Total = 317

STA 6208 – Spring 2016 – Exam 2 **PRINT** Name ANSWER KEY

Note: Conduct all tests at at $\alpha = 0.05$ significance level and show all work for any partial credit on numeric (calculation) problems.

Satterthwaite's Approximation: $\hat{\sigma}_*^2 = \sum_{i=1}^{c} g_i MS_i \Rightarrow \hat{\nu}_* = \dfrac{\left(\sum_{i=1}^{c} g_i MS_i\right)^2}{\sum_{i=1}^{c} \dfrac{(g_i MS_i)^2}{\nu_i}}$ where $\nu_i = df(MS_i)$

Q.1. A study was conducted, regarding 2 Random factors: A: Wine Judge ($a = 9$) and B: Wine Brand ($b = 10$). Each judge rated each brand once, blind to the brand label. The model fit is:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} \quad i = 1,\ldots,9;\ j = 1,\ldots,10 \quad \alpha_i \sim NID(0,\sigma_\alpha^2) \quad \beta_j \sim NID(0,\sigma_\beta^2) \quad \varepsilon_{ij} \sim NID(0,\sigma^2) \quad \{\alpha\} \perp \{\beta\} \perp \{\varepsilon\}$$

p.1.a. For this model, $SS_{\text{ERR}} = \sum_i \sum_j \left(Y_{ij} - \overline{Y}_{i\bullet} - \overline{Y}_{\bullet j} + \overline{Y}_{\bullet\bullet}\right)^2 = \sum_i \sum_j Y_{ij}^2 - b\sum_i \overline{Y}_{i\bullet}^2 - a\sum_j \overline{Y}_{\bullet j}^2 + ab\overline{Y}_{\bullet\bullet}^2 \qquad MS_{\text{ERR}} = \dfrac{SS_{\text{ERR}}}{(a-1)(b-1)}$

2 Each

p.1.a.i.

$E\{Y_{ij}\} = \mu + 0 + 0 + 0 = \mu$ $\quad E\{\overline{Y}_{i\bullet}\} = \mu$ $\quad E\{\overline{Y}_{\bullet j}\} = \mu$ $\quad E\{\overline{Y}_{\bullet\bullet}\} = \mu$

4 8 8 8

p.1.a.ii. **Derive** $V\{Y_{ij}\}$, $V\{\overline{Y}_{i\bullet}\}$, $V\{\overline{Y}_{\bullet j}\}$, $V\{\overline{Y}_{\bullet\bullet}\}$ **Be very specific on all parts of derivation.**

$V\{Y_{ij}\} = V\{\mu + \alpha_i + \beta_j + \varepsilon_{ij}\} = 0 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma^2$

$$Cov\{Y_{ij}, Y_{i'j'}\} = \begin{cases} \sigma_\alpha^2 + \sigma_\beta^2 + \sigma^2 & i = i',\ j = j' \\ \sigma_\alpha^2 & i = i',\ j \neq j' \\ \sigma_\beta^2 & i \neq i',\ j = j' \\ 0 & i \neq i',\ j \neq j' \end{cases}$$

$a = 9, b = 10$

$V\{Y_{i\cdot}\} = V\{\sum_j Y_{ij}\} = \sum_j V\{Y_{ij}\} + 2\sum\sum_{j<j'} Cov\{Y_{ij}, Y_{ij'}\} = b(\sigma_\alpha^2 + \sigma_\beta^2 + \sigma^2) + b(b-1)\sigma_\alpha^2$

8 p.1.b. Use the results from above to obtain $E\{MS_{\text{ERR}}\}$

$= b^2\sigma_\alpha^2 + b\sigma_\beta^2 + b\sigma^2 \qquad V\{\overline{Y}_{i\cdot}\} = \sigma_\alpha^2 + \frac{\sigma_\beta^2}{b} + \frac{\sigma^2}{b}$

By direct analogy: $V\{\overline{Y}_{\cdot j}\} = \frac{\sigma_\alpha^2}{a} + \sigma_\beta^2 + \frac{\sigma^2}{a}$

$V\{Y_{\cdot\cdot}\} = \sum_i \sum_j V\{Y_{ij}\} + 2\sum_i \sum\sum_{j<j'} Cov\{Y_{ij}, Y_{ij'}\} + 2\sum\sum_{i<i'} \sum_j Cov\{Y_{ij}, Y_{i'j}\} + 4\sum\sum_{i<i'}\sum\sum_{j<j'} Cov\{Y_{ij}, Y_{i'j'}\}$

$= ab(\sigma_\alpha^2 + \sigma_\beta^2 + \sigma^2) + ab(b-1)\sigma_\alpha^2 + ba(a-1)\sigma_\beta^2 + 0$

$= ab^2\sigma_\alpha^2 + a^2b\sigma_\beta^2 + ab\sigma^2 \Rightarrow V\{\overline{Y}_{\cdot\cdot}\} = \frac{\sigma_\alpha^2}{a} + \frac{\sigma_\beta^2}{b} + \frac{\sigma^2}{ab}$

44

$\rightarrow$

8

p.1.c. Write out MS_A in terms of (a subset) of the components in SS_{ERR} and obtain $E\{MS_A\}$.

(a) $E\{\sum_i \sum_j Y_{ij}^2\} = ab\mu^2 + ab\sigma_\alpha^2 + ab\sigma_\beta^2 + ab\sigma^2$

(b) $E\{b\sum_i \bar{Y}_{i.}^2\} = ab\mu^2 + ab\sigma_\alpha^2 + a\sigma_\beta^2 + a\sigma^2$

(c) $E\{a\sum_j \bar{Y}_{.j}^2\} = ab\mu^2 + b\sigma_\alpha^2 + ab\sigma_\beta^2 + b\sigma^2$

(d) $E\{ab\bar{Y}_{..}^2\} = ab\mu^2 + b\sigma_\alpha^2 + a\sigma_\beta^2 + \sigma^2$

$\Rightarrow E\{SS_{Error}\} = (a)-(b)-(c)+(d) = (ab-a-b+1)\sigma^2 = (a-1)(b-1)\sigma^2$

$\Rightarrow E\{MS_{ERR}\} = \sigma^2$

8 p.1.d. Write out MS_B in terms of (a subset) of the components in SS_{ERR} and obtain $E\{MS_B\}$.

$E\{SSA\} = (b)-(d) = (ab-b)\sigma_\alpha^2 + (a-1)\sigma^2 \Rightarrow E\{MSA\} = b\sigma_\alpha^2 + \sigma^2$

$E\{SSB\} = (c)-(d) = (ab-a)\sigma_\beta^2 + (b-1)\sigma^2 \Rightarrow E\{MSB\} = a\sigma_\beta^2 + \sigma^2$

12 p.1.e. From the published study, we obtain the following sums of squares: $SS_A = 180.32$ $SS_B = 114.90$ $SS_{ERR} = 395.12$
Obtain unbiased estimates of the model's three variances.

$MSA = \frac{180.32}{9-1} = 22.54$ $MSB = \frac{114.90}{10-1} = 12.77$

$MSE = \frac{395.12}{8(9)} = 5.49$

$\hat{\sigma}^2 = 5.49$ $\hat{\sigma}_\alpha^2 = \frac{22.54 - 5.49}{10} = 1.705$

$\hat{\sigma}_\beta^2 = \frac{12.77 - 5.49}{9} = 0.81$ $\hat{\sigma}_Y^2 = 5.49 + 1.71 + 0.81 = 8.01$

p.1.f. Out of the (estimated) total variance $V\{Y_{ijk}\} = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma^2$, what percentages are accounted by the 3 sources:

6

Judges (A) $\frac{1.71}{8.01} = .213$ Brands (B) $\frac{0.81}{8.01} = .101$ Error (AB) $\frac{5.49}{8.01} = .685$

34

Q.2. A study compared the effects of $a = 3$ brands of squash rackets (Factor A: med, high, and higher price) and $b = 2$ levels of strings (Factor B: factory or new) the speed of a ball's bounce off the racket (Y, in m/sec) after having been dropped from a standard height. There were $n = 10$ replicates per treatment (combination of racket brand and string type). Assume these are the only levels of interest to the researchers (that is, both are fixed factors).

p.2.a. Write out the statistical model assuming errors are iid Normal.

10

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad i=1,2,3 \quad j=1,2 \quad k=1,\dots,10$$

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = 0 \quad \varepsilon_{ijk} \sim NID(0, \sigma^2)$$

p.2.b. The following table gives the sample means. Compute the following least squares estimates of model parameters and σ^2:

Means			
Racket\String	Factory(j=1)	New(j=2)	Overall
Medium Price(i=1)	0.363	0.377	0.370
High Price(i=2)	0.350	0.322	0.336
Higher Price(i=3)	0.250	0.360	0.305
Overall	0.321	0.353	0.337

$$\sum_{i=1}^{3}\sum_{j=1}^{2} s_{ij}^2 = 0.0025$$

2 4

$\hat{\mu} =$ 0.337 $\hat{\sigma}^2 = \frac{9(0.0025)}{3(2)(10-1)} = .00041667$

2 each

$\hat{\alpha}_1 =$ .033 $\hat{\alpha}_2 =$ −.001 $\hat{\alpha}_3 =$ −.032 $\hat{\beta}_1 =$ −.016 $\hat{\beta}_2 =$ .016

.363 − .370 − .321 + .337 .350 − .336 − .321 + .337

3 each

$\widehat{(\alpha\beta)}_{11} =$ 0.009 $\widehat{(\alpha\beta)}_{21} =$ .030 $\widehat{(\alpha\beta)}_{31} =$ −.039

$\widehat{(\alpha\beta)}_{12} =$ −0.009 $\widehat{(\alpha\beta)}_{22} =$ −.030 $\widehat{(\alpha\beta)}_{32} =$ .039

5Y

p.2.c. Complete the following Analysis of Variance Table, including stating the null hypothesis for each row in the table.

$$SSA = 2(10)\left[.033^2 + .001^2 + .032^2\right] = .04228$$

$$SSB = 3(10)\left[.016^2 + .016^2\right] = .01536$$

$$SSAB = 10\left[2(.009)^2 + 2(.030)^2 + 2(.039)^2\right] = .05004$$

$$SSE = 9(.0025) = .0225$$

ANOVA	3 each	4 each	2 each	2 each	3 each	1 each	1 each
Source	df	SS	MS	F_obs	F(.05)	H_0:	Reject H0?
Racket	2	.04228	.02114	50.73	~3.17	$\alpha_1 = \alpha_2 = \alpha_3 = 0$	Yes / No
String	1	.01536	.01536	36.86	~4.02	$\beta_1 = \beta_2 = 0$	Yes / No
R*S	2	.05004	.02502	60.04	~3.17	$(\alpha\beta)_{ij} = 0 \ \forall i, j$	Yes / No
Error	54	.0225	.0004167	#N/A	#N/A	#N/A	#N/A
Total	59	.13018	—	#N/A	#N/A	#N/A	#N/A

57

Q.3. A study was conducted to test for effects on willingness to pay during online auctions. There were 2 factors, each with 2 levels (both fixed): Urgency (Present (i=1) /Absent (i=2)) and Contrast (3 of 6 items "Featured" (High, j=1)/all 6 items "Featured" (Low, j=2)). There were 6 watches, and the response was the amount the participant was willing to pay for the watch. Although the researchers started with 80 subjects, 9 were eliminated due to incomplete information, so N = 71. The model fit is:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad i=1,2;\ j=1,2;\ k=1,...,n_{ij} \quad \alpha_1+\alpha_2 = \beta_1+\beta_2 = \sum_{i=1}^{2}(\alpha\beta)_{ij} = \sum_{j=1}^{2}(\alpha\beta)_{ij} = 0 \quad i,j=1,2$$

p.3.a. Give the form of the (full rank) **X** matrix and **β** vector for this model. Note that although **X** has 71 rows, there are only 4 "blocks" of distinct levels, each with a particular numbers of subjects.

4 8

$$\underset{\sim}{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \beta_1 \\ (\alpha\beta)_{11} \end{bmatrix} \qquad X = \begin{bmatrix} \underset{\sim}{1}_{n_{11}} & \underset{\sim}{1}_{n_{11}} & \underset{\sim}{1}_{n_{11}} & \underset{\sim}{1}_{n_{11}} \\ \underset{\sim}{1}_{n_{12}} & \underset{\sim}{1}_{n_{12}} & -\underset{\sim}{1}_{n_{12}} & -\underset{\sim}{1}_{n_{12}} \\ \underset{\sim}{1}_{n_{21}} & -\underset{\sim}{1}_{n_{21}} & \underset{\sim}{1}_{n_{21}} & -\underset{\sim}{1}_{n_{21}} \\ \underset{\sim}{1}_{n_{22}} & -\underset{\sim}{1}_{n_{22}} & -\underset{\sim}{1}_{n_{22}} & \underset{\sim}{1}_{n_{22}} \end{bmatrix}$$

p.3.b. Obtain **X'X** as functions of the cell sample sizes (Just give the values on or above the main diagonal). $N = n_{11}+n_{12}+n_{21}+n_{22}$

10

$$X'X = \begin{bmatrix} N & n_{11}+n_{12}-n_{21}-n_{22} & n_{11}-n_{12}+n_{21}-n_{22} & n_{11}-n_{12}-n_{21}+n_{22} \\ & N & n_{11}-n_{12}-n_{21}+n_{22} & n_{11}-n_{12}+n_{21}-n_{22} \\ & & N & n_{11}+n_{12}-n_{21}-n_{22} \\ & & & N \end{bmatrix}$$

p.3.c. The authors fit the following 4 models, with approximate Error sums of squares (divided by 1000 for ease of calculation):

Model 1: $E\{Y_{ijk}\} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ Model 2: $E\{Y_{ijk}\} = \mu + \alpha_i + \beta_j$ Model 3: $E\{Y_{ijk}\} = \mu + \beta_j + (\alpha\beta)_{ij}$ Model 4: $E\{Y_{ijk}\} = \mu + \alpha_i + (\alpha\beta)_{ij}$

$SSErr_1 = 884.0$ $SSErr_2 = 937.0$ $SSErr_3 = 941.5$ $SSErr_4 = 978.3$

Test: $H_0^{AB}: (\alpha\beta)_{11} = (\alpha\beta)_{12} = (\alpha\beta)_{21} = (\alpha\beta)_{22} = 0$

$$T.S.\ F_{obs} = \frac{\dfrac{SSE_2 - SSE_1}{(71-3)-(71-4)}}{\dfrac{SSE_1}{71-4}} = \frac{\dfrac{937-884}{1}}{\dfrac{884}{67}} = \frac{67(53)}{884} = 4.017$$

4 2 4 2

$$F_{.05,1,67} \simeq 3.985$$

12 3 2

Test Statistic: $F = 4.017$ Rejection Region: $F > 3.985$ Do you conclude the Urgency effect "depends" on Contrast? Y N

40

p.3.d. The sample means for the 4 treatments are: $\overline{Y}_{11\bullet} = 216.94$ $\overline{Y}_{12\bullet} = 90.26$ $\overline{Y}_{21\bullet} = 106.12$ $\overline{Y}_{22\bullet} = 88.06$. Treating the cell sample sizes as $n_{ij} = 18$ as an approximation, use Bonferroni's method to compare all pairs of treatment means.

$$MSE = \frac{884000}{67} = 13194$$

$$t(.025, 6, 67) \approx 2.72 \quad (4) \qquad \sqrt{\frac{2(13194)}{18}} = 38.29 \quad (4)$$

$$B = 104.1 \quad (2)$$

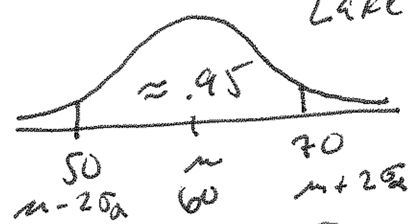
Conclude $\mu_{11} > \mu_{12}, \mu_{21}, \mu_{22}$ (2)

None others are significantly different

Q.4. Among a population of lakes, the mean adult fish lengths are normally distributed, with approximately 95% of the lake means lying between 50 and 70 centimeters. Within lakes, approximately 95 % of the fish have lengths within 12 cm of the lake mean. Consider a 1-Way random effects model, where a sample of g lakes is selected and n fish are sampled from each lake.

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \quad i = 1,\ldots,g;\ j = 1,\ldots,n \quad \alpha_i \sim NID\left(0, \sigma_\alpha^2\right) \quad \varepsilon_{ij} \sim NID\left(0, \sigma^2\right) \quad \{\alpha_i\} \perp \{\varepsilon_{ij}\}$$

Lake Means



$\Rightarrow \sigma_\alpha \approx 5 \quad \mu = 60$

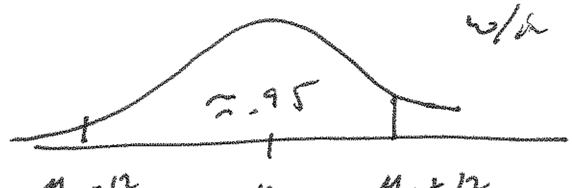
p.4.a. Obtain μ, σ_α^2, σ^2

4 $\mu = 60$

4 $\sigma_\alpha^2 = 5^2 = 25$

4 $\sigma^2 = 6^2 = 36$

w/in Lake



$\sigma \approx 6$

p.4.b. Give the mean and variance of $\overline{Y}_{\bullet\bullet}$

$$V\{Y_{..}\} = ng(\sigma_\alpha^2 + \sigma^2) + gn(n-1)\sigma_\alpha^2$$

$$= gn^2\sigma_\alpha^2 + ng\sigma^2$$

$$\Rightarrow V\{\overline{Y}_{..}\} = \left(\frac{1}{ng}\right)^2 \left[gn^2\sigma_\alpha^2 + ng\sigma^2\right] = \frac{\sigma_\alpha^2}{g} + \frac{\sigma^2}{ng} = \frac{25}{g} + \frac{36}{ng}$$

32

Q.5. Based on the 2014 WNBA season, we have the point totals (Y) by game **Location** (Home/Away) for a **sample** of 10 **Players**. Each player played 17 home games and 17 away games. Consider the model:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk} \quad i=1,\ldots,a \quad j=1,\ldots,b \quad k=1,\ldots,n \quad \sum_{i=1}^{a}\alpha_i = 0 \quad \beta_j \sim N\left(0,\sigma_\beta^2\right) \quad \alpha\beta_{ij} \sim N\left(0,\sigma_{\alpha\beta}^2\right) \quad \varepsilon_{ijk} \sim N\left(0,\sigma^2\right)$$

ANOVA	4 each		2 each	4 each	3 each
Source	df	SS	MS	F	F(0.05)
Player	10-1=9	3879.30	431.03	11.99	3.179
Location	2-1=1	1.30	1.30	0.036	5.117
P*L	9(1)=9	323.67	35.96	0.729	≈1.9
Error	2(10)(17-1)=320	15787.29	49.34	#N/A	#N/A
Total	2(10)(17)-1=339	19991.56	#N/A	#N/A	#N/A

p.5.a. Complete the partial ANOVA table.

p.5.b. Test whether there is an interaction between Player and Location (Home). H_0: $\sigma_{\alpha\beta}^2 = 0$

1

p.5.b.i. Test Stat: 0.729 p.5.b.ii. Reject H_0 if Test Stat is in the range >1.9 p.5.b.iii. P-value > or < .05?

p.5.c. Test whether there is Location (Home vs Away) Main Effect. H_0: $\alpha_1 = \alpha_2 = 0$

1

p.5.c.i. Test Stat: 0.036 p.5.c.ii. Reject H_0 if Test Stat is in the range >5.117 p.5.c.iii. P-value > or < .05?

p.5.d. Test whether there is Player Main Effect. H_0: $\sigma_\beta^2 = 0$

1

p.5.d.i. Test Stat: 11.99 p.5.d.ii. Reject H_0 if Test Stat is in the range >3.179 p.5.d.iii. P-value > or < .05?

p.5.e. Give unbiased estimates of each of the variance components:

$E\{MSAB\} = \sigma^2 + n\sigma_{\alpha\beta}^2$ $\qquad$ $E\{MSB\} = \sigma^2 + n\sigma_{\alpha\beta}^2 + an\sigma_\beta^2$

$\dfrac{35.96-49.34}{17} = -0.79$ $\qquad$ $\dfrac{431.03-35.96}{2(17)} = 11.62$

$\hat{\sigma}_{\alpha\beta}^2 =$ -0.79 (0) $\qquad$ $\hat{\sigma}_\beta^2 =$ 11.62 $\qquad$ $\hat{\sigma}^2 =$ 49.34

4 $\qquad$ 4 $\qquad$ 2

56