

More ANOVA Exercises: Solutions
 STT 422: Summer, 2004
 Vince Melfi

13.24

(a) Here is the complete ANOVA table:

Source	Degrees of freedom	Sum of squares	Mean square	F
A (Gender)	1	62.40	62.40	2.73
B (Group)	1	1599.03	1599.03	69.90
AB	1	163.80	163.80	7.16
Error	596	13633.29	22.87	
Total	599	15458.52		

Note that the problem is a bit unclear about the sample size, so a total sample size of 300, with corresponding degrees of freedom of 299 (total) and 296 (error) is also acceptable. Of course, this would change the MSE from 22.87 to 46.06, and would change the F statistics accordingly. Below I'll compute the answers assuming a total sample size of 600.

(b) $F = 7.16$. Distribution assuming H_0 is true is F with 1 numerator and 596 denominator degrees of freedom. The p-value is between 0.001 and 0.01.

(c) Null distributions are the same for all three tests. For Gender, the F statistic is 2.73, and the p-value is approximately 0.1. For Group, the F statistic is 69.90, and the p-value is less than 0.001.

13.13

(a) Below is a table with the required information. The standard deviations are somewhat different, but the sample sizes are small.

The MEANS Procedure

Analysis Variable : gpi

material	time	N		Mean	Std Dev	Minimum	Maximum
		Obs	N				
ECM1	4	3	3	65.0000000	8.6602540	55.0000000	70.0000000
	8	3	3	63.3333333	2.8867513	60.0000000	65.0000000
ECM2	4	3	3	63.3333333	2.8867513	60.0000000	65.0000000
	8	3	3	63.3333333	5.7735027	60.0000000	70.0000000
ECM3	4	3	3	73.3333333	2.8867513	70.0000000	75.0000000
	8	3	3	73.3333333	5.7735027	70.0000000	80.0000000
MAT1	4	3	3	23.3333333	2.8867513	20.0000000	25.0000000
	8	3	3	21.6666667	5.7735027	15.0000000	25.0000000
MAT2	4	3	3	6.6666667	2.8867513	5.0000000	10.0000000
	8	3	3	6.6666667	2.8867513	5.0000000	10.0000000
MAT3	4	3	3	11.6666667	2.8867513	10.0000000	15.0000000
	8	3	3	10.0000000	5.0000000	5.0000000	15.0000000

(b) The plot is in Figure 1. The plot doesn't indicate much interaction.

(c) The output from `proc glm` is given below. The F statistic for interaction is 0.06 with p-value

0.9975. The F statistic for time is 0.29 with p-value 0.5950. The F statistic for material is 251.26 with p-value less than 0.0001. There seems to be no interaction, and no main effect for time, but a highly significant main effect for material.

The GLM Procedure

Dependent Variable: gpi

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	27057.63889	2459.78535	114.26	<.0001
Error	24	516.66667	21.52778		
Corrected Total	35	27574.30556			

R-Square	Coeff Var	Root MSE	gpi Mean
0.981263	11.55937	4.639804	40.13889

Source	DF	Type I SS	Mean Square	F Value	Pr > F
material	5	27045.13889	5409.02778	251.26	<.0001
time	1	6.25000	6.25000	0.29	0.5950
material*time	5	6.25000	1.25000	0.06	0.9975

Source	DF	Type III SS	Mean Square	F Value	Pr > F
material	5	27045.13889	5409.02778	251.26	<.0001
time	1	6.25000	6.25000	0.29	0.5950
material*time	5	6.25000	1.25000	0.06	0.9975

Bonferroni exercise

The output of **proc means** is given below. This will be useful in solving the exercise. In addition, the MSE, from **proc glm**, is 1.744.

The MEANS Procedure

Analysis Variable : vitc

time	N Obs	N	Mean	Std Dev	Minimum	Maximum
0	2	2	48.7050000	1.5344217	47.6200000	49.7900000
1	2	2	41.9550000	2.1283914	40.4500000	43.4600000
3	2	2	21.7950000	0.7707464	21.2500000	22.3400000
5	2	2	12.4150000	1.0818734	11.6500000	13.1800000

(a) We have $n - I = 10 - 5 = 5$ degrees of freedom, so the multiplier t^* for the 90% confidence interval is 2.015. So the interval is

$$21.795 \pm (2.015)\sqrt{1.744/2}.$$

(b) We want intervals with confidence level $1 - \alpha = 0.9$, so $\alpha = 0.1$. Since there are five intervals, the multiplier is the $1 - \alpha/(2 \times 5) = 1 - 0.1/10 = 0.99$ percentile of the t distribution with 5 degrees of freedom. This is $t^* = 4.032$. So the intervals are

CI for μ_1 : $48.705 \pm (3.364)\sqrt{1.744/2}$

CI for μ_2 : $41.955 \pm (3.364)\sqrt{1.744/2}$

CI for μ_3 : $21.795 \pm (3.364)\sqrt{1.744/2}$

CI for μ_4 : $12.415 \pm (3.364)\sqrt{1.744/2}$

CI for μ_5 : $8.320 \pm (3.364)\sqrt{1.744/2}$

(c) Again we have $n - I = 5$ degrees of freedom, so the multiplier for a 99% confidence interval is $t^* = 3.364$. The interval is

$$(48.705 - 41.955) \pm (4.032)\sqrt{1.744 \left(\frac{1}{2} + \frac{1}{2}\right)}.$$

(d) Here $1 - \alpha = 0.99$, so $\alpha = 0.01$. There are four intervals, so the multiplier is the $1 - 0.01/(2 \times 4) = 0.99875$ percentile of the t distribution with 5 degrees of freedom. This is $t^* = 5.60417$. So the intervals are

CI for $\mu_1 - \mu_2$: $(48.705 - 41.955) \pm (5.60417)\sqrt{1.744 \left(\frac{1}{2} + \frac{1}{2}\right)}$

CI for $\mu_1 - \mu_3$: $(48.705 - 21.795) \pm (5.60417)\sqrt{1.744 \left(\frac{1}{2} + \frac{1}{2}\right)}$

CI for $\mu_1 - \mu_4$: $(48.705 - 12.415) \pm (5.60417)\sqrt{1.744 \left(\frac{1}{2} + \frac{1}{2}\right)}$

CI for $\mu_1 - \mu_5$: $(48.705 - 8.320) \pm (5.60417)\sqrt{1.744 \left(\frac{1}{2} + \frac{1}{2}\right)}$.

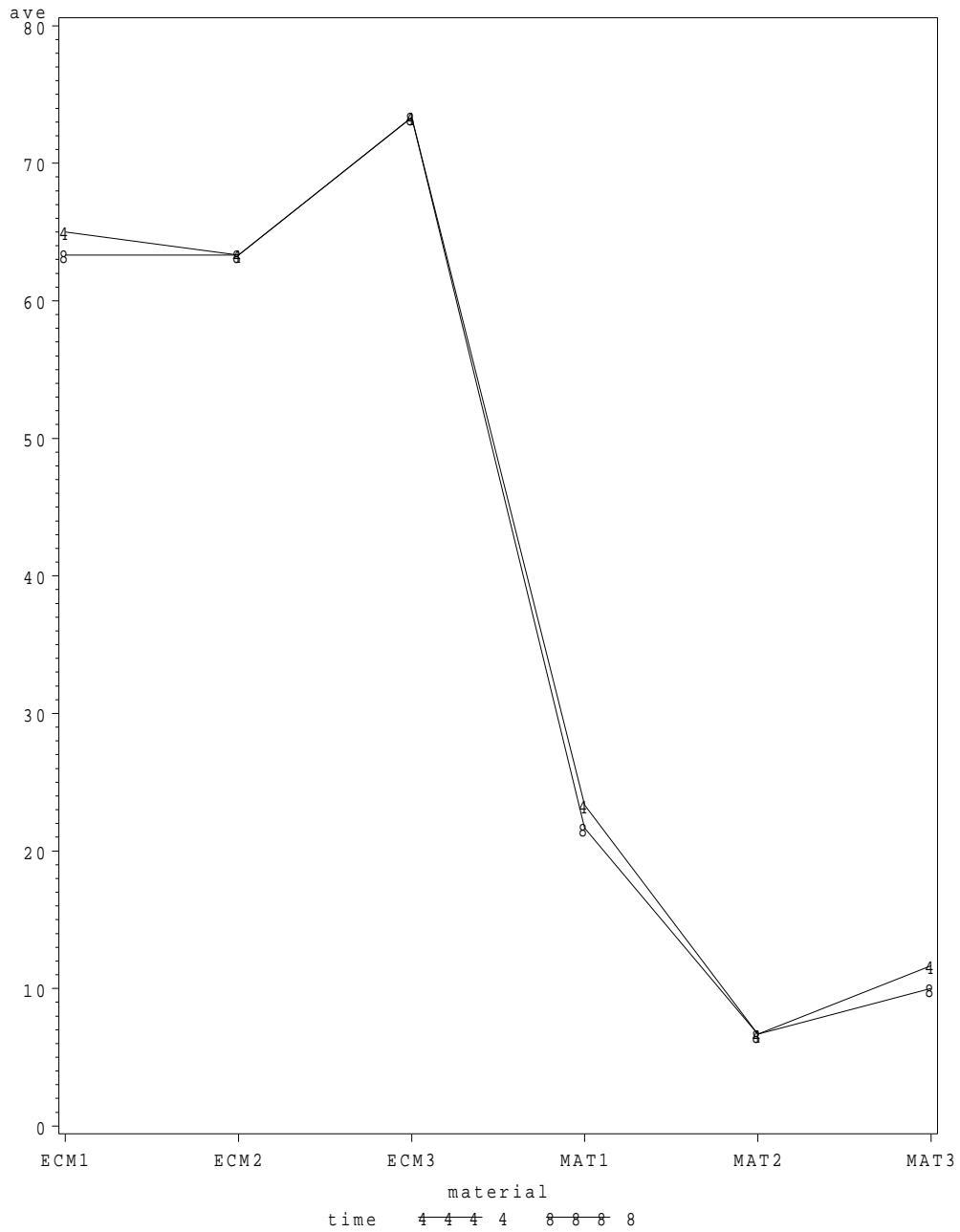


Figure 1: Interaction plot for Exercise 13.13