STAT 514 Homework 9

Due: Nov 20

1. An experiment was performed to investigate the capability of a measurement system. Ten parts were randomly selected, and two randomly selected operators measured each part three times. The tests were made in random order, and the data below resulted. You can also find data set file "measure.dat".

	Operator 1			Operator 2		
Parts	1	2	3	1	2	3
1	50	49	50	50	48	51
2	52	52	51	51	51	51
3	53	50	50	54	52	51
4	49	51	50	48	50	51
5	48	49	48	48	49	48
6	52	50	50	52	50	50
7	51	51	51	51	50	50
8	52	50	49	53	48	50
9	50	51	50	51	48	49
10	47	46	49	46	47	48

   (a) Test the variance components.
   solution:
   Here consider all two factors are random effects. We can use `test` options in `glm` procedure.

```
 Test of Hypotheses for Random Model Analysis of Variance
                Dependent Variable: measure

Source             DF   Type III SS   Mean Squares   F Value   Pr > F
operator            1      0.416667       0.416667      0.69   0.4269
part                9     99.016667      11.001852     18.28   <.0001
Error               9      5.416667       0.601852
Error: MS(operator*part)

Source             DF   Type III SS   Mean Squares   F Value   Pr > F
operator*part       9      5.416667       0.601852      0.40   0.9270
Error              40     60.000000       1.500000
Error: MS(Error)
```

Test for "part" has p-value smaller than 0.0001, we reject the Null hypothesis ($H_0 : \sigma_\tau^2 = 0$).
Test for "operator" has p-value 0.4269, we fail to reject the Null hypothesis ($H_0 : \sigma_\beta^2 = 0$).
Test for "part-operator" interaction has p-value 0.9270, we fail to reject the Null hypothesis ($H_0 : \sigma_{\tau\beta}^2 = 0$) .

(b) Find the estimates of the variance components using the analysis of variance method.
solution:
We can either computer by hand using the corrected MS term, or obtain the point estimates by `varcomp` procedure with statement `method="type 1"`.

```
       Type I Estimates
Variance Component         Estimate
Var(operator)              -0.0061728
Var(part)                   1.733333
Var(operator*part)          -0.29938
Var(Error)                   1.50000
```

For the negative estimates, we can reduce them to zero, since negative variance is not allowed.

2. Analyze the data in previous problem, assuming that the operators are fixed, using the restricted form of the mixed models.

(a) Test the variance components.
solution:
SAS can not directly handle restricted model, we obtain the ANOVA table from SAS output:

```
      Dependent Variable: measure
                       Sum of
Source          DF     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

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
```

Test for "operators":

$$F0 = MS_A/MS_{AB} = 0.4167/0.6018 = 0.692.$$

P-value is 0.4269. As the P-value is large we fail to reject null hypothesis. The fixed effect "operators" is not significant.

Test for "parts":

$$F0 = MS_B/MS_E == 11.00185/1.5 = 7.33.$$

P-value is 0.4269. As the P-value is less than 0.0001 we reject null hypothesis. The random effect "parts" is significant.
Test for "interaction":

$$F0 = MS_{AB}/MS_E = 0.40.$$

As the P-value is very big (much greater than 0.05) we fail to reject H0 and conclude that the effect due to "interaction" is not significant.

(b) Find the estimates of the variance components using the analysis of variance method.
solution:
The variance component estimates are: $\hat{\sigma}^2_\beta = (11.00185 - 1.5)/(2*3) = 1.584.$ and $\hat{\sigma}^2_{\tau\beta} = (0.60185 - 1.5)/3 = -0.299(\approx 0)$.

(c) Find an exact 95 percent confidence interval on σ^2 .
solution:
The exact 95% CI on σ^2 : $(df_E MS_E/\chi^2_{0.05/2,40}, df_E MS_E/\chi^2_{1-0.05/2,40}) = (40 \times 1.5/59.34, 40 \times 1.5/24.43) = (1.011, 2.456)$.

(d) Find approximate 95 percent confidence intervals on the variance components using the Satterthwaite method.
solution:
$\hat{\sigma}^2_\beta = (11.00185 - 1.5)/(2 \times 3) = 1.584..$
Using Satterthwaite's method the $df_\beta = (MS_B - MS_E)^2/(MS_B^2/df_B + MS_E^2/df_E) = (11.00185 - 1.5)^2/((11.00185)^2/9 + (1.5)^2/40) = 6.6852$.
Also from SAS $\chi^2_{0.025,6.6852} = 15.5256, \chi^2_{0.975,6.6852} = 1.5430$.
Hence the 95% approximate CI on σ^2_β can be given by:
$(df_\beta \hat{\sigma}^2_\beta/\chi^2_{0.05/2,6.6852}, df_\beta \hat{\sigma}^2_\beta/\chi^2_{1-0.05/2,6.6852})$
$= (6.6852 \times 1.584/15.5256, 6.6852 \times 1.584/1.5430) = (0.6821, 6.8628)$.

3. Consider a balanced three-factor ANOVA study with factors A, B, and C. Suppose both A and B have fixed effects but C has random effects. Following "Rules for Expected Mean Squares", work out the EMS table, and propose an F-test for each set of the main effects. Assume that we use the restricted mixed effects model.
solution:
Denote $q(\tau) = \frac{\Sigma\tau_i^2}{a-1}$, $q(\beta) = \frac{\Sigma\beta_i^2}{b-1}$, $q(\tau\beta) = \frac{\Sigma\tau\beta_i^2}{(a-1)(b-1)}$. The EMS term is listed in the table.
The F-tests for the main effects are:
Test H0 : all τ_i are zero vs. H1 : at least one of τ_i is not zero, with $F = MS_A/MS_{AC}$, which follows $F_{a-1,(a-1)(c-1)}$ under the null hypothesis.
Test H0 : all β_j are zero vs. H1 : at least one of β_j is not zero, with $F = MS_B/MS_{BC}$, which follows $F_{b-1,(b-1)(c-1)}$ under the null hypothesis.
Test H0 : σ^2_γ are zero vs. H1 : σ^2_γ is not zero, with $F = MS_C/MS_E$, which follows $F_{c-1,abc(n-1)}$ under the null hypothesis.

term	F a i	F b j	R c k	R n l	EMS
τ_i	0	b	c	n	$\sigma^2 + bn\sigma^2_{\tau\gamma} + bcn * q(\tau)$
β_j	a	0	c	n	$\sigma^2 + an\sigma^2_{\beta\gamma} + acn * q(\beta)$
γ_k	a	b	1	n	$\sigma^2 + abn\sigma^2_{\gamma}$
$(\tau\beta)_{ij}$	0	0	c	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn * q(\tau\beta)$
$(\tau\gamma)_{ik}$	0	b	1	n	$\sigma^2 + bn\sigma^2_{\tau\gamma}$
$(\beta\gamma)_{jk}$	a	0	1	n	$\sigma^2 + an\sigma^2_{\beta\gamma}$
$(\tau\beta\gamma)_{ijk}$	0	0	1	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma}$
ϵ_{ijkl}	1	1	1	1	σ^2

4. A rocket propellant manufacturer is studying the burning rate of propellant from three production processes. Four batches of propellant are randomly selected from the output of each process, and three determinations of burning rate are made on each batch. The results follow. You may also find data set file "rocket.dat".

	Process 1				Process 2				Process 3			
Batch	1	2	3	4	1	2	3	4	1	2	3	4
	25	19	15	15	19	23	18	35	14	35	38	25
	30	28	17	16	17	24	21	27	15	21	54	29
	26	20	14	13	14	21	17	25	20	24	50	33

(a) Explain why "batch" is nested under "process".
solution:
The reason is: (i) "batch" has four levels at each level of "process"; (ii) under the same level of "process", the levels of "batch" are comparable; (iii) under a level of "process", the levels of "batch" can be arbitrarily numbered (i.e., the levels of "batch" from different levels of "process" are not comparable).

(b) Analyze the data and draw conclusions.
solution:
The process is fixed, and batch is random effect nested within process. From SAS ouput, we have

```
                         Type 1 Analysis of Variance
Source      DF Sum of Squares  Mean Square Expected Mean Square
proc        2  676.055556      338.027778  Var(Residual) + 3 Var(batch(proc)) + Q(proc)
batch(proc) 9  2077.583333     230.842593  Var(Residual) + 3 Var(batch(proc))
Residual    24 454.000000      18.916667   Var(Residual)

Error Term Error DF F Value Pr > F
MS(batch(proc))  9  1.46    0.2815
MS(Residual)     24 12.20   <.0001
.                .  .       .


Covariance Parameter Estimates
Cov Parm        Estimate
batch(proc)     70.6420
Residual        18.9167
```

We fail to reject the test for process, reject the test for batches. Hence, our conclusion is:
Process are not different, variability due to random batches.