ANSWER KEY

STA 6208 – Exam 1 – Spring 2015 PRINT NAME ANSWER KEY

For Each Problem, assume the Completely Randomized Design (1-Way ANOVA): 289

$$Y_{ij} = \mu_i + e_{ij} = \mu + \tau_i + e_{ij} \quad i = 1,\dots,t;\ j = 1,\dots,r_i \quad e_{ij} \sim NID\left(0,\sigma^2\right) \quad \sum_{i=1}^{t} r_i \tau_i = 0$$

F(A=.95)	df1							
df2	1	2	3	4	5	6	7	8
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097
70	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056

Studentized Range							
df\#trts	2	3	4	5	6	7	8
1	17.969	26.976	32.819	37.082	40.408	43.119	45.397
2	6.085	8.331	9.798	10.881	11.734	12.435	13.027
3	4.501	5.910	6.825	7.502	8.037	8.478	8.852
4	3.926	5.040	5.757	6.287	6.706	7.053	7.347
5	3.635	4.602	5.218	5.673	6.033	6.330	6.582
6	3.460	4.339	4.896	5.305	5.628	5.895	6.122
7	3.344	4.165	4.681	5.060	5.359	5.606	5.815
8	3.261	4.041	4.529	4.886	5.167	5.399	5.596
9	3.199	3.948	4.415	4.755	5.024	5.244	5.432
10	3.151	3.877	4.327	4.654	4.912	5.124	5.304
11	3.113	3.820	4.256	4.574	4.823	5.028	5.202
12	3.081	3.773	4.199	4.508	4.750	4.950	5.119
13	3.055	3.734	4.151	4.453	4.690	4.884	5.049
14	3.033	3.701	4.111	4.407	4.639	4.829	4.990
15	3.014	3.673	4.076	4.367	4.595	4.782	4.940
16	2.998	3.649	4.046	4.333	4.557	4.741	4.896
17	2.984	3.628	4.020	4.303	4.524	4.705	4.858
18	2.971	3.609	3.997	4.276	4.494	4.673	4.824
19	2.960	3.593	3.977	4.253	4.468	4.645	4.794
20	2.950	3.578	3.958	4.232	4.445	4.620	4.768
21	2.941	3.565	3.942	4.213	4.424	4.597	4.743
22	2.933	3.553	3.927	4.196	4.405	4.577	4.722
23	2.926	3.542	3.914	4.180	4.388	4.558	4.702
24	2.919	3.532	3.901	4.166	4.373	4.541	4.684
25	2.913	3.523	3.890	4.153	4.358	4.526	4.667
26	2.907	3.514	3.880	4.141	4.345	4.511	4.652
27	2.902	3.506	3.870	4.130	4.333	4.498	4.638
28	2.897	3.499	3.861	4.120	4.322	4.486	4.625
29	2.892	3.493	3.853	4.111	4.311	4.475	4.613
30	2.888	3.486	3.845	4.102	4.301	4.464	4.601
40	2.858	3.442	3.791	4.039	4.232	4.388	4.521
48	2.843	3.420	3.764	4.008	4.197	4.351	4.481
60	2.829	3.399	3.737	3.977	4.163	4.314	4.441
80	2.814	3.377	3.711	3.947	4.129	4.277	4.402

Bonferroni t				
df\#Comparis	1	3	6	10
1	12.706	38.188	76.390	127.321
2	4.303	7.649	10.886	14.089
3	3.182	4.857	6.232	7.453
4	2.776	3.961	4.851	5.598
5	2.571	3.534	4.219	4.773
6	2.447	3.287	3.863	4.317
7	2.365	3.128	3.636	4.029
8	2.306	3.016	3.479	3.833
9	2.262	2.933	3.364	3.690
10	2.228	2.870	3.277	3.581
11	2.201	2.820	3.208	3.497
12	2.179	2.779	3.153	3.428
13	2.160	2.746	3.107	3.372
14	2.145	2.718	3.069	3.326
15	2.131	2.694	3.036	3.286
16	2.120	2.673	3.008	3.252
17	2.110	2.655	2.984	3.222
18	2.101	2.639	2.963	3.197
19	2.093	2.625	2.944	3.174
20	2.086	2.613	2.927	3.153
21	2.080	2.601	2.912	3.135
22	2.074	2.591	2.899	3.119
23	2.069	2.582	2.886	3.104
24	2.064	2.574	2.875	3.091
25	2.060	2.566	2.865	3.078
26	2.056	2.559	2.856	3.067
27	2.052	2.552	2.847	3.057
28	2.048	2.546	2.839	3.047
29	2.045	2.541	2.832	3.038
30	2.042	2.536	2.825	3.030
40	2.021	2.499	2.776	2.971
50	2.009	2.477	2.747	2.937
60	2.000	2.463	2.729	2.915
70	1.994	2.453	2.715	2.899
80	1.990	2.445	2.705	2.887

Q.1 Derive E{MS_{Trts}} for the balanced 1-Way ANOVA with t treatments and r replicates per treatment by completing the following parts. **SHOW ALL WORK to Receive any Credit!!!!!!!!!**

p.1.a. Show: $SS_{Trts} = \sum_{i=1}^{t}\sum_{j=1}^{r}\left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right)^2 = r\sum_{i=1}^{t}\left(\bar{y}_{i\bullet}\right)^2 - rt\left(\bar{y}_{\bullet\bullet}\right)^2$ 12

$$= \sum_i\sum_j \left(\bar{y}_{i.}^2 + \bar{y}_{..}^2 - 2\bar{y}_{i.}\bar{y}_{..}\right) = \sum_i\sum_j \bar{y}_{i.}^2 + \sum_i\sum_j \bar{y}_{..}^2 - 2\sum_i\sum_j \bar{y}_{i.}\bar{y}_{..}$$

$$= r\sum_i \bar{y}_{i.}^2 + tr\bar{y}_{..}^2 - 2r\bar{y}_{..}\sum_i \bar{y}_{i.} \qquad \frac{\sum_i \bar{y}_{i.}}{t} = \bar{y}_{..} \Rightarrow \sum_i \bar{y}_{i.} = t\bar{y}_{..}$$

$$= r\sum_i \bar{y}_{i.}^2 + tr\bar{y}_{..}^2 - 2tr\bar{y}_{..}^2$$

$$= r\sum_i \bar{y}_{i.}^2 - tr\bar{y}_{..}^2$$

4 4 4 4 4 4

p.1.b. Derive $E\{\bar{y}_{i\bullet}\}$, $V\{\bar{y}_{i\bullet}\}$, $E\{(\bar{y}_{i\bullet})^2\}$, $E\{\bar{y}_{\bullet\bullet}\}$, $V\{\bar{y}_{\bullet\bullet}\}$, $E\{(\bar{y}_{\bullet\bullet})^2\}$

$E\{y_{ij}\} = \mu_i \quad V\{y_{ij}\} = \sigma^2 \quad Cov\{y_{ij}, y_{i'j'}\} = 0 \ \forall\ i \neq i'$ and/or $j \neq j'$

$$E\left\{\frac{1}{r}\sum_j y_{ij}\right\} = \frac{1}{r}\sum_j E\{y_{ij}\} = \frac{1}{r}r\mu_i = \mu_i \qquad V\left\{\frac{1}{r}\sum_j y_{ij}\right\} = \frac{1}{r^2}r\sigma^2 = \frac{\sigma^2}{r}$$

$$\Rightarrow E\{(\bar{y}_{i.})^2\} = \mu_i^2 + \frac{\sigma^2}{r} = \left(\bar{\mu}_.^2 + \tau_i^2 + 2\bar{\mu}_.\tau_i\right) + \frac{\sigma^2}{r}$$

$$E\left\{\frac{1}{tr}\sum_i\sum_j y_{ij}\right\} = \frac{1}{tr}\sum_i\sum_j E\{y_{ij}\} = \frac{1}{tr}\sum_i r\mu_i = \frac{1}{t}\sum_i \mu_i = \bar{\mu}_.$$

$$V\left\{\frac{1}{tr}\sum_i\sum_j y_{ij}\right\} = \left(\frac{1}{tr}\right)^2 tr\sigma^2 = \frac{\sigma^2}{tr}$$

$$\Rightarrow E\{\bar{y}_{..}^2\} = \bar{\mu}_.^2 + \frac{\sigma^2}{tr}$$

36

p.1.c. Use results from p.1.a. and p.1.b. to obtain: $E\{SS_{\text{TRTS}}\}$ and $E\{MS_{\text{TRTS}}\}$ 12

$$E\{SS_{TRTS}\} = r\sum_i\left[\left(\bar{\mu}_.^2 + \tau_i^2 + 2\bar{\mu}_.\tau_i\right) + \frac{\sigma^2}{r}\right] - tr\left[\bar{\mu}_.^2 + \frac{\sigma^2}{tr}\right]$$

$$= \left(\cancel{tr\bar{\mu}_.^2} + r\sum_i\tau_i^2 + 2\bar{\mu}_. r\overbrace{\sum_i\tau_i}^{0} + t\sigma^2\right) - \left(\cancel{tr\bar{\mu}_.^2} + \sigma^2\right)$$

$$= r\sum_i\tau_i^2 + (t-1)\sigma^2$$

$$\Rightarrow E\{MS_{TRTS}\} = E\left\{\frac{SS_{TRTS}}{t-1}\right\} = \frac{r\sum_i\tau_i^2}{t-1} + \sigma^2$$

Q.1.d. For which scenario of the parameter values, will the power of the F-test be highest (assume $t = 3$, $r_1 = r_2 = r_3 = r$):

Scenario 1: $\mu_1 = 80$, $\mu_2 = 100$, $\mu_3 = 120$, $\sigma = 10$ or **Scenario 2:** $\mu_1 = 90$, $\mu_2 = 100$, $\mu_3 = 110$, $\sigma = 5$

Compute $E\{MS_{Trts}\}$ for each scenario (as a function of r). 12

(1) $\tau_1 = -20 \quad \tau_2 = 0 \quad \tau_3 = 20 \Rightarrow \sum_i \tau_i^2 = 800$

$$\Rightarrow E\{MS_{TRTS}\} = \frac{800r}{2} + 100 = 100 + 400r$$

(2) $\tau_1 = -10 \quad \tau_2 = 0 \quad \tau_3 = 10 \Rightarrow \sum \tau_i^2 = 200$

$$E\{MS_{TR}\} = \frac{200r}{2} + 25 = 100r + 25$$

24

Q.2. A published report, based on a **balanced 1-Way ANOVA** reports means (SDs) for the three treatments as:

Trt 1: 70 (8) Trt 2: 75 (6) Trt 3: 80 (10) Tragically, the authors fail to give the treatment sample sizes.

p.2.a The Treatment degrees of freedom is the same for each sample size. The error degrees of freedom are:

$$\sum_i (\bar{y}_{i.} - \bar{y}_{..})^2 = 25 + 0 + 25 = 50$$

$$\sum_i s_i^2 = 64 + 36 + 100 = 200$$

3 3 3 3

df_{TRTS} = 3-1=2 $r_i = 2$: df_{ERR} = 3(2-1)=3 $r_i = 6$: df_{ERR} = 3(6-1)=15 $r_i = 10$: df_{ERR} = 3(10-1)=27

p.2.b. Complete the following table, given arbitrary levels of the number of replicates per treatment:

	4 each	4 each	2 each	2 each	2 each	3 each
r	SSTrt	SSErr	MSTrt	MSErr	F_obs	F(0.05)
2	2(50) = 100	1(200) = 200	50	66.67	0.75	9.552
6	6(50) = 300	5(200) = 1000	150	66.67	2.25	3.682
10	10(50) = 500	9 (200) = 1800	250	66.67	3.75	3.354

p.2.c. The smallest r_i, so that these means are significantly different is: 3

i) $r_i <= 2$ ii) $2 < r_i <= 6$ iii) $6 < r_i <= 10$ iv) $r_i > 10$

66

Q.3. A study is conducted to compare 4 menus in terms of numbers of calories ordered by restaurant customers (in 100s of calories). The treatments (menus) are (consider them increasing in order of information provided):

1) No Calories Reported 2) Calories Reported 3) Rank-Ordered Calories 4) Color-Coded Calories

The sample sizes are all based on samples of r = 20 customers per menu. The sample means and estimated variance are:
$\overline{Y}_{1\bullet} = 17.6$ $\overline{Y}_{2\bullet} = 16.8$ $\overline{Y}_{3\bullet} = 16.0$ $\overline{Y}_{4\bullet} = 14.4$ $\overline{Y}_{\bullet\bullet} = 16.2$ $s^2 = \text{MSErr} = 196.0$.

p.3.a. Complete the following table to test H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

2 each SSE = 14896 6 each 2 each 4 3 2

ANOVA	df	SS	MS	F_obs	F(.05)	Reject H0?
Source						Yes / No
Trts (Menus)	3	~~98.40~~ 112	~~32.80~~ 37.33	#N/A	#N/A	#N/A
Error	76	~~1132076~~	~~14896~~ 196	#N/A	#N/A	#N/A
Total	79		#N/A	#N/A	#N/A	#N/A

~~$F^* = 0.167$~~

$F^* = 0.190$

$F(.05, 3, 76) \approx 2.725$

$$SS_{TRTS} = 20\left[1.4^2 + 0.6^2 + (-0.2)^2 + (-1.8)^2\right]$$

$$= 20\left[1.96 + 0.36 + 0.04 + 3.24\right] = \cancel{98.40}\ 112$$

$$SSE = 76(196) = \cancel{1132096}$$

p.3.b.: Give 3 orthogonal contrasts, and their estimates and estimated standard errors

Contrast 1: Menu 1 vs Menus {2, 3, 4} Contrast 2: Menu 2 vs Menus {3, 4} Contrast 3: Menu 3 vs Menu 4

3 3 3

Contrast 1: $C_1 = 3\mu_1 - \mu_2 - \mu_3 - \mu_4$ $c_1 = 52.8 - 16.8 - 16.0 - 14.4 = 5.6$ $SE[c_1] = \sqrt{196\left[\frac{9+1+1+1}{20}\right]} = \sqrt{117.6} = 10.84$

3 3 3

Contrast 2: $C_2 = 2\mu_2 - \mu_3 - \mu_4$ $c_2 = 33.6 - 16 - 14.4 = 3.2$ $SE[c_2] = \sqrt{196\left(\frac{4+1+1}{20}\right)} = \sqrt{58.8} = 7.67$

3 3 3

Contrast 3: $C_3 = \mu_3 - \mu_4$ $c_3 = 16.0 - 14.4 = 1.6$ $SE[c_3] = \sqrt{196\left(\frac{1+1}{20}\right)} = \sqrt{19.6} = 4.43$

p.3.c. Obtain Simultaneous 95% CIs for each (population) contrast, based on Bonferroni's method

$$t\left(\frac{.025}{3}, 76\right) \approx 2.448$$

4 Contrast 1: $5.6 \pm 2.448(10.84) \equiv 5.6 \pm 26.5 = \cancel{(-12.1, 40.9)}\ (-20.9, 32.1)$

4 Contrast 2: $3.2 \pm 2.448(7.67) \equiv 3.2 \pm 18.8 = (-15.6, 22.0)$

4 Contrast 3: $1.6 \pm 2.448(4.43) \equiv 1.6 \pm 10.84 = (-9.2, 12.4)$

66

Q.4. Players in the English Premier Soccer League are classified as one of 4 positions (Defender, Forward, Goalie, or Midfielder). Random samples of 8 players were selected from each position for players during the 2014 season. The players' Body Mass Indices were measured. The within position (Error) sum of squares is SSE = 68. The means for the 4 positions are: 22.77 (D), 22.19 (F), 24.11 (G), and 23.16 (M).

$df_E = 32-4=28$

p.4.a. Compute the standard error of the difference between 2 means: $s\{\overline{Y}_{i\bullet}-\overline{Y}_{i'\bullet}\}$

$MSE = \frac{68}{28} = 2.43$

$\sqrt{2.43\left(\frac{1}{8}+\frac{1}{8}\right)} = \sqrt{.6071} = 0.78$

8

$s\{\overline{Y}_{i\bullet}-\overline{Y}_{i'\bullet}\}=$ 0.78

Max diff = 24.11 - 22.19 = 1.92

p.4.b. Compute Tukey's Honest Significant Difference for simultaneously comparing all pairs of positions, with a family-wise error rate of 0.05. Identify significant differences (if any) among all pairs of means.

$q(.05, 4, 28) = 3.861$

$HSD = \frac{3.861}{\sqrt{2}}(0.78) = 2.13$

8

Tukey's HSD = 2.13

None significantly different

p.4.c. Compute Bonferroni's Minimum Significant Difference for simultaneously comparing all pairs of positions, with a family-wise error rate of 0.05. Identify significant differences (if any) among all pairs of means.

$t\left(\frac{.025}{6}, 28\right) = 2.839$ $\quad B = 2.839(0.78) = 2.21$

8

Bonferroni's MSD = 2.21

None different

p.4.d. Compute Scheffe's Minimum Significant Difference for simultaneously comparing all pairs of positions, with a family-wise error rate of 0.05. Identify significant differences (if any) among all pairs of means.

$\sqrt{(4-1)F(.05,3,28)} = \sqrt{3(2.947)} = 2.97$

$2.97(.78) = 2.32$

8

Scheffe's MSD = 2.32

None Different

32

Q.5. Consider the following 2 models:

Model 1: $y_{ij} = \mu_i + e_{ij} \quad i = 1,\ldots,t;\ j = 1,\ldots,r_i$ Model 2: $y_{ij} = \mu + e_{ij} \quad i = 1,\ldots,t;\ j = 1,\ldots,r_i$

p.5.a. Derive the least squares estimator of μ_k for Model 1. SHOW ALL WORK!!!!!

12✓

$$Q = \sum_i \sum_j (y_{ij} - \mu_i)^2$$

$$\frac{\partial Q}{\partial \mu_k} = -2 \sum_j (y_{kj} - \mu_k) \overset{set}{=} 0$$

$$\Rightarrow \sum_j y_{kj} = r_k \hat{\mu}_k \Rightarrow \hat{\mu}_k = \bar{y}_{k.}$$

p.5.b. Derive the least squares estimator of μ for Model 2. SHOW ALL WORK!!!!!

12✓

$$Q = \sum_i \sum_j (y_{ij} - \mu)^2$$

$$\frac{\partial Q}{\partial \mu} = -2 \sum_i \sum_j (y_{ij} - \mu) \overset{set}{=} 0$$

$$\Rightarrow \sum_i \sum_j y_{ij} = \hat{\mu} \sum_i r_i \Rightarrow \hat{\mu} = \frac{\sum_i \sum_j y_{ij}}{\sum_i r_i} = \bar{y}_{..}$$

24

p.5.c. Three brands of cellphones are to be compared in terms of sound quality. Samples of $r_1 = 2$, $r_2 = 3$, $r_3 = 2$ are obtained, with the following sound quality scores: Brand 1: 8, 9; Brand 2: 8, 7, 6; Brand 3: 5, 6.

Brand1	Brand2	Brand3
8	8	5
9	7	6
	6	

Mod 1 8.5 7 5.5 $\hat{\mu}$
Mod 2 7 7 7

Obtain the Error Sum of Squares and Degrees of Freedom for each model:

10

Model 2: $SSE_2 = (8-7)^2 + (9-7)^2 + (8-7)^2 + (7-7)^2 + (6-7)^2 + (5-7)^2 + (6-7)^2$ $df_{E2} =$ $7-1=6$ 3

$= 1+4+1+0+1+4+1 = 12$

10

Model 1: $SSE_1 =$ $df_{E1} =$ $7-3=4$ 3

$= (-0.5)^2 + (0.5)^2 + (1)^2 + 0^2 + (-1)^2 + (-0.5)^2 + (0.5)^2$

$= 0.25 + 0.25 + 1 + 0 + 1 + 0.25 + 0.25 = 3$

p.5.d. Conduct the General Linear Test to test $H_0: \mu_1 = \mu_2 = \mu_3$

12

$$F_{obs} = \frac{\frac{SSE_2 - SSE_1}{df_{E2} - df_{E1}}}{\frac{SSE_1}{df_{E1}}} = \frac{\frac{12-3}{6-4}}{\frac{3}{4}} = \frac{9}{2}\left(\frac{4}{3}\right) = 6$$

$F(.05; 2,4) = 6.944$

3

Test Statistic: 6 Rejection Region: $F_{obs} \geq 6.944$

41