Solution for Exam 2010 Problem 5

ANYM, 20191111

We have three variables y_1, y_2, and y_3, and two effects 'Treatment' and 'Block'.

As we are interested in the effects it is a *two-sided MANOVA,* which we will use for the analysis, see section 4.3.2 page 304.

We are further supplied with a SAS-script, that we run. I will show how to solve the exercise by hand and show the relevant parts in the SAS-output as well as general comments at the end of the document.

Q 5.1

The usual test statistic for treatment effect. We use theorem 4.26 page 306.

Theorem 4.26

The ratio test at level α for test of H_0 against H_1 is given by the critical region

$$\{y_{11},\dots,y_{km} \left| \frac{\det(\mathbf{q}_1)}{\det(\mathbf{q}_1+\mathbf{q}_2)} \le U(p,k-1,(k-1)(m-1))_\alpha \right.\}.$$

The ratio test at level α for test of K_0 against K_1 is given by the critical region

$$\{y_{11},\dots,y_{km} \left| \frac{\det(\mathbf{q}_1)}{\det(\mathbf{q}_1+\mathbf{q}_3)} \le U(p,m-1,(k-1)(m-1))_\alpha \right.\}.$$

p is the dimensionality of our observations, i.e. p=3. If we denote treatment by k, and blocks by h, we get k=4 and h=3:

The test for treatment is then:

$$U(p,k-1,(k-1)(m-1)) = U(3,4-1,(4-1)(3-1)) = U(3,3,6)$$

The answer is thus 3.

Q 5.1 - SAS

We now look at the SAS-output:

The GLM Procedure
Multivariate Analysis of Variance

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Treatment Effect H = Type III SSCP Matrix for Treatment E = Error SSCP Matrix S=3 M=-0.5 N=1					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
NOTE: F Statistic for Roy's Greatest Root is an upper bound.					
Wilks' Lambda	0.00202344	12.95	9	9.8856	0.0002
Pillai's Trace	2.41901846	8.33	9	18	<.0001
Hotelling-Lawley Trace	29.34255673	13.04	9	4	0.0124
Roy's Greatest Root	16.69341645	33.39	3	6	0.0004

We see here that SAS specifies the degrees of freedom as S, M and N. If we let U(p,q,r), they are given by:

$$S = \min(p, q) = \min(3,3) = 3$$

$$M = \frac{|p-q|-1}{2} = \frac{|3-3|-1}{2} = \frac{-1}{2} = -0.5$$

$$N = \frac{r-p-1}{2} = \frac{6-3-1}{2} = \frac{2}{2} = 1$$

More details can be found in the link below.

https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_introreg_sect012.htm

More important is the interpretation! We see that the treatment effect is significant in all tests, and thus conclude that treatments affects the contents of the apples!

Q 5.2

The transformation of the test statistic for “no block effect” to the F-distribution?

We can write it up from the theorem above:

$$U\big(p, m-1, (k-1)(m-1)\big) = U\big(3, 3-1, (4-1)(3-1)\big) = U(3,2,6)$$

For the transformation between the U- and F-distribution we use theorem 4.22 on page 296:

Theorem 4.22

Let U be U(s,r,n-k)-distributed and let

$$t = \begin{cases} 1 & s^2 + r^2 = 5 \\ \sqrt{\frac{s^2r^2-4}{s^2+r^2-5}} & s^2 + r^2 \neq 5 \end{cases}$$

$$v = \frac{1}{2}(2(n-k) + r - s - 1).$$

Then

$$F = \frac{1 - U^{\frac{1}{t}}}{U^{\frac{1}{t}}} \cdot \frac{vt + 1 - \frac{1}{2}sr}{sr}$$

is approximately distributed as

$$F(sr, vt + 1 - \frac{1}{2}sr).$$

If either s or r are equal to 1 or 2, then the approximation is exact.

As we have $v_2 = 2$ the transformation is exact, and the answer is 5.

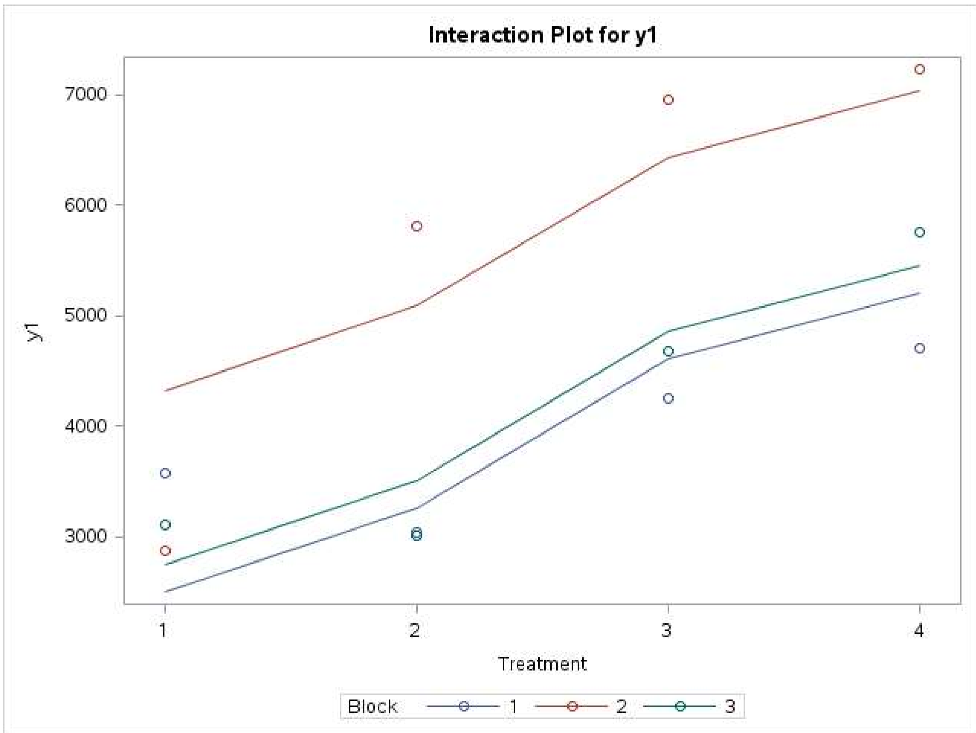
In SAS we have:

The GLM Procedure
Multivariate Analysis of Variance

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Block Effect H = Type III SSCP Matrix for Block E = Error SSCP Matrix S=2 M=0 N=1					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
NOTE: F Statistic for Roy's Greatest Root is an upper bound.					
NOTE: F Statistic for Wilks' Lambda is exact.					
Wilks' Lambda	0.19344994	1.70	6	8	0.2387
Pillai's Trace	1.11090541	2.08	6	10	0.1459
Hotelling-Lawley Trace	2.59599315	1.73	6	4	0.3099
Roy's Greatest Root	1.63190102	2.72	3	5	0.1545

We see that the test is exact as expected. We also see that the block effect is not significant at the usual 5 % level.

This might be a slight surprise, as the interaction plot for y1 seems to indicate some effect.



SAS-output

When doing a MANOVA in SAS we both get the univariate output, as well as the MANOVA.

The univariate output should already be familiar to you – take note of colored numbers:

The GLM Procedure
Dependent Variable: y1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	21547408.33	4309481.67	5.12	0.0357
Error	6	5052016.67	842002.78		
Corrected Total	11	26599425.00			

R-Square	Coeff Var	Root MSE	y1 Mean
0.810070	20.02416	917.6071	4582.500

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	3	13729558.33	4576519.44	5.44	0.0380
Block	2	7817850.00	3908925.00	4.64	0.0605

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Treatment	3	13729558.33	4576519.44	5.44	0.0380
Block	2	7817850.00	3908925.00	4.64	0.0605

The GLM Procedure
Dependent Variable: y2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	27414166.67	5482833.33	10.96	0.0056
Error	6	3002000.00	500333.33		
Corrected Total	11	30416166.67			

R-Square	Coeff Var	Root MSE	y2 Mean
0.901302	7.044074	707.3424	10041.67

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	3	24868100.00	8289366.67	16.57	0.0026
Block	2	2546066.67	1273033.33	2.54	0.1584

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Treatment	3	24868100.00	8289366.67	16.57	0.0026
Block	2	2546066.67	1273033.33	2.54	0.1584

The GLM Procedure
Dependent Variable: y3

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	34268.25000	6853.65000	14.77	0.0026
Error	6	2784.66667	464.11111		
Corrected Total	11	37052.91667			

R-Square	Coeff Var	Root MSE	y3 Mean
0.924846	11.58758	21.54324	185.9167

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	3	31659.58333	10553.19444	22.74	0.0011
Block	2	2608.66667	1304.33333	2.81	0.1376

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Treatment	3	31659.58333	10553.19444	22.74	0.0011
Block	2	2608.66667	1304.33333	2.81	0.1376

The new part is the MANOVA output. We have already seen the test blocks, but we also have:

The GLM Procedure
Multivariate Analysis of Variance

E = Error SSCP Matrix			
	y1	**y2**	**y3**
y1	5052016.6667	3065116.6667	-38485
y2	3065116.6667	3002000	-45766.66667
y3	-38485	-45766.66667	2784.6666667

H = Type III SSCP Matrix for Treatment			
	y1	**y2**	**y3**
y1	13729558.333	9832783.3333	-495547.5
y2	9832783.3333	24868100	-554688.3333
y3	-495547.5	-554688.3333	31659.583333

H = Type III SSCP Matrix for Block			
	y1	**y2**	**y3**
y1	7817850	4025850	-49845
y2	4025850	2546066.6667	-58583.33333
y3	-49845	-58583.33333	2608.6666667

That means that we here have all the information we need to do univariate tests of significance!

Further these matrices have a direct connection to section 4.3.2 page 304. We see that:

Q1 : **E = Error SSCP Matrix**

Q2 : **H = Type III SSCP Matrix for Treatment**

Q3 : **H = Type III SSCP Matrix for Block**

Let us try it out. The test-statistic for treatment is given in theorem 4.26 on page 306:

$$\frac{\det(q_1)}{\det(q_1+q_2)} = \frac{1.1840 \cdot 10^{16}}{5.8515 \cdot 10^{18}} = 0.002$$

Which is the Wilks Lambda!

MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall Treatment Effect
H = Type III SSCP Matrix for Treatment
E = Error SSCP Matrix

S=3 M=-0.5 N=1

Statistic	Value	F Value	Num DF	Den DF	Pr > F
NOTE: F Statistic for Roy's Greatest Root is an upper bound.					
Wilks' Lambda	0.00202344	12.95	9	9.8856	0.0002
Pillai's Trace	2.41901846	8.33	9	18	<.0001
Hotelling-Lawley Trace	29.34255673	13.04	9	4	0.0124
Roy's Greatest Root	16.69341645	33.39	3	6	0.0004