-factor ANOVA must have a quantitative response variable. Here the response is whether the subjects suffered prostate cancer or not, a categorical variable. A two-factor ANOVA is not appropriate.

b) We could use a chi-square analysis to test whether amount of *fish* in the diet and incidence of *prostate cancer* are independent.

15. Baldness and heart disease again.

a) A chi-square test of independence gives a chi-square statistic of 14.510 with a P-value equal to 0.0023. We reject the hypothesis that baldness and heart disease are independent.

b) No, the fact that these are not independent does not mean that one causes the other. There could be a lurking variable (such as age) that influences both.

16. Fish and prostate again.

A chi-square test of independence finds a chi-square statistic of 3.677 with 3 degrees of freedom and a P-value equal to 0.2985. We fail to reject the null hypothesis that *fish* consumption and incidence of *prostate* cancer are independent. The data provide no evidence that these are associated.

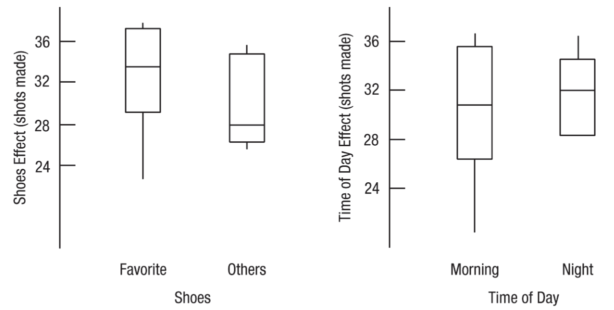
17. Basketball shots.

a) H_0: The time of day has no effect on the number of shots made. $(\gamma_M = \gamma_N)$
H_A: The time of day does have an effect on the number of shots made.

H_0: The type of shoe has no effect on the number of shots made. $(\tau_F = \tau_O)$
H_A: The type of shoe does have an effect on the number of shots made.

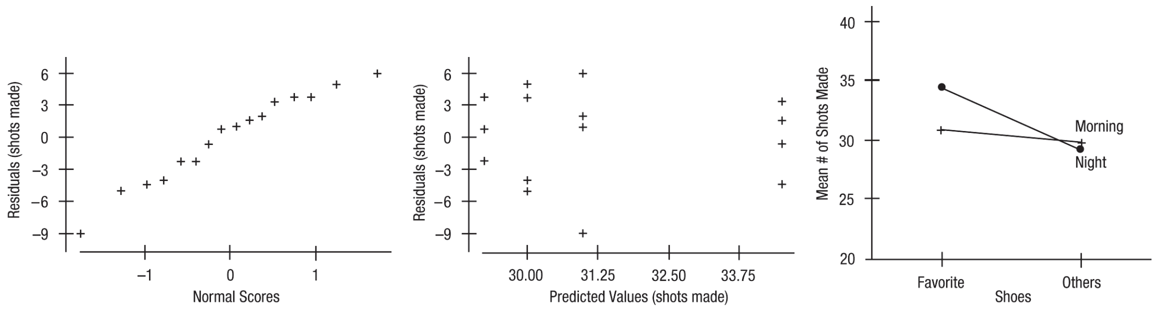
17. (continued)

b)



The partial boxplots show little effect of either *time of day* or *shoes* on the number of shots made.

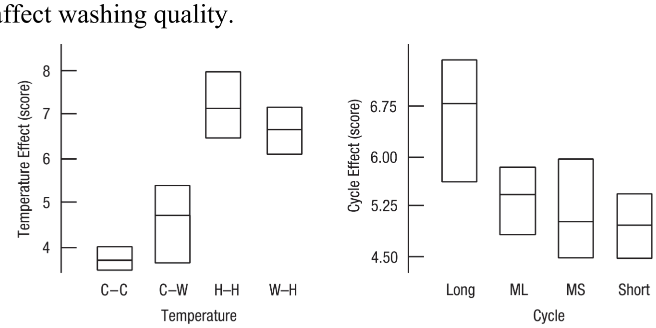
Randomization condition: We assume that the number of shots made were independent from one treatment condition to the next.
Similar Variance condition: The side-by-side boxplots have similar spreads. The residuals plot shows no pattern, and no systematic change in spread.
Nearly Normal condition: The Normal probability plot of the residuals is straight.



The interaction plot shows a possible interaction effect. It looks as though the favorite shoes may make more of a difference at night. However, we fail to reject the null hypothesis that there is no interaction effect. In fact, none of the effects appears to be significant. It looks as though she cannot conclude that either *shoes* or *time of day* affect her mean free throw percentage.

18. Washing.

We will test the effects of *temperature* settings and *cycle* lengths on washing quality. The partial boxplots show that both factors may affect washing quality.

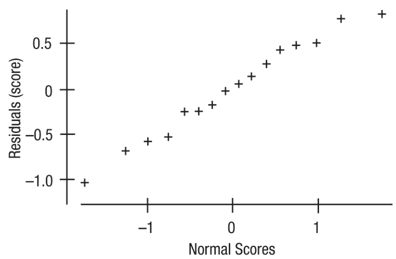


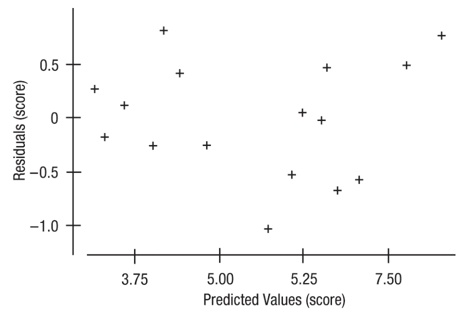
H_0: The temperature has no effect on the washing quality. $(\gamma_{CC} = \gamma_{CW} = \gamma_{WH} = \gamma_{HH})$
H_A: The temperature does have an effect on the washing quality.

H_0: The cycle has no effect on the washing quality. $(\tau_S = \tau_{MS} = \tau_{ML} = \tau_L)$
H_A: The cycle does have an effect on the washing quality.

18. (continued)

Randomization condition: We assume that he randomly assigned handkerchiefs to treatments.
Similar Variance condition: The side-by-side boxplots have relatively similar spreads. The residuals plot does not show enough of a pattern to cause concern.
Nearly Normal condition: The Normal probability plot of the residuals is reasonably straight.





We assume that the interaction effects are negligible. The experiment was unreplicated, so no degrees of freedom are available to estimate them.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F-ratio	P-Value
Temp	3	33.2519	11.084	23.468	0.0001
Cycle	3	7.19688	2.39896	5.0794	0.0250
Error	9	4.25062	0.472292		
Total	15	44.6994			

The ANOVA table confirms what we saw in the boxplots: both *temperature* and *cycle* have significant effects on washing quality. With such small P-values we reject both null hypotheses and conclude that both the *temperature* settings and *cycle* lengths affect the cleanliness score.

19. Sprouts sprouted.

a) H_0: The temperature level has no effect on the number of sprouts. $(\gamma_{32} = \gamma_{34} = \gamma_{36})$
H_A: The temperature level does have an effect on the number of sprouts.

H_0: The salinity level has no effect on the number of sprouts. $(\tau_0 = \tau_4 = \tau_8 = \tau_{12})$
H_A: The salinity level does have an effect on the number of sprouts.

b) **Randomization condition:** We assume that the sprouts were representative of all sprouts.
Similar Variance condition: There appears to be more spread in the number of sprouts for the lower salinity levels. This is cause for some concern, but most likely does not affect the conclusions.
Nearly Normal condition: The Normal probability plot of the residuals should be straight.

The partial boxplots show that *salinity* level appears to have an effect on the number of bean sprouts while *temperature* does not. The ANOVA supports this. With a P-value less than 0.0001 for *salinity*, there is strong evidence that the salinity level affects the number of bean sprouts. However, it appears that neither the interaction term nor *temperature* have a significant effect on the number of bean sprouts. The interaction term has a P-value equal to 0.7549 and *temperature* has a P-value equal to 0.3779.

20. Containers revisited.

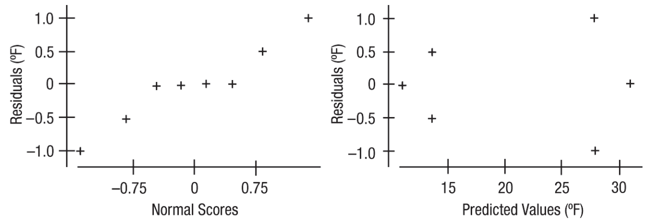
a) H_0: The type of liquid has no effect on the change in temperature. $(\gamma_W = \gamma_C)$
H_A: The type of liquid does have an effect on the change in temperature.

H_0: The environment has no effect on the change in temperature. $(\tau_O = \tau_R)$
H_A: The environment does have an effect on the change in temperature.

20. (continued)

b) Both *liquid* and *environment* are significant with P-values of 0.0079 and less than 0.0001, respectively. The interaction of *liquid* and *environment* is not significant.

c) The Normal probability plot is reasonably straight and the residuals plot shows no definite pattern. There appear to be no violations of the Nearly Normal condition nor the Similar Variance condition.



d) Both the type of *liquid* and the *environment* seem to affect how well the containers maintain heat. On average the change is about 2.75 degrees less for coffee than water. On average, the change is about 17.25 degrees less at room temperature than outside.

21. Gas additives.

H_0: The car type has no effect on gas mileage. $(\gamma_H = \gamma_M = \gamma_S)$
H_A: The car type does have an effect on gas mileage.

H_0: The mean gas mileage is the same for both additives. $(\tau_G = \tau_R)$
H_A: The mean gas mileage is different for the additives.

The ANOVA table shows that both *car type* and *additive* affect gas mileage, with P-values less than 0.0001. There is a significant interaction effect as well that makes interpretation of the main effects problematic. However, the residual plot shows a strong increase in variance, which makes the whole analysis suspect. The Similar Variance condition appears to be violated.

22. Chromatography

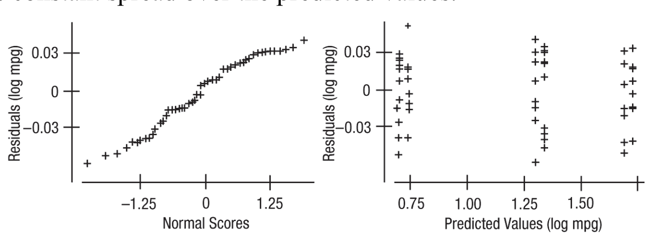
H_0: The mean counts are the same for both flow rates. $(\gamma_F = \gamma_S)$
H_A: The mean counts are not the same for the flow rates.

H_0: The mean counts are the same for each concentration. $(\tau_L = \tau_M = \tau_H)$
H_A: The mean counts are not all the same for the concentrations.

The ANOVA table shows that both *concentration* and *flow rate* affect counts, with P-values less than 0.0001. There is a significant interaction effect as well that makes interpretation of the main effects problematic. However, the residual plot shows a strong increase in variance, which makes the whole analysis suspect. The Similar Variance condition appears to be violated.

23. Gas additives again.

After the re-expression of the response, gas mileage, the Normal probability plot of residuals looks straight and the residuals plot shows constant spread over the predicted values.



After the re-expression, the ANOVA table only shows the main effects to be significant, while the interaction term is not. We can conclude that both the *car type* and *additive* have an effect on mileage and that the effects are constant (in log(*mpg*)) over the values of the various levels of the other factor.

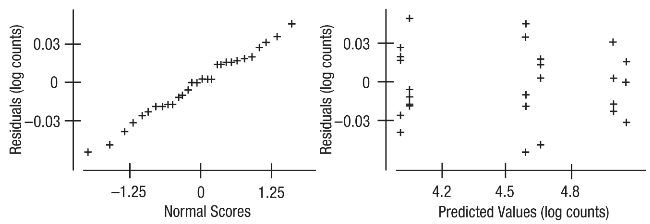
23. (continued)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F-ratio	P-Value
Type	2	10.1254	5.06268	5923.1	< 0.0001
Additive	1	0.026092	0.0260915	30.526	< 0.0001
Interaction	2	7.57E-05	3.78E-05	0.044265	0.9567
Error	54	0.046156	8.55E-04		
Total	59	10.1977			

24. Chromatography again.

After the re-expression of the response, total counts, the Normal probability plot of residuals looks straight and the residuals plot shows constant spread over the predicted values.



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F-ratio	P-Value
Concentration	2	5.14171	2.57085	2992	< 0.0001
Flow rate	1	0.022372	0.0223722	26.038	< 0.0001
Interaction	2	2.88E-04	1.44E-04	0.16736	0.8469
Error	24	0.020622	8.59E-04		
Total	29	5.18499			

After the re-expression, the ANOVA table only shows the main effects to be significant, while the interaction term is not. We can conclude that both the *concentration* and *flow rate* have an effect on counts and that the effects are constant (in log(*total counts*)) over the values of the various levels of the other factors.

25. Batteries.

a) H_0: The mean times are the same under both environments. $(\gamma_C = \gamma_{RT})$
H_A: The mean times are different under the two environments.

H_0: The brand of battery has no effect on time. $(\tau_A = \tau_B = \tau_C = \tau_D)$
H_A: The brand of batter does have an effect on time.

b) From the partial boxplots it appears that *environment* does have an effect on time, but it is unclear whether *brand* has an effect on time.

c) Yes, the ANOVA does match our intuition based on the boxplots. The *brand* effect has a P-value equal to 0.099. While this is not significant at the $\alpha = 0.05$ level, it is significant at the $\alpha = 0.10$ level. As it appeared on the boxplot, the *environment* is clearly significant with a P-value less than 0.0001.

d) There is also an interaction, however, which makes the statement about *brands* problematic. Not all *brands* are affected by the environment in the same way. Brand C, which works best in the cold, performs worst at room temperature.

e) I would be uncomfortable recommending brand C because it performs the worst of the four at room temperature.

26. Peas.

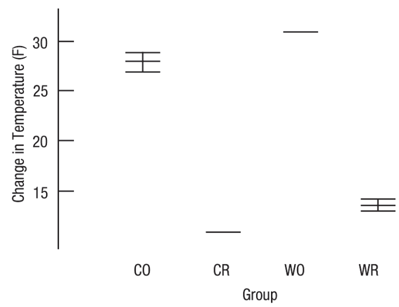
Varying levels of *Quickgrow* affect the growth of sweet peas as measured by weight (P-value = 0.0013). The levels of *water* used did not change the weight enough to be statistically significant (P-value = 0.095). Using *little* amounts of *Quickgrow* increased weight by an average 369 mg over using *moderate* or *full* amounts. Using *full* amounts of water increased weight an average of almost 200 mg over *little* water, but this increase was not statistically significant.

27. Batteries once more.

In this one-way ANOVA, we can see that the means vary across treatments. (However, boxplots with only 2 observations are not appropriate.) By looking closely, it seems obvious that the four flashlights at room temperature lasted much longer than the ones in the cold. It is much harder to see whether the means of the four brands are different, or whether they differ by the same amounts across both environmental conditions. The two-way ANOVA with interaction makes these distinctions clear.

28. Containers one more time.

The boxplots are shown below, but with only two observations in each group, a boxplot is not appropriate. There are clearly differences (as shown in the ANOVA table below), but it is difficult to assess how much change is due to changes in *environment* and how much is due to the *liquid*. A two-factor ANOVA makes this much clearer by making it possible to separate and better interpret the effects of brand and environment.



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F-ratio	P-Value
Treatment	3	610.375	203.4583	325.533	< 0.0001
Error	4	2.5	0.625		
Total	7	612.575			