COVER PAGE

STATISTICS 608 - EXAM 1

June 12/13 2017

Student's Name (PRINT) RAJAN KAPOOR

Student's Email Address (PRINT) r.kapoor@tamu.edu

INSTRUCTIONS TO STUDENTS:

(1) Write your answers in the spaces provided on the examination paper. The last two (blank) pages can be used for rough work or as additional space for answers. You have exactly 75 minutes to complete the exam within the time frame 12:01 PM, CDT, 6/12/2017 to 12:01 PM, CDT 6/13/2017. **The exam starts AFTER you have downloaded and printed it**. If you have been granted extra time by Disability Services, your proctor will have been informed accordingly.

(2) Upon accessing the exam, you have a 30-minute buffer in which to download, print, scan and upload it to Webassign. You may not work on the exam during this time.

(3) You may use your own computer together with a pocket calculator and/or any software package already loaded on your computer to do calculations.

(4) The exam is OPEN BOOK. You may make use of the textbook and any other material that you saw fit to prepare beforehand, either as hard copy or on your PC/Laptop.

(5) **You may not access the internet other than to download and to upload the exam.**

I attest that I spent no more than 75 minutes to complete the exam. I did not access the internet during the exam nor did I receive assistance from anyone during the exam. I promise not to discuss or provide any information to anyone concerning any aspect of this exam until after 3:31 PM on 6/13/2017.

Student's Signature Rajan

INSTRUCTIONS TO PROCTOR:

The exam starts only **after** the student has downloaded and printed it. **Immediately** after the exam ends, have the student scan the exam **with this cover sheet on top** to a PDF file and upload it to Webassign.

(1) I certify that the student's exam start time was 5:16 pm, and the exam completion time was 6:24 pm

(2) I certify that the student has followed all the **INSTRUCTIONS TO STUDENTS** listed above.

(3) I certify that the exam was scanned into a PDF and uploaded to Webassign **in my presence**.

Proctor's Printed Name Remote Proctor

Proctor's Signature Rajan

Date 06/12/17

STATISTICS 608 - Examination 1

June 12/13, 2017 - Duration: **75 MINUTES**

Total points available: **22** (**20** points =100%)

$\frac{21}{20}$

PARTIAL CREDIT WILL ACCRUE FOR ALL RELEVANT WORK SHOWN.
This paper consists of six (6) pages (including two pages for rough work).

Question 1 [1/2+1/2+1/2+1/2+2=4 points] $5\frac{1}{2}$

This question is based on the attached data set *Exam1data.txt* (also available as *Exam1data.xls* and *Exam1data.csv*):

Y is the response variable, x is the predictor variable.
In the following, you *need not show* any R, SAS or other code.

Fit a simple linear regression model $Y = \alpha + \beta x + e$ to the data and, without regard to whether the model fit is good or not,

1.1 write the least squares estimates of the intercept and slope;

$\hat{\beta}_0 = -286.21$ (int) $\hat{\beta}_1 = 372.87$ (slope)

1.2 write the standard error of the estimated slope.

35.57

1.3 write the numerical value of R^2. $1 - \frac{RSS}{SST}$, $1 - \frac{918277}{918277 + 2147272} = 0.7005$

1.4 write your **estimate** of the error variance, i.e. of the variance of e; $\frac{RSS}{n-2} = 19538$

1.5 Write the 95% **confidence interval** for $E[Y|x = 1.7]$.

(301.0577, 394.2956)

1.6 Write a 90% **prediction interval** for a new observation Y^* which is to be made at $x = 1.7$.

9 typo?

(107.9386, 585.4148)

Question 2 [2 points] If $\boxed{2}$

$$n = 16, \sum x_i^2 = 978.142 \text{ and } \bar{x} = \quad 1.027,$$

then $SXX =$ $\sum x_i^2 - n\bar{x}^2 = 978.142 - 16(1.027) = 961.71$ ✓

Question 3 [1/2+1+1+1+1/2=4 points] $\boxed{3\frac{1}{2}}$

A linear regression model

$$Y = \beta_0 + \beta_1 x + e$$

was fit to 21 pairs of (Y, x) observations. For these data

$\sum x_i^2 - n\bar{x}^2$

$$RSS = 261.855; \; \bar{x} = \quad 1.03, \; SXX = 961.266.$$

To test the hypothesis $H_0 : \beta_1 = 0.1$ at the 5% level of significance, (two tailed t test)

a the degrees of freedom are $\nu =$ 19 $(21-2)$ ✓

b the critical value = $t_{0.05/2, 19} = 2.093$ ✓

c the unbiased estimate of the error variance = $\frac{RSS}{n-2} = \frac{261.855}{19} = 13.7818$ ✓

d the numerical value of the test statistic =

$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{SXX} = \frac{\sum(x_i - \bar{x})y_i}{SXX} = \frac{\sum x_i y_i - \bar{x}\sum \bar{y}}{SXX}$ ✓

e The outcome of the test is (a) reject H_0; (b) accept H_0. ?

$\frac{\hat{\beta}_1 - 0.1}{se(\hat{\beta}_1)} = \frac{\hat{\beta}_1 - 0.1}{\frac{S}{\sqrt{SXX}}} = \frac{\hat{\beta}_1 - 0.1}{\frac{3.712}{31}}$ ✓

Question 4 [4 points] $\boxed{4}$

Suppose Y has mean $\mu > 0$ and variance $g(\mu)$ and that $1/Y$ has variance equal to 1. Find the function g.

$$Var(f(Y)) = [f'(\mu)]^2 \, Var(Y)$$ ✓

$f(Y) = \frac{1}{Y} \qquad f' = -1/y^2$ ✓

$$Var\left(\frac{1}{Y}\right) = \left.\frac{1}{y^4}\right|_{y=\mu} g(\mu)$$

$$\Rightarrow 1 = \frac{1}{\mu^4} g(\mu) \Rightarrow g(\mu) = \mu^4$$ ✓

Question 5 [1+1+4=6 points] 6

In a simple linear regression model

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

set

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

for $i = 1, \ldots, n$ and

$$\bar{y} = \frac{y_1 + \cdots + y_n}{n}, \quad \bar{x} = \frac{x_1 + \cdots + x_n}{n}.$$

You may take the expressions for $\hat{\beta}_0$ and $\hat{\beta}_1$ on pages 18 and 19 in the textbook as given. Show that

5.1 $\quad y_i - \hat{y}_i = (y_i - \bar{y}) - \hat{\beta}_1(x_i - \bar{x})$

$$y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$$
$$= y_i - \bar{y} + \hat{\beta}_1 \bar{x} - \hat{\beta}_1 x_i$$
$$= (y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x}) \checkmark$$

and that

5.2 $\quad \hat{y}_i - \bar{y} = \hat{\beta}_1(x_i - \bar{x})$.

$$\hat{y}_i - \bar{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i - (\hat{\beta}_0 + \hat{\beta}_1 \bar{x})$$
$$= \hat{\beta}_0 + \hat{\beta}_1 x_i - \cancel{\hat{\beta}_0} - \hat{\beta}_1 \bar{x}$$
$$= \hat{\beta}_1 (x_i - \bar{x}) \checkmark$$

Continued on next page.

Hence, show that

5.3 $\quad \sum_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = 0.$

$$LHS = \sum_{i=1}^{n} \left[(y_i - \bar{y}) - \hat{\beta}_1 (x_i - \bar{x})\right]\left[\hat{\beta}_1 (x_i - \bar{x})\right]$$

$$= \hat{\beta}_1 \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x}) - \hat{\beta}_1^2 \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \hat{\beta}_1 \underbrace{\left(\hat{\beta}_1 \times \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)} - \hat{\beta}_1^2 \times \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\hat{\beta}_1 = \frac{Sxy}{Sxx} \Rightarrow Sxy = \hat{\beta}_1 (Sxx).$$

$$= 0 \quad ✓$$

Space for rough work or additional calculations

$\Sigma x_i y_i -$ $\quad x_i (y_i - \bar{y}) - \bar{x}$

$\Sigma y_i (x_i - \bar{x})$

$RSS = \Sigma (y_i - \hat{y}_i)^2$ $\quad Sxx = \Sigma (x_i - \bar{x})^2$

$= \Sigma x_i^2 - n\bar{x}^2$

$\Sigma y_i^2 +$ $\quad \Rightarrow \Sigma x_i^2 = n\bar{x}^2 + Sxx$

$-$

$RSS = \Sigma (y_i - \bar{y})^2 + -$

$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$

$\Sigma x_i y_i - \bar{x} \Sigma y_i$

$\Sigma (y_i - \hat{y}_i)^2 = \Sigma (y_i - \hat{\beta}_0 - \hat{\beta}_1 x)^2$

$\downarrow$

$\hat{\beta}_0 - \hat{\beta}_1 x$ $\quad \Sigma (y_i - \bar{y})^2 + \Sigma (\hat{y}_i - \bar{y})^2$

Space for rough work or additional calculations