COVER PAGE

STATISTICS 608 - EXAM 2

July 10/11 2017

Student's Name (PRINT) RAJAN KAPOOR
Student's Email Address (PRINT) t. 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 90 minutes to complete the exam within the time
frame 12:01 PM, CDT, 7/10/2017 to 12:01 PM, CDT 7/11/2017. 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 completing the exam, you have a 30-minute buffer in which to 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 <u>OPEN BOOK</u> . 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 90 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 7/11/2017.
Student's Signature 'Rojan
INSTRUCTIONS TO PROCTOR:
The exam starts only <u>after</u> the student has downloaded and printed it. <u>Immediately</u> after the exam ends, have the student scan the exam <u>with this cover sheet on top</u> to a PDF file and upload it to Webassign.
(1) I certify that the student's exam start time was 4:39 km, and the exam completion time was 6:08 AM
(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 Rojan Date 07/11/2017

图 \$46 和人人 一版作为语言读篇\$P\$ \$16 多音 " INC 17:51 - 404

	.ett.	OF	1114	A.		CT POST	ing and Armatha	
	* * * *		The second		3 4	· · · · · · · · · · · · · · · · · · ·	on the second secondary.	

ે કરો કે જ	nen ert een nogdig Madies and suit ertt benjog voorwaaren, op ope handereng met op not in geenhood. Dit onder est os
	म् वोद्येक्तर कारक प्रतार व्यवद्वतरक क र (कार्नु केंट्र) एविकालक कार कार भी उपलब्धक करें कार्यक्रिक प्रतार कारक कार कार्यक्र
\$ve-67€	विक्रिक स्थापक के कारण प्रयोगनायाँ वृद्धिर्विद्धपनि पूर्व के इसके स्थापक एक के प्राप्त करते हैं। यह कार्यक्षिक सम्बद्धि से पहले
	and the second of the second
Zaruta	en en la completa de la companya de
	e de la companya della companya dell
	र्वहरूका वर्षकार अपने करकार मुख्य केवा का पृत्य प्रावशास कुर न असूर का या कुम्पून अस्तुमारीस उठाएक किया का करक करान किया । विकास के प्रावस के
	water of professional and the second professional and the
orașev-	givesti was dongotti metatati. Listu yar ipare noditari berite merusirat yina beri oligifisi jeliki yelashire di eshi e
	ાલ કર્યું કરો છે. જે સામાનું તાલ કરતા મુખ્યત્વે કરા કરાકો કર કરાકો કરા કેક્સ સામાનો માટે કરો છે. જે માટે કરો ક
	্ত প্ৰত্য হৈ হৈছে গল্পুৰ কৰ্ম হৈছে। ইয়াৰ ইয়াৰ কাৰ্যাৰ ক্ষেত্ৰৰ কৰা সন্তান প্ৰতীপ ক্ষেত্ৰৰ ইয়াৰ কুনাই কৰি সি
i Wes	rangangan penggahaki dalahkat di proposisson fidi kangarangan sekrangangan di pengrapangan kepingan kepingan p
	es <mark>angan de edunario de ejul distensa de esa este on este adredis francos de guidas diapejas</mark> eleder estado estado
	· · · · · · · · · · · · · · · · · · ·

Section 1			F3*	5745	 · · · ·	;;3 s

The constant of progress of the constant of the constant of the progress of the constant of th

MARCH Process and reductions assess to the first first was not been now a consider the estimation by

· Professional of TRUCKER Of Section of the FRESSION of Basewall the Section of Section Of Fresh (Section 1)

AND THE THE SECOND CONTRACTOR OF A PROCESSING AND A PARTY OF THE PROPERTY OF A SECOND CONTRACTOR OF THE PROPERTY OF THE PROPER

STEED TO THE CONTRACT OF SERVICE STATES OF

SM STATE		ng j	nîza,()

STATISTICS 608 Examination 2 Summer 2017

Duration: 90 MINUTES

Total points available: 35 (32 points = 100%)

SHOW ALL CALCULATIONS AND EXPLANATIONS. PARTIAL CREDIT WILL ACCRUE FOR ALL RELEVANT WORK SHOWN.

This paper consists of seven (7) pages.

covariate linear model with intercept:

Question 1 [3] Look at the following output from fitting to observed responses a three -

covariate	est. coef.	st.err.	t	sig
X_1	0.5	0.5	1.0	0.33
X_2	1.0	0.25	4.0	0.00
X_3	0.4	0.4	1.0	0.33

The statement: "Both X_1 and X_3 should be dropped from the model because neither has a significant t-statistic" is (choose one) (a) justified (b) not justified. Explain *briefly* your reasoning.

X; and X3 donot have significant t-statistic but they (both of them) cannot be dropped simultaneously on the basis of t-statistic because the significance might change when one of them is dropped. Partial-f test should be used to make this decision.

(b) not justified

ALANDERS SEELES SEELES SEELES

THE ENGINE PROOF SE STATE OF THE SECOND SECO

ABBOR AND ONESKATIONS AND EMPLANATIONS BY BY THE CHEST WEST VOICE SERVE ALL GERG GOVERN WE AN ANDWELL

alogo pri i i i na anolica como casa casa con consiste de consiste

Assault a seconomy to meson or product over ample compatible ledit to deed for a self-basella or necesiality observations made in a

			643 394			
[34.0]	400		23	- f- · · · ·	, y	
140.0		5. A	Ç. Î	Ann ann an		

s sai mitina marami tehun ketaman hangana berganak berganak 18. Sai 18. Sai ketak banasak ketak berak ketak 1 Sajahan pangan dining iku interimpun pengalah berakan ketah berak interimpuntah diningan ketah ketah ketah sa Banasan **Question 2** [12] The *Question 2 Exam 2* data show the numerical values of a predictor y and three covariates x_1 , x_2 and x_3 . Assume that a valid linear regression

$$E[y|x_1,x_2] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

is in force.

Test the null hypothesis $H_0: \beta_2 = 2\beta_1$ at the 10% level of significance. For partial credit you need to show your calculations and briefly outline your method: numerator and denominator of your test statistic [4+2=6]; applicable degrees of freedom [2+1=2], numerical value of your test statistic [1], critical value [1] and decision [1]. $\beta_0 + \beta_1 (\alpha_1 + 2\alpha_2) + \beta_3 \alpha_4 \qquad \alpha = 0.$

of = 20-4 = 16/

RSS fuy = 12-7244

ofreg = 20-3=17/

RSSnea = 16.9028/

F = (RSSner-RSSfuy) = RSSfum dfrum

= $(4.1784) \div (0.7953)$

z 5.2539 V

fc = F,6,17,01 = 1.8997 & 1.9 (critical value)

F>fc = reject null hypothesis

ng mendhenera, ke apad a bahasanan oda kerka adal di remai ada sasa partif bila. Dan di sasa sasa pi Bahasangan sasa bahasa sasa sasa da da da da da da sasa sasa bahasa sasa bahasa sasa di sasa sasa bahasa sasa

Barrier of Arthur High Barrier

医多类形式

And the second second of the second of the second s



Question 3 [4+4=8] Consider the Question 3 Exam 2 data. These data come from Table 8.1 in Section 8.1 of the textbook. The questions to follow have nothing to do with logistic regression.

Set
$$z_i = y_i/m_i$$
 and

$$V_i = \arcsin(\sqrt{z_i}). \tag{1}$$

It is known that $var(V_i)$ is, to good approximation, of the form C/m_i , for some constant C which does not depend upon x and that, also to good approximation,

$$E[V_i|x] = arcsin(\sqrt{\theta_i(x)})$$

where $\theta_i(x) = E[z_i|x]$. The following linear model has been proposed:

$$\arcsin(\sqrt{\theta_i(x)}) = \gamma_0 + \gamma_1 x$$

where x denotes the Michelin Guide food rating.

3.1 Find the least squares estimates of γ_0 and γ_1 and of the constant C. Explain briefly how you go about this.

Note: the arcsine function is denoted by asin in \mathbf{R} .

$$V_0 = -2.94$$

 $V_1 = 0.244$
1. C = 0.0834

 $Y_0 = -2.94$ \times 1 $(+ asin(10)) \times m$ \times 1 (-2.94) \times 1 \times

3.2 Notwithstanding your answers in 3.1, assume for this part of the question that

$$\hat{\gamma}_0 = 1.75, \ \hat{\gamma}_1 = 0.117, \ \hat{C} = 0.254.$$

Find a 90% prediction interval for z^* , the predicted z-value for a new restaurant with food rating $x^* = 21$.

Africa Common de de la composição de la

n de la companya de la co

dansamen karak aliku kelisti kemit adalah laberah lambangan indugi keladi perkalah mendalah di Ki Kalandarah labarah dan menggan teragai mengan kanda ingan kampan indapatan kalandarah di perkasa keladi di perk

Character State

i ge<mark>rmangan</mark> et i noof meel heletoore meeletoo germandoloofili on ja Personali in oo oo

The state of the s

Constant provides the All and the telephone for the section of the

a Brand and Rhynde () and the confined for the part of the confined notices and part of the confined of the co

the stable of behaviors making both in

n destinations und est direct and armenes in the information of the contract of all the contract of the contra

Little annualisticamenta sui substituta militare periodo de la comencialidade escribilidade in 172. La comencialidade de la comencialid Question 4 [2+3=5] The Question 4 Exam 2 data show the numerical values of a response variable y and two covariates x_1 and x_2 . Show a plot which indicates that the response variable y should be transformed if a linear regression model is to be fit (a rough, hand drawn version, is acceptable). Then find an appropriate transformation. You need not show any R or SAS code.

Scatter plot of y VIS X, and y VS X2 should be linear. but it is concentrated around 0.

These two plots o printe.

Residuels va fitted value

Wide variation

Non constraniance

2

je logg Voing inverse reponse pur

600

ent a la medien la dividuation del mada enclui, anno en ela elle discolario di la comencia di additiva delle di additiva di la confide dividuatione della secolaria di la dividuatione di la dividuatione di la dividuatione di additiva di la dividuatione di la di



Question 5 [4+3=7]

In the general linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ (X is $n \times q$ and $\boldsymbol{\beta}$ is $q \times 1$) denote the true, unknown to you, value of the parameter vector $\boldsymbol{\beta}$ by $\boldsymbol{\beta}^0$. The residual vector is defined by $\hat{\mathbf{e}} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ where $\hat{\boldsymbol{\beta}}$ denotes the least squares estimator of $\boldsymbol{\beta}^0$. Assume that $\mathbf{X}'\mathbf{X} = \mathbf{I}$, the identity matrix.

5.1 Show that

$$\mathbf{e}'\mathbf{e} = \hat{\mathbf{e}}'\hat{\mathbf{e}} + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0)'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0).$$

5.2 Suppose q=3. Then there are three possible fitted models if exactly two predictors are used. Given that

$$\hat{\beta}_0 = 1.1, \ \hat{\beta}_1 = 2.2, \ \hat{\beta}_2 = 0.5,$$

are the least squares estimators when all three predictors are used, write the equation of the best fitting among the three two -predictor models. Explain clearly your argument.

(5) $e'e = (Y - X\beta)'(Y - X\beta)'(Y - X\beta)'$ $= (\hat{e} + X\hat{\beta} - X\beta)'(\hat{e} + X\hat{\beta}) - (\hat{e} + X\hat{\beta})' + (\hat{e} + X\beta)'(\hat{e} + X\beta)' + (\hat{e} + X\beta)' + (\hat{e} + X\beta)'(\hat{e} + X\beta)' + (\hat{e} + X\beta)$

(5.3)

1.12, + 2.222 + 0.523

Best fitting should be

1.1(2,+22) + 0.523

Bo(21+222) + B223

Correlation among covariates I variance of estimated coefficients if x_1 and x_2 are taken separately there will be correlation high.

Space for rough work or additional calculations

energiateliko lomekilian arazaren beraria batuar.

Space for rough work or additional calculations

$$V_{i} = Y_{0} + Y_{1} \times Y_{2} \times Y_{3} \times Y_{4} \times Y_{5} \times Y_{5$$

$$e'e = (Y - x\beta')(Y - x\beta)$$

$$= (\hat{e} + x\hat{\beta} - x\beta')(\hat{e} + x\hat{\beta} - x\beta')$$

$$= (\hat{e}' + \hat{\beta}'x' - \beta'x')(\hat{e} + x\hat{\beta} - x\beta')$$

$$= (\hat{e}' + (\hat{\beta}' - \beta')x')(\hat{e} + x(\hat{\beta} - \beta'))$$

$$= (\hat{e}' + (\hat{\beta}' - \beta')x')(\hat{e} + x(\hat{\beta} - \beta')x')(\hat{e} + x(\hat{\beta} - \beta')x')(\hat{e} + x(\hat{\beta} - \beta'))(\hat{e} + x(\hat{\beta} - \beta')x')(\hat{e} + x($$