

AMS 315, Fall 2020, Examination 2, October 15, 2020

Instructions:

This examination has 6 problems, worth a total of 320 points. The point value of each question is given at the end of the problem. You may use your notes, any texts of your choice, and any calculator of your choice. Work your problems on separate sheets of paper. Please put your name, Stony Brook identification number, and where you are while taking this examination in the upper right-hand corner of your first sheet of your submitted paper. You may not use any other assistance.

The forms change between sets of questions. The form is given on the top of the page of each set of questions. Make sure to specify the number and form of the problem that you are solving for each problem.

At the end of the examination, please show each page of your submitted examination work to the camera, make a pdf file of your work, and submit it. In the event that there is a problem, please e-mail the pdf of your work to Stephen.finch@stonybrook.edu, with a brief explanation of the problems that you encountered.

You are on your honor not to use any other assistance during this examination. There will be no partial credit given for a problem unless you show your work.

Since the course satisfies requirements for actuarial credentials, academic integrity standards will be enforced strictly.

AA

1. A research team took a sample of 6 observations from the random variable Y , which had a normal distribution $N(\mu, \sigma^2)$. They observed $\bar{y}_6 = 798.4$, where \bar{y}_6 was the average of the 6 sampled observations, and $s^2 = 1673.1$ was the observed value of the unbiased estimate of σ^2 , based on the sample values. Find the 99% confidence interval for σ^2 . This problem is worth 50 points.

5DF

$$.99 = P_{\lambda} \left\{ .4117 < \frac{5s^2}{\sigma^2} < 16.75 \right\} \text{ FROM TABLE 7.}$$

$$= P_{\lambda} \left\{ \frac{1}{16.75} < \frac{\sigma^2}{5s^2} < \frac{1}{.4117} \right\}$$

$$= P_{\lambda} \left\{ \frac{5s^2}{16.75} < \sigma^2 < \frac{5s^2}{.4117} \right\}$$

HENCE 99% CI FOR σ^2 IS

$$\frac{5(1673.1)}{16.75} = 499.4 \text{ TO } \frac{5(1673.1)}{0.4117} = 20,319.4.$$

- 25 WRONG DF WITH CORRECT CHOICE OF PERCENTILES.
- 50 WRONG PROCEDURE
- 20 NO 16.75
- 20 NO .4117
- 50 USED T DISTRIBUTION
- 10 WRONG DF

[illegible]

1. A research team took a sample of 3 observations from the random variable Y , which had a normal distribution $N(\mu, \sigma^2)$. They observed $\bar{y}_3 = 228.9$, where \bar{y}_3 was the average of the 3 sampled observations, and $s^2 = 2275.6$ was the observed value of the unbiased estimate of σ^2 , based on the sample values. Find the 95% confidence interval for σ^2 . This problem is worth 50 points.

$$DF = 2$$

$$P_2 \{ .05064 < \frac{2S^2}{\sigma^2} < 7.378 \} = 0.95$$

$$0.95 = P_n \left\{ \frac{1}{7.378} < \frac{G^2}{29^2} < \frac{1}{.05064} \right\}$$

$$= \rho_n \left\{ \frac{2S^2}{7.378} < \sigma^2 < \frac{2S^2}{.05064} \right\}$$

HENCE 95% CI FOR G^2 IS

$$\frac{2(2275.6)}{7.378} = 616.9 \text{ TO } \frac{2(2275.6)}{.05064} = 89,873.6$$

= 10 WRONG DF.

- 50 T TEST

-10 NO .05064

-10 NO 7.378

-50 USED ∇ DISTRIBUTION

-15 RECIPROCAL WRONG PERCENTAGE

2. A research team took a random sample of 5 observations from a normally distributed random variable Y and observed that $\bar{y}_5 = 231.2$ and $s_Y^2 = 1,138.7$, where \bar{y}_5 was the average of the five observations sampled from Y and s_Y^2 was the unbiased estimate of $\text{var}(Y)$. A second research team took a random sample of 3 observations from a normally distributed random variable X and observed that $\bar{x}_3 = 1491.8$ and $s_X^2 = 6891.3$, where \bar{x}_3 was the average of the three observations sampled from X and s_X^2 was the unbiased estimate of $\text{var}(X)$. Find the 95% confidence interval for $\frac{\text{var}(X)}{\text{var}(Y)}$.

$$TS = \frac{S_y^2 / \sigma_y^2}{S_x^2 / \sigma_x^2} \sim F(4, 2)$$

$$P\left\{\frac{1}{10.65} < F_{4,2} < 39.25\right\} = 0.95.$$

$$= P_r \left\{ \frac{1}{10.65} < \frac{S_y^2}{S_x^2} \cdot \frac{\sigma_x^2}{\sigma_y^2} < 39.25 \right\} = 0.95$$

$$= P_n \left\{ \frac{1}{10.65} \frac{S_X^2}{S_Y^2} < \frac{\sigma_X^2}{\sigma_Y^2} < 39.25 \frac{S_X^2}{S_Y^2} \right\} = 0.95$$

NOTE $\frac{S_x^2}{S_y^2} = \frac{6891.3}{1138.7} = 6.052$

THE 95% CI FOR $\frac{\text{VAR}(X)}{\text{VAR}(Y)}$ IS

$$\frac{1}{10.65}(6.052) = 0.568 \text{ TO } 39.25(6.052) = 237.5$$

- 20 EACH DE ERROR.

-15 RECIPROCAL WRONG PERCENTAGE

-50 two SAMPLE T-TEST

-10 NO 10.65

-10 NO 39.25.

2. A research team took a random sample of 6 observations from a normally distributed random variable Y and observed that $\bar{y}_6 = 638.7$ and $s_Y^2 = 4,638.7$, where \bar{y}_6 was the average of the six observations sampled from Y and s_Y^2 was the unbiased estimate of $var(Y)$. A second research team took a random sample of 4 observations from a normally distributed random variable X and observed that $\bar{x}_4 = 499.8$ and $s_X^2 = 8,491.3$, where \bar{x}_4 was the average of the four observations sampled from X and s_X^2 was the unbiased estimate of $var(X)$. Find the 99% confidence interval for $\frac{var(X)}{var(Y)}$.

$$TS = \frac{S_Y^2 / \sigma_Y^2}{S_X^2 / \sigma_X^2} \sim F(5, 3)$$

$$0.99 = P\left\{ \frac{1}{16.53} < F(5,3) < 45.39 \right\}$$

$$= P_n \left\{ \frac{1}{16.53} < \frac{S_y^2}{S_x^2} \cdot \frac{\sigma_x^2}{\sigma_y^2} < 45.39 \right\}$$

$$= P_{\alpha} \left\{ 16.53 \left(\frac{S_X^2}{S_Y^2} \right) < \frac{\sigma_X^2}{\sigma_Y^2} < 45.39 \left(\frac{S_X^2}{S_Y^2} \right) \right\}$$

NOTE $\frac{S_{\Sigma}^2}{S_r^2} = \frac{8491.3}{4638.7} = 1.83$

THE 99% CI FOR $\frac{\text{VAR}(X)}{\text{VAR}(Y)}$ IS

$$\frac{1}{16.53} (1.83) = 0.11 \quad \text{TO} \quad 45.39 (1.83) = 83.1$$

- 20 EACH DE ERROR

-15 RECIPROCAL WRONG PERCENTILE

-50 T-TEST

-10 NO 45.39

-10 NO 16.53

[illegible]

3. A research team studied the response of a participant to a dosage of medication. Dosages were randomly assigned to participants. The research team then measured each participant's response for $n = 365$ participants. The average response was 145.3, with an observed standard deviation of 38.4 (the divisor in the underlying variance calculation was $n - 1$). The average dosage was 62.7, with an observed standard deviation of 11.5 (the divisor in the underlying variance calculation was also $n - 1$). The correlation coefficient between the two variables was 0.56. The team sought to estimate the regression of participant response on the dosage of medication.
- Complete the analysis of variance table for the regression of participant response on the dosage of medicine given the participant. Test the null hypothesis that the slope of this regression is zero at levels of significance 0.10, 0.05, and 0.01. This part is worth 30 points.
 - Find the estimated regression equation of participant response on dosage. Find the 95% confidence interval for the slope in this equation. 20 points.
 - Use the least-squares equation to estimate the response for participants whose dosage was 90.0. What is the 95% confidence interval for this response score? This part is worth 20 points.

A. DV = RESPONSE, IV = DOSAGE $\sum (x_i - \bar{x}_m)^2 = 48,139.$

$SS_{TOTAL} = (n-1)(SD_{RESPONSE})^2 = 364 \times (38.4)^2 = 536,738.84$

$SS_{REG} = r^2 SS_{TOTAL} = (.56)^2 SS_{TOTAL} = 168,321.61$

$SS_{ERROR} = (1-r^2) SS_{TOTAL} = (1-.56^2) SS_{TOTAL}$

$= 368,418.23$

$MSE = SS_{ERROR} / 363 = 1014.93 = (31.86)^2$

ANOVA TABLE

SOURCE	DF	SUM OF SQUARES	MS	F
DOSAGE	1	168,321.61	168,321.61	165.85
ERROR	363	368,418.23	1,014.93	
TOTAL	364	536,739.84		

α	$F(1, 363)$	
0.10	2.719	REJECT
0.05	3.867	REJECT
0.01	6.705	REJECT

REJECT $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$

AT $\alpha = .01$ (AND .05 AND .10).

-15 DP ERROR

-10 REVERSE E DV +EV

- 10 USE $(n-1)$ SDOV AS TOPSS ENTRY.

- 10 USE (N-1)SDOV
- 10 INCORRECT ANOVA ENTRY

-25 NO OR INCONSESTENT DECISION

F2020 EXAMINATION 2 3E CONTINUED.

3b $\hat{y}(x) = \bar{y}_n + \hat{\beta}_1(x - \bar{x}_n)$.

$$\hat{\beta}_1 = \frac{38.4}{11.5} (0.56) = 1.87$$

$$\hat{y}(x) = 145.3 + 1.87(x - 62.7)$$

$$= 28.1 + 1.87x, \quad \sum (x_i - \bar{x}_n)^2 = 364(11.5)^2 = 48,139$$

95% CI FOR β_1 :

$$\hat{\beta}_1 \pm t_{1.960, 363} \sqrt{\frac{MSE}{\sum (x_i - \bar{x}_n)^2}}$$

$$= 1.87 \pm 1.967 \sqrt{\frac{1014.93}{48,139}}$$

$$= 1.87 \pm 0.286$$

95% CI FOR β_1 IS 1.584 TO 2.156.

NO FURTHER PENALTY FOR REVERSING IV + DV.
-10 FOR SUBSTANTIVE ERROR.

3c $\hat{y}(90) = 145.3 + 1.87(90 - 62.7)$

$$= 145.3 + 1.87(27.3) = 145.3 + 51.1 = 196.4$$

95% CI FOR $\beta_0 + 90\beta_1$:

$$196.4 \pm 1.967 \sqrt{\left(\frac{1}{365} + \frac{(27.3)^2}{48,139}\right) (1,014.93)}$$

$$196.4 \pm 1.967 \sqrt{(1,014.93)(.00274 + 0.0155)}$$

$$196.4 \pm 1.967 \sqrt{18.49} = 196.4 \pm 8.46$$

THE 95% CI FOR $\beta_0 + 90\beta_1$ IS 187.9 TO 204.9.

-20 FOR PREDICTION INTERVAL

-10 FOR EACH SUBSTANTIVE ERROR.

FF

3. A research team studied the response of a participant to a dosage of medication. Dosages were randomly assigned to participants. The research team then measured each participant's response $n = 245$ participants. The average response was 514.5, with an observed standard deviation of 87.4 (the divisor in the underlying variance calculation was $n - 1$). The average dosage was 118.6, with an observed standard deviation of 21.7 (the divisor in the underlying variance calculation was also $n - 1$). The correlation coefficient between the two variables was 0.66. The team sought to estimate the regression of participant response on the dosage of medication.
- Complete the analysis of variance table for the regression of participant response on the dosage of medicine given the participant. Test the null hypothesis that the slope of this regression is zero at levels of significance 0.10, 0.05, and 0.01. This part is worth 30 points.
 - Find the estimated regression equation of participant response on dosage. Find the 99% confidence interval for the slope in this equation. 20 points.
 - Use the least-squares equation to estimate the response for a participant whose dosage was 170.0. What is the 99% prediction interval for this participant's response? This part is worth 20 points.

A. $n = 245$ DV = RESPONSE IV = DOSE

	MEAN	SD	
RESPONSE	514.5	87.4	$r = 0.66$
DOSE	118.6	21.7	

$$\sum (x_i - \bar{x}_n)^2 = (n-1)SD_{IV}^2 = 244(21.7)^2 = 114,897.16$$

$$\sum (y_i - \bar{y}_n)^2 = (n-1)SD_{DV}^2 = 244(87.4)^2 = 1,863,857.44 \text{ ON } 244 \text{ DF.}$$

$$SS_{REG} = r^2(SS_{TOTAL}) = (.66)^2(1,863,857.44) = 811,896.3 \text{ ON } 1 \text{ DF.}$$

$$SS_{ERR} = (1 - r^2)(SS_{TOTAL}) = (1 - .66^2)(1,863,857.44) = 1,051,961.1 \text{ ON } 243 \text{ DF}$$

$$MS_{ERR} = \frac{SS_{ERROR}}{DF_{ERROR}} = \frac{1,051,961.1}{243} = 4329.06 = (65.8)^2$$

ANOVA TABLE				
SOURCE	DF	SS	MS	F
DOSE	1	811,896.3	811,896.3	187.5
ERROR	243	1,051,961.1	4,329.1	
TOTAL	244	1,863,857.4		

α F(1, 243)

α	F(1, 243)	Decision
.10	2.726	REJECT
.05	3.840	REJECT
.01	6.740	REJECT

REJECT $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$ AT $\alpha = .01$ (AND .05 AND .01).

F2020 EXAMINATION 2 3F CONTINUED.

- 3A PENALTIES
- 15 DE ERROR
 - 10 REVERSE DV + IV.
 - 10 USE $(n-1)SD_{IV}$ AS TOT SS
 - 10 EACH INCORRECT ANOVA ENTRY.
 - 25 NO OR INCONSISTENT DECISION.

3B. $\hat{Y}(x) = \bar{y}_m + \hat{\beta}_1(x - \bar{x}_m)$, $\hat{\beta}_1 = \frac{87.4}{21.7} \cdot 0.66 = 2.66$

$\hat{Q}(x) = 514.5 + \hat{\beta}_1(x - 118.6) = 514.5 + 2.66(x - 118.6)$

$= 199.0 + 2.66x$; RECALL $\sum(x_i - \bar{x}_m)^2 = 114,897.16$

99% CI FOR β_1

$$2.66 \pm t_{2.576, 243} \sqrt{\frac{MSE}{\sum(x_i - \bar{x}_m)^2}}$$

$$2.66 \pm 2.596 \sqrt{\frac{4329.1}{114,897.16}} = 2.66 \pm 2.596 \sqrt{0.0377}$$

$$2.66 \pm 2.596(0.194) = 2.66 \pm 0.504$$

THE 99% CI FOR β_1 IS 2.16 TO 3.16

NO FURTHER PENALTY FOR REVERSING DV + IV.

-10 EACH SUBSTANTIVE ERROR.

3C. $\hat{Y}(170) = 514.5 + 2.66(170 - 118.6)$

$= 514.5 + 2.66(51.4) = 651.22$

99% PREDICTION MARGIN = $2.596 \sqrt{MSE(1 + \frac{1}{n} + \frac{(x - \bar{x}_m)^2}{\sum(x_i - \bar{x}_m)^2})}$

$$= 2.596 \sqrt{4,329.1(1 + \frac{1}{245} + \frac{(51.4)^2}{114,897.16})}$$

$$= 2.596 \sqrt{4,329.1(1 + 0.00408 + 0.0230)}$$

$$= 2.596 \sqrt{4,329.1(1.027)} = 2.596 \sqrt{4446.3}$$

$$= 173.1$$

99% PREDICTION INTERVAL FOR $Y_F(170)$:

$$651.22 \pm 173.1 = 478.1 \text{ TO } 824.3$$

-20 FOR CI FOR $\beta_0 + 170\beta_1$

-10 EACH SUBSTANTIVE ERROR.

[illegible]

4. A research team wishes to test the null hypothesis $H_0: \rho = 0$ at $\alpha = 0.025$ against the alternative $H_1: \rho > 0$ using Fisher's transformation of the Pearson product moment correlation coefficient as the test statistic. They have asked their consulting statistician for a sample size n such that $\beta = 0.05$ when $\rho = 0.25$. What is this value? This problem is worth 50 points.

$$F(.25) = \frac{1}{2} \ln \left(\frac{1+.25}{1-.25} \right) = \frac{1}{2} \ln (1.666)$$

$$= \frac{1}{2} (.5108) = 0.2554$$

$$\sqrt{n-3} \geq \frac{1.960(1) + 1.645}{.2554 - 0} = \frac{3.605}{.2554} = 14.11$$

$$n-3 \geq 200, \quad n \geq 203$$

+15 CORRECT FL.25)

-20 NO 1.960

-20 NO 1.645.

-20 NO 1.645.
-40 REPORT 14.4 AS n , $\sqrt{14.4}$ AS n ,
SET UP

-25 WRONG PROBLEM SET UP

4. A research team wishes to test the null hypothesis $H_0: \rho = 0$ at $\alpha = 0.005$ against the alternative $H_1: \rho > 0$ using Fisher's transformation of the Pearson product moment correlation coefficient as the test statistic. They have asked their consulting statistician for a sample size n such that $\beta = 0.01$ when $\rho = 0.33$. What is this value? This problem is worth 50 points.

$$\sqrt{n-3} > \frac{2.576 \cdot 1 + 2.326 \cdot 1}{.343 - 0} = \frac{4.902}{0.343} = 14.29.$$

$n \geq 208$

-20 NO 2.576

-20 NO 2.326

-20 NO 2.326
-40 REPORT 14.29 AS m, $\sqrt{14.29}$ AS m.

-25 WRONG PROBLEM SETUP

////////////////////////////////////

5. The correlation matrix of the random variables Y_1, Y_2, Y_3, Y_4 is $\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$, $0 <$

$\rho < 1$, and each random variable has variance σ^2 . Let $W_1 = -3Y_1 - Y_2 + Y_3 + 3Y_4$, and let $W_2 = Y_1 - Y_2 - Y_3 + Y_4$. Find the variance covariance matrix of $(W_1, W_2)^T$.

This problem is worth 50 points.

$$VCV(MY) = M VCV(Y) M^T \text{ WHERE } M = \begin{bmatrix} -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$VCV(MY) = \begin{bmatrix} -3 & -1 & 1 & 3 \\ 1 & -1 & -1 & 1 \end{bmatrix} \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} M^T = \sigma^2 \begin{bmatrix} -3+3\rho & -1+\rho & 1-\rho & 3-3\rho \\ 1-\rho & -1+\rho & -1+\rho & 1-\rho \end{bmatrix} \begin{bmatrix} -3 & 1 \\ -1 & -1 \\ 1 & -1 \\ 3 & 1 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 20-20\rho & 0 \\ 0 & 4-4\rho \end{bmatrix}$$

+15 CORRECT M.

-20 EACH INCORRECT ENTRY.

-10 FORGET σ^2 THROUGHOUT, -5 MISSING AT END

-5 CALCULATION ERROR

XX

5. The correlation matrix of the random variables Y_1, Y_2, Y_3, Y_4 is $\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$,

$0 < \rho < 1$, and each random variable has variance σ^2 . Let $W_1 = -3Y_1 - Y_2 + Y_3 + 3Y_4$, and let $W_2 = -Y_1 + 3Y_2 - 3Y_3 + Y_4$. Find the variance covariance matrix of $(W_1, W_2)^T$. This problem is worth 50 points.

$$\text{vcv}(MY) = M \text{vcv}(Y) M^T \text{ WHERE } M = \begin{bmatrix} -3 & -1 & 1 & 3 \\ -1 & 3 & -3 & 1 \end{bmatrix}$$

$$\text{vcv}(MY) = \begin{bmatrix} -3 & -1 & 1 & 3 \\ -1 & 3 & -3 & 1 \end{bmatrix} \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} M^T$$

$$= \sigma^2 \begin{bmatrix} -3+3\rho & -1+\rho & 1-\rho & 3-3\rho \\ -1+\rho & 3-3\rho & -3+3\rho & 1-\rho \end{bmatrix} \begin{bmatrix} -3 & -1 \\ 3 & -3 \\ 1 & 3 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 20-20\rho & 0 \\ 0 & 20-20\rho \end{bmatrix}$$

+15 CORRECT M.

-20 EACH INCORRECT ENTRY,

-10 FORGET σ^2 THROUGHOUT; -5 MISSING AT END

-5 CALCULATION ERROR

6. A research team collected data $(y_i, x_i, w_i), i = 1, \dots, n$. They seek to fit the model $E(Y_i) = \beta_1 x_i + \beta_2 w_i$, using Ordinary Least Squares by minimizing the function

a. Find $\frac{\partial}{\partial b_1} [SS(b_1, b_2)]$ and $\frac{\partial}{\partial b_2} [SS(b_1, b_2)]$. This part is worth 20 points.

- b. Specify the system of two normal equations whose solution is $(\hat{\beta}_1, \hat{\beta}_2)$. Do not solve this system. This part is worth 30 points.

A.
$$\begin{aligned}\frac{\partial SS}{\partial b_1} &= \sum 2(y_i - b_1 x_i - b_2 w_i)(-x_i) \\ &= \sum -2(y_i - b_1 x_i - b_2 w_i)x_i \\ \frac{\partial SS}{\partial b_2} &= \sum 2(y_i - b_1 x_i - b_2 w_i)(-w_i) \\ &= \sum -2(y_i - b_1 x_i - b_2 w_i)w_i\end{aligned}$$

B NORMAL EQUATIONS

$$\sum (-2)(y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 w_i) x_i = 0$$
$$\sum (-2)(y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 w_i) w_i = 0$$

PARTIAL

A -10 FOR EACH INCORRECT PARTIAL.

A -10 FOR EACH INCORRECT PARTIALS IN A
B -15 FOR EACH INCORRECT NORMAL EQUATION

-15 FOR EACH COLOR.
NO POINTS IF PARTIALS IN A
DATE:

NO FORMS
ARE INCORRECT.