$\boxed{\text{TOTAL: 209}}$

STA 6207 – Exam 1 – Fall 2016 PRINT Name ANSWER KEY

Unless stated otherwise, Questions are based on the following 2 regression models, and X is full rank.

Model 1: $Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i \quad i = 1,...,n \quad \varepsilon_i \sim NID\left(0,\sigma^2\right)$

Model 2: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \mathbf{X} \equiv n \times 2 \quad \boldsymbol{\beta} \equiv 2 \times 1 \quad \boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$

Given:

$$\frac{d(\mathbf{a'x})}{d\mathbf{x}} = \mathbf{a} \qquad \frac{d(\mathbf{x'Ax})}{d\mathbf{x}} = 2\mathbf{Ax} \text{ (A symmetric)} \qquad SS_{XX} = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad SS_{XY} = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)$$

Conduct all Tests at $\alpha = 0.05$ significance level.

---

t and F Critical Values Indexed by Numerator Degrees of Freedom

df	$t_{.05}$	$t_{.025}$	$\chi^2_{.05}$	$F_{.05,1}$	$F_{.05,2}$	$F_{.05,3}$	$F_{.05,4}$	$F_{.05,5}$	$F_{.05,6}$	$F_{.05,7}$	$F_{.05,8}$
1	6.314	12.706	3.841	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883
2	2.920	4.303	5.991	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371
3	2.353	3.182	7.815	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845
4	2.132	2.776	9.488	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041
5	2.015	2.571	11.070	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818
6	1.943	2.447	12.592	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147
7	1.895	2.365	14.067	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726
8	1.860	2.306	15.507	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438
9	1.833	2.262	16.919	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230
10	1.812	2.228	18.307	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072
11	1.796	2.201	19.675	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948
12	1.782	2.179	21.026	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849
13	1.771	2.160	22.362	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767
14	1.761	2.145	23.685	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699
15	1.753	2.131	24.996	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641
16	1.746	2.120	26.296	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591
17	1.740	2.110	27.587	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548
18	1.734	2.101	28.869	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510
19	1.729	2.093	30.144	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477
20	1.725	2.086	31.410	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447
21	1.721	2.080	32.671	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420
22	1.717	2.074	33.924	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397
23	1.714	2.069	35.172	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375
24	1.711	2.064	36.415	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355
25	1.708	2.060	37.652	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337
26	1.706	2.056	38.885	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321
27	1.703	2.052	40.113	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305
28	1.701	2.048	41.337	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291
29	1.699	2.045	42.557	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278
30	1.697	2.042	43.773	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266
40	1.684	2.021	55.758	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180
50	1.676	2.009	67.505	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130
60	1.671	2.000	79.082	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097
70	1.667	1.994	90.531	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074
80	1.664	1.990	101.879	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056
90	1.662	1.987	113.145	3.947	3.098	2.706	2.473	2.316	2.201	2.113	2.043
100	1.660	1.984	124.342	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032
110	1.659	1.982	135.480	3.927	3.079	2.687	2.454	2.297	2.182	2.094	2.024
120	1.658	1.980	146.567	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016
130	1.657	1.978	157.610	3.914	3.066	2.674	2.441	2.284	2.169	2.081	2.010
140	1.656	1.977	168.613	3.909	3.061	2.669	2.436	2.279	2.164	2.076	2.005
150	1.655	1.976	179.581	3.904	3.056	2.665	2.432	2.274	2.160	2.071	2.001
160	1.654	1.975	190.516	3.900	3.053	2.661	2.428	2.271	2.156	2.067	1.997
170	1.654	1.974	201.423	3.897	3.049	2.658	2.425	2.267	2.152	2.064	1.993
180	1.653	1.973	212.304	3.894	3.046	2.655	2.422	2.264	2.149	2.061	1.990
190	1.653	1.973	223.160	3.891	3.043	2.652	2.419	2.262	2.147	2.058	1.987
200	1.653	1.972	233.994	3.888	3.041	2.650	2.417	2.259	2.144	2.056	1.985
∞	1.645	1.960	---	3.841	2.995	2.605	2.372	2.214	2.099	2.010	1.938

Q.1. A simple linear regression model is to be fit, relating Mean Annual Temperature (Y, in °F) to Year – 1957 (X). That is, the "origin" is 1957. The data are for the years 1957-2014 (n = 58). The sample means and sums of squares and cross-products are given below for Model 1.

$$\overline{X}=28.5 \quad \overline{Y}=68.2759 \quad \sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)=559.6083 \quad \sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=16254.5 \quad \sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2=83.4361$$

p.1.a. Compute the least squares estimates of β_1 and β_0, and write out the predicted equation.

5 $\hat{\beta}_1 = \dfrac{559.6083}{16254.5} = .0344$

5 $\hat{\beta}_0 = 68.2759 - .0344(28.5) = 68.2759 - 0.9804 = 67.2955$

5 $\hat{Y} = 67.2955 + .0344X$

p.1.b. $\sum_{i=1}^{n}\left(Y_i-\hat{Y}_i\right)^2 = 64.1699$ Use this to obtain an unbiased estimate of σ^2.

5 $s^2 = \dfrac{64.1699}{58-2} = 1.1459$

p.1.c. Obtain a 95% Confidence Interval for β_1 (the amount, on average, that mean temperature increases per ~~year~~). 7

3 $t_{.025,56} \approx 2.004$ $\quad \hat{SE}\{\hat{\beta}_1\} = \sqrt{\dfrac{1.1459}{16254.5}} = 0.0084$

$0.0344 \pm \underbrace{2.004(0.0084)}_{.0168} = (0.0176, 0.0512)$ 5

p.1.d. For the year 2014, the (observed) average temperature was 67.65. Obtain the predicted temperature and the residual for that year.

$X_{2014} = 2014 - 1957 = 57$ 2 $\quad \hat{Y}_{2014} = 67.2955 + 0.0344(57)$

$= 67.2955 + 1.9608 = 69.2563$ 5

$e_{2014} = 67.65 - 69.2563 = -1.6063$ 4

46

Q.2. Based on Model 1, derive the two normal equations by minimizing $Q = \sum_{i=1}^{n} \varepsilon_i^2$ with respect to β_0 and β_1.

Show all Work. $Q = \sum_i (Y_i - \beta_0 - \beta_1 X_i)^2$

$$\frac{\partial Q}{\partial \beta_0} = 2 \sum_i (Y_i - \beta_0 - \beta_1 X_i)(-1) \overset{set}{=} 0$$

12

$$\Rightarrow \sum_i Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum_i X_i$$

$$\frac{\partial Q}{\partial \beta_1} = 2 \sum_i (Y_i - \beta_0 - \beta_1 X_i)(-X_i) \overset{set}{=} 0$$

12

$$\Rightarrow \sum_i X_i Y_i = \hat{\beta}_0 \sum_i X_i + \hat{\beta}_1 \sum_i X_i^2$$

Q.3. The fitted value for Y_j can be written as $\hat{Y}_j = \sum_{i=1}^{n} \left[\frac{1}{n} + \frac{(X_i - \bar{X})}{SS_{XX}} (X_j - \bar{X}) \right] Y_i$

p.3.a. Give $\mathrm{COV}\{Y_j, \hat{Y}_j\}$

$$\mathrm{Cov}\{Y_j, \hat{Y}_j\} = \mathrm{Cov}\left\{Y_j, \left[\frac{1}{n} + \frac{(X_j - \bar{X})^2}{SS_{XX}}\right] Y_j\right\}$$

$$= \sigma^2 \left[\frac{1}{n} + \frac{(X_j - \bar{X})^2}{SS_{XX}} \right]$$

10

p.3.b. Give $\mathrm{COV}\{Y_j, \hat{Y}_k\}$ $j \neq k$

$$\mathrm{Cov}\{Y_j, \hat{Y}_k\} = \mathrm{Cov}\left\{Y_j, \left[\frac{1}{n} + \frac{(X_j - \bar{X})}{SS_{XX}}(X_k - \bar{X})\right] Y_j\right\}$$

$$= \sigma^2 \left[\frac{1}{n} + \frac{(X_j - \bar{X})(X_k - \bar{X})}{SS_{XX}} \right]$$

10

44

Q.4. A simple linear regression model is fit relating fairway accuracy (Y, in percent) to average drive distance (X, in yards) for a random sample of n = 20 professional women golfers from the 2009 season, based on Model 1.

p.4.a. Complete the following table cells.

Regression Statistics					
R Square	,2005 4				
Observations	20				
ANOVA					
Source	*df*	*SS*	*MS*	*F*	*F(.05)*
Regression	1	81.7	81.7	4.514 4	4.414 4
Residual	2 20-2 = 18	4 407.5-81.7 = 325.8	4 18.1	#N/A	#N/A
Total	2 20-1=19	407.5	#N/A	#N/A	#N/A
Parameter	*Coefficients*	*Standard Error*	*t Stat*	*t(.025)*	
Intercept	131.517	29.552	4.450	2.101 4	
drive	-0.250	0.118	-2.119 4	2.101	

$$r^2 = \frac{81.7}{407.5} = .2005$$

3

p.4.b. What is your conclusion regarding the test of $H_0: \beta_1 = 0 \quad H_A: \beta_1 \neq 0$? **Reject H₀** / **Fail to Reject H₀**

p.4.c. Give the sample correlation, r, between Y and X.

$$r = \text{sgn}\{\hat{\beta}\}\sqrt{r^2} = -\sqrt{.2005} = -.4478$$

4

35

Q.5. In the matrix form of the simple linear regression model, the least squares estimator for β is $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ where the elements of **X** are fixed constants in a controlled laboratory experiment.

$E\{Y\} = E\{X\beta + \varepsilon\} = X\beta$ 5

p.5.a. Derive $E\{\hat{\boldsymbol{\beta}}\}$ **Show all work** 15

$Y = X\beta + \varepsilon \qquad E\{\varepsilon\} = 0 \Rightarrow E\{Y\} = X\beta$

$\hat{\beta} = AY \qquad A = (X'X)^{-1}X' \qquad E\{AY\} = AE\{Y\}$

$= (X'X)^{-1}X'X\beta = I\beta = \beta$

p.5.b. Derive $V\{\hat{\boldsymbol{\beta}}\}$ **Show all work** 15 $V\{Y\}$ 5

$V\{Y\} = V\{X\beta + \varepsilon\} = 0 + V\{\varepsilon\} = \sigma^2 I$

$V\{AY\} = AV\{Y\}A' = (X'X)^{-1}X'\sigma^2 I\left((X'X)^{-1}X'\right)'$

$= (X'X)^{-1}X'\sigma^2 I X(X'X)^{-1}$

$= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} = \sigma^2(X'X)^{-1}$

30

Q.6. Based on Model 1, demonstrate the following results for summations.

p.6.a. Show that $\sum_{i=1}^{n}\left(X_i-\overline{X}\right)=0$

10

$$\sum_i (X_i-\overline{X}) = \sum_i X_i - n\overline{X} = \sum_i X_i - n\frac{\sum_i X_i}{n} = \sum_i X_i - \sum_i X_i = 0$$

p.6.b. Show that $\sum_{i=1}^{n}\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)=\sum_{i=1}^{n}\left(X_i-\overline{X}\right)Y_i$

10

$$\sum_i (X_i-\overline{X})(Y_i-\overline{Y}) = \sum_i (X_i-\overline{X})Y_i - \sum_i (X_i-\overline{X})\overline{Y}$$

$$= \sum_i (X_i-\overline{X})Y_i - \overline{Y}\sum_i (X_i-\overline{X}) = \sum_i (X_i-\overline{X})Y_i - 0$$

Q.7. An experiment is conducted, relating weekly sales for a food delivery (Y) service to the amount of advertising (X) during the week. The results for a sample of n = 6 weeks are given below. Fit Model 2 in matrix form by filling in the following matrices.

$\sum_i 1 = 6$ $\quad \sum_i Y_i = 270$

$\sum_i X_i = 24$ $\quad \sum_i X_i Y_i = 1240$

$\sum_i X_i^2 = 112$

Week	1	2	3	4	5	6
Ad Spend (X)	2	2	4	4	6	6
Sales (Y)	20	30	40	50	70	60

8 $\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 4 \\ 1 & 4 \\ 1 & 6 \\ 1 & 6 \end{bmatrix}$ 3 $\mathbf{Y} = \begin{bmatrix} 20 \\ 30 \\ 40 \\ 50 \\ 70 \\ 60 \end{bmatrix}$ 5 $\mathbf{X'X} = \begin{bmatrix} 6 & 24 \\ 24 & 112 \end{bmatrix}$ 5 $\mathbf{X'Y} = \begin{bmatrix} 270 \\ 1240 \end{bmatrix}$

$|X'X| = \frac{1}{672-576} = \frac{1}{96}$

8 $(\mathbf{X'X})^{-1} = \frac{1}{96}\begin{bmatrix} 112 & -24 \\ -24 & 6 \end{bmatrix}$ 5 $\hat{\boldsymbol{\beta}} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$ 5 $\hat{\mathbf{Y}} = \begin{bmatrix} 25 \\ 25 \\ 45 \\ 45 \\ 65 \\ 65 \end{bmatrix}$ 5 $\mathbf{e} = \begin{bmatrix} -5 \\ 5 \\ -5 \\ 5 \\ 5 \\ -5 \end{bmatrix}$

$SSE = \mathbf{e'e} =$

$= 6(5)^2 = 6(25) = 150$

$$\hat{\beta} = \frac{1}{96}\begin{bmatrix} 112(270) - 24(1240) \\ -24(270) + 6(1240) \end{bmatrix} = \frac{1}{96}\begin{bmatrix} 480 \\ 960 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

54 5