STAT571B
Ken Youens-Clark
Homework 7

Montgomery 13.2

(a) Analyze the data from this experiment, assuming that both parts and operators are random effects.

```
data d13_2;
input part inspector y @@;
cards;
1 1 37 1 1 38 1 1 37 1 2 41 1 2 41 1 2 40 1 3 41 1 3 42 1 3 41
2 1 42 2 1 41 2 1 43 2 2 42 2 2 42 2 2 42 2 3 43 2 3 42 2 3 43
3 1 30 3 1 31 3 1 31 3 2 31 3 2 31 3 2 31 3 3 29 3 3 30 3 3 28
4 1 42 4 1 43 4 1 42 4 2 43 4 2 43 4 2 43 4 3 42 4 3 42 4 3 42
5 1 28 5 1 30 5 1 29 5 2 29 5 2 30 5 2 29 5 3 31 5 3 29 5 3 29
6 1 42 6 1 42 6 1 43 6 2 45 6 2 45 6 2 45 6 3 44 6 3 46 6 3 45
7 1 25 7 1 26 7 1 27 7 2 28 7 2 28 7 2 30 7 3 29 7 3 27 7 3 27
8 1 40 8 1 40 8 1 40 8 2 43 8 2 42 8 2 42 8 3 43 8 3 43 8 3 41
9 1 25 9 1 25 9 1 25 9 2 27 9 2 29 9 2 28 9 3 26 9 3 26 9 3 26
10 1 35 10 1 34 10 1 34 10 2 35 10 2 35 10 2 34 10 3 35 10 3 34 10 3 35
;
run;

/* random effects model*/
proc glm data=d13_2;
class inspector part;
model y=inspector|part;
random inspector part inspector*part/test;
output out=diag r=res p=pred; run;

proc plot;
data=diag;
plot res*pred;
run;
```

The GLM Procedure
Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	29	4023.733333	138.749425	271.47	<.0001
Error	60	30.666667	0.511111		
Corrected Total	89	4054.400000			

R-Square	Coeff Var	Root MSE	y Mean
0.992436	1.996984	0.714920	35.80000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
inspector	2	39.266667	19.633333	38.41	<.0001
part	9	3935.955556	437.328395	855.64	<.0001
inspector*part	18	48.511111	2.695062	5.27	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
inspector	2	39.266667	19.633333	38.41	<.0001
part	9	3935.955556	437.328395	855.64	<.0001
inspector*part	18	48.511111	2.695062	5.27	<.0001

From the SAS output, the overall effect (top table) is highly significant (<.0001). The third table show that each factor, "inspector" and "part," are highly significant (<.0001) as well as their interaction, so for each we would reject the null hypothesis and state that they do affect the outcome.

The GLM Procedure
Tests of Hypotheses for Random Model Analysis of Variance
Dependent Variable: y

Source	DF	Type III SS	Mean Square	F Value	Pr > F
inspector	2	39.266667	19.633333	7.28	0.0048
part	9	3935.955556	437.328395	162.27	<.0001
Error	18	48.511111	2.695062		
Error: MS(inspector*part)					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
inspector*part	18	48.511111	2.695062	5.27	<.0001
Error: MS(Error)	60	30.666667	0.511111		

(b) Estimate the variance components using the analysis of variance method.

```
proc varcomp data=d13_2 method=type1;
class inspector part;
model y=inspector part inspector*part;
run;
```

Variance Components Estimation Procedure

Class Level Information		
Class	**Levels**	**Values**
inspector	3	1 2 3
part	10	1 2 3 4 5 6 7 8 9 10

Number of Observations Read	90
Number of Observations Used	90

Dependent Variable:	y

Type 1 Analysis of Variance				
Source	**DF**	**Sum of Squares**	**Mean Square**	**Expected Mean Square**
inspector	2	39.266667	19.633333	Var(Error) + 3 Var(inspector*part) + 30 Var(inspector)
part	9	3935.955556	437.328395	Var(Error) + 3 Var(inspector*part) + 9 Var(part)
inspector*part	18	48.511111	2.695062	Var(Error) + 3 Var(inspector*part)
Error	60	30.666667	0.511111	Var(Error)
Corrected Total	89	4054.400000		

Type 1 Estimates	
Variance Component	**Estimate**
Var(inspector)	0.56461
Var(part)	48.29259
Var(inspector*part)	0.72798
Var(Error)	0.51111

```
MS(E)   = 0.51
Var(T)  = [MS(A) - MS(AB)]/bn = (19.6333 - 2.695) / 10(3) =  0.5646
Var(B)  = [MS(B) - MS(AB)]/an = (437.3284 - 2.695) / 3(3) = 48.2926
Var(TB) = [MS(AB) - MS(E)]/n  = (2.695 - 0.51) / 3        =  0.7283
```

The VARCOMP method gives estimates greater than 0 for every factor, supporting the previous analysis that all are significant; however, the estimates for "inspector" and the interaction between "inspector/part" are quite small while that for "part" is quite large.

(c) Estimate the variance components using the REML method. Use the confidence intervals on the variance components to assist in drawing conclusions.

```
proc mixed data=d13_2 cl maxiter=20 covtest method=reml;
class inspector part;
model y = ;
random inspector part inspector*part; run;
```

Covariance Parameter Estimates							
Cov Parm	Estimate	Standard Error	Z Value	Pr > Z	Alpha	Lower	Upper
inspector	0.5646	0.6551	0.86	0.1944	0.05	0.1344	67.3537
part	48.2926	22.9067	2.11	0.0175	0.05	22.7624	162.52
inspector*part	0.7280	0.3011	2.42	0.0078	0.05	0.3717	2.0164
Residual	0.5111	0.09332	5.48	<.0001	0.05	0.3682	0.7575

These results match the VARCOMP almost exactly, supporting the decision to reject the null hypotheses for the parts ($p=0.0175 < \alpha=0.05$) and the interaction between inspectors and parts ($p=0.0078 < \alpha=0.05$) but to fail to reject the null hypotheses for inspector ($p=0.19 > \alpha=0.05$). Further, none of the confidence intervals (lower/upper) include 0 but the one for inspector (which we suspect has no bearing on outcome) very nearly includes 0.

Montgomery 13.7

Reanalyze the measurement system experiment in Problem 13.2, assuming that operators are a fixed factor. Estimate the appropriate model components using the ANOVA method.

```
proc mixed data=d13_2 cl maxiter=20 covtest method=type1;
class inspector part;
model y = ;
random part inspector*part; run;
```

Type 1 Analysis of Variance								
Source	DF	Sum of Squares	Mean Square	Expected Mean Square	Error Term	Error DF	F Value	Pr > F
part	9	3935.955556	437.328395	Var(Residual) + 3 Var(inspector*part) + 9 Var(part)	MS(inspector*part)	20	99.64	<.0001
inspector*part	20	87.777778	4.388889	Var(Residual) + 3 Var(inspector*part)	MS(Residual)	60	8.59	<.0001
Residual	60	30.666667	0.511111	Var(Residual)	.	.	.	.

Covariance Parameter Estimates							
Cov Parm	Estimate	Standard Error	Z Value	Pr Z	Alpha	Lower	Upper
part	48.1044	22.9070	2.10	0.0357	0.05	3.2074	93.0013
inspector*part	1.2926	0.4637	2.79	0.0053	0.05	0.3838	2.2014
Residual	0.5111	0.09332	5.48	<.0001	0.05	0.3682	0.7575

Changing the operators to fixed (leaving them out of the "random" category) increase the confidence interval for "part" (upper bound from 67 to 93) as some of the variability has been shifted to this factor. P-values for "part" and the interaction of "part"/"inspector" remain well below 0.05, so we still reject the null hypotheses for them.

Montgomery 13.26

Analyze the data in Problem 13.1, assuming that operators are fixed, using both the unrestricted and the restricted forms of the mixed models. Compare the results obtained from the two models.

Unrestricted

```
data dat;
input part operator y @@;
cards;
1 1 50 1 1 49 1 1 50 1 2 50 1 2 48 1 2 51
2 1 52 2 1 52 2 1 51 2 2 51 2 2 51 2 2 51
3 1 53 3 1 50 3 1 50 3 2 54 3 2 52 3 2 51
4 1 49 4 1 51 4 1 50 4 2 48 4 2 50 4 2 51
5 1 48 5 1 49 5 1 48 5 2 48 5 2 49 5 2 48
6 1 52 6 1 50 6 1 50 6 2 52 6 2 50 6 2 50
7 1 51 7 1 51 7 1 51 7 2 51 7 2 50 7 2 50
8 1 52 8 1 50 8 1 49 8 2 53 8 2 48 8 2 50
9 1 50 9 1 51 9 1 50 9 2 51 9 2 48 9 2 49
10 1 47 10 1 46 10 1 49 10 2 46 10 2 47 10 2 48
;
run;

proc glm data=dat;
class operator part;
model y=operator|part;
random part operator*part / test;
means operator / tukey lines E=operator*part;
lsmeans operator / adjust=tukey E=operator*part tdiff stderr;
run;
```

The GLM Procedure
Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	19	104.8500000	5.5184211	3.68	0.0003
Error	40	60.0000000	1.5000000		
Corrected Total	59	164.8500000			

R-Square	Coeff Var	Root MSE	y Mean
0.636033	2.451942	1.224745	49.95000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
operator	1	0.41666667	0.41666667	0.28	0.6011
part	9	99.01666667	11.00185185	7.33	<.0001
operator*part	9	5.41666667	0.60185185	0.40	0.9270

Source	DF	Type III SS	Mean Square	F Value	Pr > F
operator	1	0.41666667	0.41666667	0.28	0.6011
part	9	99.01666667	11.00185185	7.33	<.0001
operator*part	9	5.41666667	0.60185185	0.40	0.9270

The p-values in the above ANOVA table for "operator" (0.6) and "operator/part" (0.927) are well above α=0.05, so we accept the null hypotheses for those factors, but the p-value for "part" (<0.0001) is well below, so we reject the null hypothesis for this factor and state that it has a significant effect.

We need to incorporate additional tests. I got very confused on this part. As I understand the lecture notes, we need to compute the following:

Testing hypotheses:

$$H_0 : \tau_1 = \tau_2 = \ldots = 0 \rightarrow \mathrm{MS_A}/\mathrm{MS_{AB}}$$
$$H_0 : \sigma_\beta^2 = 0 \rightarrow \mathrm{MS_B}/\mathrm{MS_E}$$
$$H_0 : \sigma_{\tau\beta}^2 = 0 \rightarrow \mathrm{MS_{AB}}/\mathrm{MS_E}$$

```
"operator"    H0: F = MS(A)/MS(AB) = 0.4167/0.602 = 0.6914
"part"        H0: F = MS(B)/MS(E)  = 11.0019/1.5  = 7.3346
"interaction" H0: F = MS(AB)/MS(E) = 0.6019/1.5   = 0.4013
```

But I think that is for the "restricted" model, in which case we need to make the change to the test for beta:

Test $H_0 : \sigma_\beta^2 = 0$ using $\mathrm{MS_{AB}}$ in denominator

Which puts the F value much higher:

```
"part"        H0: F = MS(B)/MS(AB) = 11.0019/0.602 = 18.2756
```

The first F value for "part" matches that the ANOVA as does the one for the interaction; while the F value for "operator" is slightly higher, the resulting p-value (0.28 vs 0.69) is still high enough to support the above conclusions.

Restricted

```
proc mixed data=dat alpha=0.05 cl covtest;
class operator part;
model y=operator / ddfm=kr;
random part operator*part;
lsmeans operator / alpha=0.05 cl diff adjust=tukey; run;
```

Covariance Parameter Estimates							
Cov Parm	Estimate	Standard Error	Z Value	Pr > Z	Alpha	Lower	Upper
part	1.6111	0.8655	1.86	0.0313	0.05	0.7020	6.7389
operator*part	0	.	.	.	.	.	.
Residual	1.3350	0.2697	4.95	<.0001	0.05	0.9316	2.0731

The confidence intervals for "part" does not include 0, so we can assume that this factor does affect the outcome.

Least Squares Means									
Effect	operator	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
operator	1	50.0333	0.4534	11.3	110.34	<.0001	0.05	49.0384	51.0283
operator	2	49.8667	0.4534	11.3	109.97	<.0001	0.05	48.8717	50.8616

The numbers for the two operators across this chart are extremely close, therefore we could assume that there is very little variation caused by the operators, a finding supported by our failure to reject the null hypothesis for this factor above.