Alexandra Chang

Homework 5

1. The Candle Experiment

An experiment was run to determine whether different colored candles (red, white, blue, yellow) burn at different speeds. Each experimenter collected four observations on each color in a random order, and "experimenter" was used as a blocking factor. The design was a general complete block design with $v = 4$, $k = 16$, $b = 4$, and $s = 4$. The resulting burning times (in seconds) are shown in Table 17.21 in the book and can be downloaded at `http://deanvossdraguljic.ietsandbox.net/DeanVossDraguljic/SASdata/candle.sas`.

(a) [2 pts] Analyze the experiment as though the experimenters represent a random sample from a large population of people who might use these candles in practice. Use a two-way mixed model with interaction where the experimenters are blocks with random effects and the interactions of block and color are random as well. The color has fixed effects. Provide the 95% simultaneous confidence intervals for all pairwise comparisons of color using Tukey's method.

(b) [1 pt] Provide the SAS code for 1(a).

(c) [2 pts] In 1(a), suppose we do not treat these experimenters as a random sample from a large population and consequently the model does not have random effects. Provide the 95% simultaneous confidence intervals for all pairwise comparisons of color using Tukey's method.

(d) [1 pt] Provide the SAS code for 1(c).

(e) [2 pts] Which model provides shorter confidence intervals? Give a justification if you could.

(f) [2 pts] If a two-way main-effects model was used in 1(a) and 1(c), do you think the two models (one with random effects and one without) would produce the simultaneous confidence intervals of different lengths?

2. Completely Randomized Design with Three Factors

Suppose in a completely randomized designs experiment there are 3 crossed factors A, B and C. Their levels are $a = 4$, $b = 3$ and $c = 3$. Each treatment is replicated $r = 2$ times. The following model is applied to the experiment:

$$Y_{ijkt} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijkt}, \quad i = 1, 2, 3, 4; \quad j, k = 1, 2, 3,$$

where ϵ_{ijkt}'s are i.i.d. $N(0, \sigma^2)$, α_i's represent the fixed effects of A, β_j's and γ_k's are the random effects of Factors B and C, respectively, and $(\beta\gamma)_{jk}$'s are the random interaction effects. Furthermore, assume all random terms are independent and

$$\beta_j \sim N(0, \sigma_B^2), \quad \gamma_k \sim N(0, \sigma_C^2), \quad (\beta\gamma)_{jk} \sim N(0, \sigma_{BC}^2).$$

Partial SAS output is provided below:

Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	4.20183285	0.38198480	0.35	0.9712
Error	60	66.31427369	1.10523789		
Corrected Total	71	70.51610654			

R-Square	Coeff Var	Root MSE	y Mean
0.059587	10.43951	1.051303	10.07043

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	3	0.95832853	0.31944284	0.29	0.8331
B	2	0.34514265	0.17257132	0.16	0.8558
C	2	0.10863270	0.05431635	0.05	0.9521
B*C	4	2.78972897	0.69743224	0.63	0.6423

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	3	0.95832853	0.31944284	0.29	0.8331
B	2	0.34514265	0.17257132	0.16	0.8558
C	2	0.10863270	0.05431635	0.05	0.9521
B*C	4	2.78972897	0.69743224	0.63	0.6423

(a) [1 pt] The F-ratio for testing the hypothesis that A has equal main effects is ____________.

(b) [1 pt] The degree of freedom of the denominator for the hypothesis testing in the previous question is ____________.

(c) [1 pt] The F-ratio for testing the hypothesis that B has no effects is ____________.

(d) [1 pt] The degree of freedom of the denominator for the hypothesis testing in the previous question is ____________.

1. The Candle Experiment

An experiment was run to determine whether different colored candles (red, white, blue, yellow) burn at different speeds. Each experimenter collected four observations on each color in a random order, and "experimenter" was used as a blocking factor. The design was a general complete block design with $v = 4$, $k = 16$, $b = 4$, and $s = 4$. The resulting burning times (in seconds) are shown in Table 17.21 in the book and can be downloaded at `http://deanvossdraguljic.ietsandbox.net/DeanVossDraguljic/SASdata/candle.sas`.

Table 17.21 Data for the candle experiment (seconds)

Person	Color							
	Red		White		Blue		Yellow	
1	989	1032	1044	979	1011	951	974	998
	1077	1019	987	1031	928	1022	1033	1041
2	899	912	847	880	899	800	886	859
	911	943	879	830	820	812	901	907
3	898	840	840	952	909	790	950	992
	955	1005	961	915	871	905	920	890
4	993	957	987	960	864	925	949	973
	1005	982	920	1001	824	790	978	938

(a) [2 pts] Analyze the experiment as though the experimenters represent a random sample from a large population of people who might use these candles in practice. Use a two-way mixed model with interaction where the experimenters are blocks with random effects and the interactions of block and color are random as well. The color has fixed effects. Provide the 95% simultaneous confidence intervals for all pairwise comparisons of color using Tukey's method.

The Mixed Procedure

Model Information	
Data Set	WORK.CANDLE
Dependent Variable	time
Covariance Structure	Variance Components
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment

Class Level Information

Class	Levels	Values
person	4	1 2 3 4
color	4	1 2 3 4

Dimensions	
Covariance Parameters	3
Columns in X	5
Columns in Z	20
Subjects	1
Max Obs per Subject	64

Number of Observations	
Number of Observations Read	64
Number of Observations Used	64
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	681.33629906	
1	1	638.39368124	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
person	3049.70
person*color	12.2483
Residual	1708.85

Fit Statistics	
-2 Res Log Likelihood	638.4
AIC (Smaller is Better)	644.4
AICC (Smaller is Better)	644.8
BIC (Smaller is Better)	642.6

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
color	3	9	11.44	0.0020

Least Squares Means

Effect	color	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
color	1	963.56	29.5346	9	32.62	<.0001	0.05	896.75	1030.37
color	2	938.31	29.5346	9	31.77	<.0001	0.05	871.50	1005.12
color	3	882.56	29.5346	9	29.88	<.0001	0.05	815.75	949.37
color	4	949.31	29.5346	9	32.14	<.0001	0.05	882.50	1016.12

Differences of Least Squares Means

Effect	color	_color	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
color	1	2	25.2500	14.8233	9	1.70	0.1227	Tukey-Kramer	0.3757	0.05	-8.2827	58.7827	-21.0253	71.5253
color	1	3	81.0000	14.8233	9	5.46	0.0004	Tukey-Kramer	0.0018	0.05	47.4673	114.53	34.7247	127.28
color	1	4	14.2500	14.8233	9	0.96	0.3615	Tukey-Kramer	0.7739	0.05	-19.2827	47.7827	-32.0253	60.5253
color	2	3	55.7500	14.8233	9	3.76	0.0045	Tukey-Kramer	0.0192	0.05	22.2173	89.2827	9.4747	102.03
color	2	4	-11.0000	14.8233	9	-0.74	0.4770	Tukey-Kramer	0.8778	0.05	-44.5327	22.5327	-57.2753	35.2753
color	3	4	-66.7500	14.8233	9	-4.50	0.0015	Tukey-Kramer	0.0066	0.05	-100.28	-33.2173	-113.03	-20.4747

(b) [1 pt] Provide the SAS code for 1(a).

```
proc mixed data=candle;
  class person color;
  model time = color;
  random person person*color;
  lsmeans color / cl pdiff adjust=tukey;
run;
```

```
data candle;
  do person = 1 to 4;
    do row = 1,2;
      do color = 1 to 4;
        do col = 1,2;
          input time @@;
          output;
          drop row col;
        end;
      end;
    end;
  end;
  lines;
   989 1032 1044  979 1011  951  974  998
  1077 1019  987 1031  928 1022 1033 1041
   899  912  847  880  899  800  886  859
   911  943  879  830  820  812  901  907
   898  840  840  952  909  790  950  992
   955 1005  961  915  871  905  920  890
   993  957  987  960  864  925  949  973
  1005  982  920 1001  824  790  978  938
;
run;
```

(c) [2 pts] In 1(a), suppose we do not treat these experimenters as a random sample from a large population and consequently the model does not have random effects. Provide the 95% simultaneous confidence intervals for all pairwise comparisons of color using Tukey's method.

The GLM Procedure

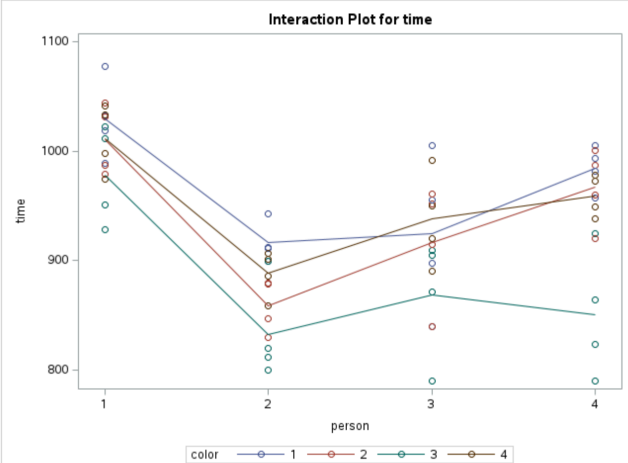
Dependent Variable: time

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	227824.7500	15188.3167	8.89	<.0001
Error	48	82025.0000	1708.8542		
Corrected Total	63	309849.7500			

R-Square	Coeff Var	Root MSE	time Mean
0.735275	4.428608	41.33829	933.4375

Source	DF	Type I SS	Mean Square	F Value	Pr > F
person	3	151659.1250	50553.0417	29.58	<.0001
color	3	60345.0000	20115.0000	11.77	<.0001
person*color	9	15820.6250	1757.8472	1.03	0.4315

Source	DF	Type III SS	Mean Square	F Value	Pr > F
person	3	151659.1250	50553.0417	29.58	<.0001
color	3	60345.0000	20115.0000	11.77	<.0001
person*color	9	15820.6250	1757.8472	1.03	0.4315



The GLM Procedure

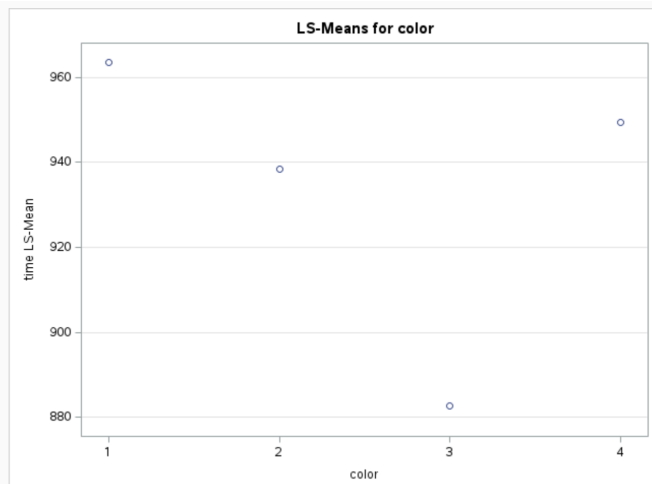
Source	Type III Expected Mean Square
person	Var(Error) + 4 Var(person*color) + 16 Var(person)
color	Var(Error) + 4 Var(person*color) + Q(color)
person*color	Var(Error) + 4 Var(person*color)

The GLM Procedure
Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: time

Source	DF	Type III SS	Mean Square	F Value	Pr > F
person	3	151659	50553	28.76	<.0001
color	3	60345	20115	11.44	0.0020
Error	9	15821	1757.847222		
Error: MS(person*color)					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
person*color	9	15821	1757.847222	1.03	0.4315
Error: MS(Error)	48	82025	1708.854167		



Least Squares Means
Adjustment for Multiple Comparisons: Tukey

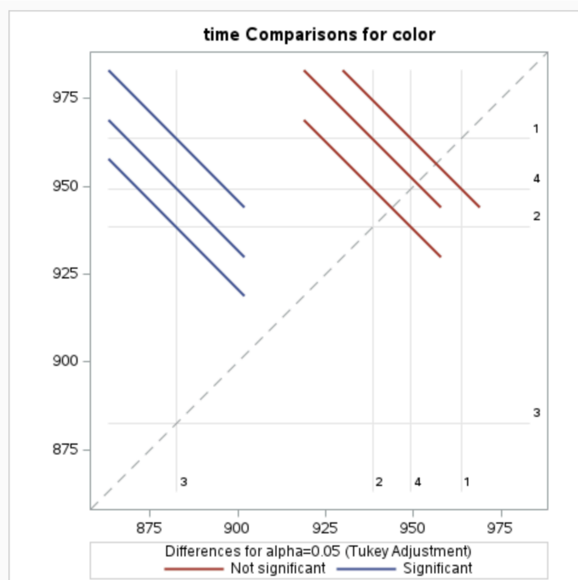
color	time LSMEAN	LSMEAN Number
1	963.562500	1
2	938.312500	2
3	882.562500	3
4	949.312500	4

Least Squares Means for effect color
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: time

i/j	1	2	3	4
1		0.3209	<.0001	0.7642
2	0.3209		0.0021	0.8751
3	<.0001	0.0021		0.0002
4	0.7642	0.8751	0.0002	

color	time LSMEAN	95% Confidence Limits	
1	963.562500	942.783450	984.341550
2	938.312500	917.533450	959.091550
3	882.562500	861.783450	903.341550
4	949.312500	928.533450	970.091550



Least Squares Means for Effect color

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	25.250000	-13.646740	64.146740
1	3	81.000000	42.103260	119.896740
1	4	14.250000	-24.646740	53.146740
2	3	55.750000	16.853260	94.646740
2	4	-11.000000	-49.896740	27.896740
3	4	-66.750000	-105.646740	-27.853260

fixed effects model

(d) [1 pt] Provide the SAS code for 1(c).

```
proc glm data=candle;
  class person color;
  model time = person color person*color;
  random person person*color/test;
  lsmeans color / cl pdiff adjust=tukey E=person*color;
run;
```

Assume no random effects

(e) [2 pts] Which model provides shorter confidence intervals? Give a justification if you could.

The model in 1(c) provides shorter intervals. That is the model where the experimenter effects were treated as fixed.

(f) [2 pts] If a two-way main-effects model was used in 1(a) and 1(c), do you think the two models (one with random effects and one without) would produce the simultaneous confidence intervals of different lengths?

No! In this case, they will produce the intervals of the same length. This is because in both models the degrees of freedom for error is used.

(next page for Q2)

↓

2. Completely Randomized Design with Three Factors

Mixed-Effects Model

Suppose in a completely randomized designs experiment there are 3 crossed factors A, B and C. Their levels are $a = 4$, $b = 3$ and $c = 3$. Each treatment is replicated $r = 2$ times. The following model is applied to the experiment:

$$Y_{ijkt} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijkt}, \quad i = 1, 2, 3, 4; \quad j, k = 1, 2, 3,$$

where ϵ_{ijkt}'s are i.i.d. $N(0, \sigma^2)$, α_i's represent the fixed effects of A, β_j's and γ_k's are the random effects of Factors B and C, respectively, and $(\beta\gamma)_{jk}$'s are the random interaction effects. Furthermore, assume all random terms are independent and

$$\beta_j \sim N(0, \sigma_B^2), \quad \gamma_k \sim N(0, \sigma_C^2), \quad (\beta\gamma)_{jk} \sim N(0, \sigma_{BC}^2).$$

Fixed: A

Random: B, C, Therefore, BC is random

Partial SAS output is provided below:

Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	4.20183285	0.38198480	0.35	0.9712
Error	60	66.31427369	1.10523789		
Corrected Total	71	70.51610654			

R-Square	Coeff Var	Root MSE	y Mean
0.059587	10.43951	1.051303	10.07043

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	3	0.95832853	0.31944284	0.29	0.8331
B	2	0.34514265	0.17257132	0.16	0.8558
C	2	0.10863270	0.05431635	0.05	0.9521
B*C	4	2.78972897	0.69743224	0.63	0.6423

✓

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	3	0.95832853	0.31944284	0.29	0.8331
B	2	0.34514265	0.17257132	0.16	0.8558
C	2	0.10863270	0.05431635	0.05	0.9521
B*C	4	2.78972897	0.69743224	0.63	0.6423

Effect	EMS	F-ratio
A	$Q(A) + \sigma^2$	msA / msE
B	$\sigma_B^2 + c\sigma_{BC}^2 + \sigma^2$	ms B / ms BC
C	$\sigma_C^2 + b\sigma_{BC}^2 + \sigma^2$	msC / ms BC
BC	$\sigma_{BC}^2 + \sigma^2$	msBC / ms E
Error	σ^2	–

(a) [1 pt] The F-ratio for testing the hypothesis that A has equal main effects is msA / ms E. ✓ 0.29

(b) [1 pt] The degree of freedom of the denominator for the hypothesis testing in the previous question is 60.

(c) [1 pt] The F-ratio for testing the hypothesis that B has no effects is ms B / ms BC. ✓ 0.25

(d) [1 pt] The degree of freedom of the denominator for the hypothesis testing in the previous question is 4.

Homework 5

1. The Candle Experiment

An experiment was run to determine whether different colored candles (red, white, blue, yellow) burn at different speeds. Each experimenter collected four observations on each color in a random order, and "experimenter" was used as a blocking factor. The design was a general complete block design with $v = 4$, $k = 16$, $b = 4$, and $s = 4$. The resulting burning times (in seconds) are shown in Table 17.21 in the book and can be downloaded at `http://deanvossdraguljic.ietsandbox.net/DeanVossDraguljic/SASdata/candle.sas`.

(a) [2 pts] Analyze the experiment as though the experimenters represent a random sample from a large population of people who might use these candles in practice. Use a two-way mixed model with interaction where the experimenters are blocks with random effects and the interactions of block and color are random as well. The color has fixed effects. Provide the 95% simultaneous confidence intervals for all pairwise comparisons of color using Tukey's method.

The last two columns are for the simultaneous confidence intervals.

Differences of Least Squares Means

Effect	color	_color	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
color	1	2	25.2500	14.8233	9	1.70	0.1227	Tukey-Kramer	0.3757	0.05	-8.2827	58.7827	-21.0253	71.5253
color	1	3	81.0000	14.8233	9	5.46	0.0004	Tukey-Kramer	0.0018	0.05	47.4673	114.53	34.7247	127.28
color	1	4	14.2500	14.8233	9	0.96	0.3615	Tukey-Kramer	0.7739	0.05	-19.2827	47.7827	-32.0253	60.5253
color	2	3	55.7500	14.8233	9	3.76	0.0045	Tukey-Kramer	0.0192	0.05	22.2173	89.2827	9.4747	102.03
color	2	4	-11.0000	14.8233	9	-0.74	0.4770	Tukey-Kramer	0.8778	0.05	-44.5327	22.5327	-57.2753	35.2753
color	3	4	-66.7500	14.8233	9	-4.50	0.0015	Tukey-Kramer	0.0066	0.05	-100.28	-33.2173	-113.03	-20.4747

(b) [1 pt] Provide the SAS code for 1(a).

```
data candle;
    do person=1 to 4;
        do row=1,2;
            do color=1 to 4;
                do col=1,2;
                input time @@;
                output;
                drop row col;
    end; end; end; end;
    lines;
    989 1032 1044 979 1011 951 974 998
    1077 1019 987 1031 928 1022 1033 1041
    899 912 847 880 899 800 886 859
    911 943 879 830 820 812 901 907
    898 840 840 952 909 790 950 992
    955 1005 961 915 871 905 920 890
    993 957 987 960 864 925 949 973
```

```
        1005 982 920 1001 824 790 978 938
        ;
run;

/* 1(a)-1(b)*/
proc mixed data=candle;
    class person color;
    model time=color;
    random person person*color;
    lsmeans color /cl pdiff adjust=Tukey;
run;

/* Use proc glm to get the same answers */
proc glm data=candle;
    class person color;
    model time=person color person*color;
    random person person*color/test;
    lsmeans color /cl pdiff adjust=Tukey E=person*color;
run;
```

(c) [2 pts] In 1(a), suppose we do not treat these experimenters as a random sample from a large population and consequently the model does not have random effects. Provide the 95% simultaneous confidence intervals for all pairwise comparisons of color using Tukey's method.

The last two columns are for the simultaneous confidence intervals.

Least Squares Means for Effect color				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	25.250000	-13.646740	64.146740
1	3	81.000000	42.103260	119.896740
1	4	14.250000	-24.646740	53.146740
2	3	55.750000	16.853260	94.646740
2	4	-11.000000	-49.896740	27.896740
3	4	-66.750000	-105.646740	-27.853260

(d) [1 pt] Provide the SAS code for 1(c).

```
/* 1(c)-(d) fixed effects model */
proc glm data=candle;
    class person color;
    model time=person color person*color;
    lsmeans color /cl pdiff adjust=Tukey;
run;
```

(e) [2 pts] Which model provides shorter confidence intervals? Give a justification if you could.

The model in 1(c) provides shorter intervals. That is the model where the experimenter effects were treated as fixed.

(f) [2 pts] If a two-way main-effects model was used in 1(a) and 1(c), do you think the two models (one with random effects and one without) would produce the simultaneous confidence intervals of different lengths?

No! In this case, they will produce the intervals of the same length. This is because in both models the degrees of freedom for error is used.

2. Completely Randomized Design with Three Factors

Suppose in a completely randomized designs experiment there are 3 crossed factors A, B and C. Their levels are $a = 4$, $b = 3$ and $c = 3$. Each treatment is replicated $r = 2$ times. The following model is applied to the experiment:

$$Y_{ijkt} = \mu + \alpha_i + \beta_j + \gamma_k + (\beta\gamma)_{jk} + \epsilon_{ijkt}, \quad i = 1, 2, 3, 4; \quad j, k = 1, 2, 3,$$

where ϵ_{ijkt}'s are i.i.d. $N(0, \sigma^2)$, α_i's represent the fixed effects of A, β_j's and γ_k's are the random effects of Factors B and C, respectively, and $(\beta\gamma)_{jk}$'s are the random interaction effects. Furthermore, assume all random terms are independent and

$$\beta_j \sim N(0, \sigma_B^2), \quad \gamma_k \sim N(0, \sigma_C^2), \quad (\beta\gamma)_{jk} \sim N(0, \sigma_{BC}^2).$$

Partial SAS output is provided below:

Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	11	4.20183285	0.38198480	0.35	0.9712
Error	60	66.31427369	1.10523789		
Corrected Total	71	70.51610654			

R-Square	Coeff Var	Root MSE	y Mean
0.059587	10.43951	1.051303	10.07043

Source	DF	Type I SS	Mean Square	F Value	Pr > F
A	3	0.95832853	0.31944284	0.29	0.8331
B	2	0.34514265	0.17257132	0.16	0.8558
C	2	0.10863270	0.05431635	0.05	0.9521
B*C	4	2.78972897	0.69743224	0.63	0.6423

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	3	0.95832853	0.31944284	0.29	0.8331
B	2	0.34514265	0.17257132	0.16	0.8558
C	2	0.10863270	0.05431635	0.05	0.9521
B*C	4	2.78972897	0.69743224	0.63	0.6423

(a) [1 pt] The F-ratio for testing the hypothesis that A has equal main effects is 0.29.

(b) [1 pt] The degree of freedom of the denominator for the hypothesis testing in the previous question is 60.

(c) [1 pt] The F-ratio for testing the hypothesis that B has no effects is 0.25.

(d) [1 pt] The degree of freedom of the denominator for the hypothesis testing in the previous question is 4.