

AA

AMS 315, Examination 1
March 14, 2019

Name: SOLUTION

ID:

Directions: Write your name in the space provided. Work each problem in the space underneath the problem and on the back side of the page. You are on your honor not to use any other assistance during this examination.

You may use only the paper in this form. If you un-staple your examination, please put your name on each sheet of your examination. You may use a calculator but not a computer or cell-phone. You may also use a single sheet of notes in your handwriting that is the size of the paper in this examination. Do not make marks on the tables given to you to work this examination. Turn in your paper, your notes, and your tables at the end of the examination.

There will be no partial credit given for a problem unless you show your work. This examination is worth 300 points. There are 6 problems, and the value of each problem or part is given at the end of the problem. In the event of a fire alarm, please take your papers, exit the room, find a private place to work, and turn in your examination to me in my office (Math Tower 1-113) by 9:00 pm today. In this event, you are still on your honor not to give or receive assistance.

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

AA

1. A research team took a random sample of 3 observations from a normally distributed random variable Y and observed that $\bar{y}_3 = 40.6$ and $s_y^2 = 120.7$, where \bar{y}_3 was the average of the three observations sampled from Y and s_y^2 was the unbiased estimate of $\text{var}(Y)$ (i.e., the divisor in the variance was $n-1$). A second research team took a random sample of 4 observations from a normally distributed random variable X and observed that $\bar{x}_4 = 15.4$ and $s_x^2 = 130.8$, where \bar{x}_4 was the average of the four observations sampled from X and s_x^2 was the unbiased estimate of $\text{var}(X)$ (i.e., the divisor in the variance was $n-1$). Test the null hypothesis $H_0 : E(X) = E(Y)$ against the alternative $H_1 : E(X) \neq E(Y)$ at the 0.10, 0.05, and 0.01 levels of significance using the pooled variance t-test. This problem is worth 40 points.

$$S_p^2 = \frac{2(120.7) + 3(130.8)}{5} = \frac{633.8}{5} = 126.76 \text{ on 5DF.}$$

$$t_5 = \frac{15.4 - 40.6 - 0}{\sqrt{126.76 \left(\frac{1}{3} + \frac{1}{4} \right)}} = \frac{-25.2}{\sqrt{73.94}} = \frac{-25.2}{8.60} = -2.93$$

α	Z	t_5	
.10	1.645	2.015	R
.05	1.960	2.571	R
.01	2.576	4.032	A

REJECT $H_0 : E(X) = E(Y)$ VS

$H_1 : E(X) \neq E(Y)$ AT $\alpha = .10$ AND $.05$

ACCEPT AT $\alpha = .01$.

-30 USE NORMAL FOR STUDENT T

-20 USE ONE SIDED VALUES.

-25 DON'T INCLUDE $\frac{1}{n} + \frac{1}{m}$

-20 WRONG DF

AA

2. 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 $\bar{y}_3 = 67.2$, where \bar{y}_3 was the average of the three sampled observations and $s^2 = 143.7$ was the observed value of the unbiased estimate of σ^2 based on the sample values (i.e., the divisor in the variance was $n-1$). Find the 95% confidence interval for σ^2 . This problem is worth 40 points.

$$DF = 2.$$

$$P\{.05064 < \chi^2_2 < 7.378\} = 0.95.$$

$$P\{.05064 < \frac{\sum (Y_i - \bar{Y}_n)^2}{\sigma^2} < 7.378\} = 0.95.$$

$$P\{\frac{1}{7.378} < \frac{\sigma^2}{\sum (Y_i - \bar{Y}_n)^2} < \frac{1}{.05064}\}$$

95% CI FOR σ^2 :

$$\frac{2(143.7)}{7.378} \text{ TO } \frac{2(143.7)}{.05064}$$

38.95 TO 5675.36. IS THE

$$95\% \text{ CI FOR } \sigma^2.$$

-20 WRONG DF.

-15 DON'T USE .05064; -5 FOR USING 0.5064.

-15 DON'T USE 7.378.

-40 ONE SAMPLE T-TEST

BB

2. A research team took a random sample of 5 observations from a normally distributed random variable X and observed that $\bar{x}_5 = 382.5$ and $s_x^2 = 67.2$, where \bar{x}_5 was the average of the five observations sampled from X and s_x^2 was the unbiased estimate of $\text{var}(X)$ (i.e., the divisor in the variance was $n-1$). A second research team took a random sample of 3 observations from a normally distributed random variable Y and observed that $\bar{y}_3 = 331.8$ and $s_y^2 = 740.8$, where \bar{y}_3 was the average of the three observations sampled from Y and s_y^2 was the unbiased estimate of $\text{var}(Y)$ (i.e., the divisor in the variance was $n-1$). Test the null hypothesis $H_0 : \text{var}(Y) = \text{var}(X)$ against the alternative $H_1 : \text{var}(Y) > \text{var}(X)$ at the 0.10, 0.05, and 0.01 levels of significance. This problem is worth 40 points.

$$TS = \frac{s_y^2}{s_x^2} = \frac{740.8}{67.2} = 11.02.$$

α	$F(2, 4)$		REJECT $H_0 : \text{VAR}(X) = \text{VAR}(Y)$
.10	4.32	R	VS $H_1 : \text{VAR}(Y) > \text{VAR}(X)$ AT
.05	6.94	R	$\alpha = .10$ AND .05. ACCEPT AT 0.01.
.01	18.00	A	

-20 WRONG NUMERATOR DF

-20 WRONG DENOMINATOR DF

-40 TWO SAMPLE T.

-30 CRITICAL VALUES FROM OTHER THAN AN F DISTRIBUTION

-20 TWO SIDED TEST

-30 USE $\frac{s_y^2}{s_x^2}$ WITH (4, 2) DEGREES OF FREEDOM

A) -18 WRONG OR NO CONCLUSION

-10 EACH INCORRECT ANOVA TABLE ENTRY.

-15 FOR REVERSING IV AND DV (TSS = 1322.336)

NO FURTHER POINTS OFF IF CONSISTENT (I.E. $F = 24.41$).
 AA

3. A research team collected data on $n = 216$ participants, who were between 50 and 55 years of age. Each participant reported the average time per week spent watching a screen (TV screen, electronic game, computer, smart phone, etc.) in the last month. The average screen time reported was 11.7 hours, with an observed standard deviation of 2.48 hours (the divisor in the underlying variance calculation was $n - 1$). Each participant also took a test of memory. The average memory score was 98.7, with an observed standard deviation of 28.4 (the divisor in the underlying variance calculation was $n - 1$). The Pearson product moment correlation coefficient between the two variables was -0.32. The research team seeks to estimate the regression of memory score on hours spent watching a screen.

- Complete the analysis of variance table for the regression of memory score on average time watching a screen and test the null hypothesis that the slope is zero at levels of significance 0.10, 0.05, and 0.01. This part is worth 20 points.
- Find the estimated regression equation of memory score on average time spent watching a screen. Find the 99% confidence interval for the slope in this equation. This part is worth 20 points.
- Use the least-squares prediction equation to estimate the memory score for a participant whose average time spent watching a screen was 16 hours. Give the 99% prediction interval for the memory score of this participant (whose average time spent watching a screen was 16 hours). This part is worth 20 points.

$$A. TSS = 215 (28.4)^2 = 173,410.4 \quad \text{NOTE } \sum (x_2 - \bar{x}_n)^2 = 215 (2.48)^2 = 1322.336$$

$$SS_{REG} = r^2 TSS = (-0.32)^2 TSS = 17,757.225$$

$$SS_{ERR} = (1 - r^2) TSS = (1 - (-0.32)^2) TSS = 155,653.175$$

$$MSE = \frac{SS_{ERR}}{214} = 727.351 = (26.969)^2$$

ANOVA TABLE

SOURCE	DF	SS	MS	F
REG	1	17,757.225	17,757.225	24.41
ERROR	214	155,653.175	727.351	
TOTAL	215	173,410.4		

α	$F(1, 215)$	$F(1, 240)$	$F(1, \infty)$	REJECT
.10	2.729 R	2.73	2.71	$H_0: \beta_1 = 0$ $vs H_1: \beta_1 \neq 0$
.05	3.885 R	3.88	3.84	At $\alpha = .01$
.01	6.754 R	6.74	6.63	(.05 + .10)

B).

$$\hat{\beta}_1 = r \frac{s_{DV}}{s_{XV}} = (-0.32) \frac{28.4}{248} = (-3.665)$$

AA

$$\hat{Y}(x) = 98.7 + (-3.665)(x - 11.7)$$

$$= 141.6 - 3.665x$$

$$\text{VAR}(\hat{\beta}_1) = \frac{\text{MSE}}{\sum(x_i - \bar{x}_m)^2} = \frac{727.351}{1322.338} = 0.5500$$

$$\text{SE}(\hat{\beta}_1) = \sqrt{.5500} = 0.742$$

99% CI FOR β_1

$$-3.665 \pm 2.600(0.742) = -3.665 \pm 1.928$$

$$-5.593 \text{ TO } -1.737$$

$$c) \hat{Y}(16) = 98.7 - (3.665)(16 - 11.7) = 98.7 - 15.76 = 82.9$$

$$\sqrt{1 + \frac{1}{216} + \frac{(16 - 11.7)^2}{1322.338}} = \sqrt{1 + .00463 + \frac{18.49}{1322.338}}$$

$$= \sqrt{1 + .00463 + 0.01398} = \sqrt{1.0186} = 1.0093$$

99% PI FOR $Y_F(16)$:

$$82.9 \pm 2.600 \sqrt{727.351} (1.0093)$$

$$82.9 \pm 70.8$$

$$12.1 \text{ TO } 153.7$$

DEDUCTIONS.

B NO POINTS OFF IF B+C CALCULATIONS ARE CONSISTENT WITH A.

-10 WRONG $\hat{\beta}_1$, WRONG $\hat{\beta}_0$

-10 WRONG $\text{SE}(\hat{\beta}_1)$

-10 WRONG CI FOR $\hat{\beta}_1$.

C -20 USE CI FOR $\beta_0 + 16\beta_1$

-10 SUBSTANTIVE FORMULA ERRORS

A -18 WRONG OR NO CONCLUSION.

-10 EACH INCORRECT ANOVA TABLE ENTRY.

-15 FOR REVERSING IV+DV (I.E. TSS = 1783.782)

NO FURTHER DEDUCTION IF CONSISTENT (I.E. F = 76.62)
BB

3. A research team collected data on $n = 456$ participants, who were between 50 and 55 years of age. Each participant reported the average time per week spent watching a screen (TV screen, electronic game, computer, smart phone, etc.) in the last month. The average screen time reported was 8.7 hours, with an observed standard deviation of 1.98 hours (the divisor in the underlying variance calculation was $n - 1$). Each participant also took a test of memory. The average memory score was 128.7, with an observed standard deviation of 31.4 (the divisor in the underlying variance calculation was $n - 1$). The Pearson product moment correlation coefficient between the two variables was -0.38. The research team seeks to estimate the regression of memory score on hours spent watching a screen.

- Complete the analysis of variance table for the regression of memory score on average time watching a screen and test the null hypothesis that the slope is zero at levels of significance 0.10, 0.05, and 0.01. This part is worth 20 points.
- Find the estimated regression equation of memory score on average time spent watching a screen. Find the 99% confidence interval for the slope in this equation. This part is worth 20 points.
- Use the least-squares prediction equation to estimate the expected memory score for participants whose average screen watching time was 15 hours per week. Give the 99% confidence interval for the expected memory score for the participants who spent 15 hours per week on average watching screens. (20 points)

$$A) TSS = (455)(31.4)^2 = 448,611.8 \quad \sum (x_i - \bar{x}_m)^2 = 455(1.98)^2 = 1783.782$$

$$SS_{REG} = (r^2)TSS = (-0.38)^2 TSS = 64,779.5439$$

$$SSE = (1 - r^2)TSS = (0.8556)TSS = 383,832.2561$$

$$MSE = \frac{SSE}{454} = 845.4455 \text{ ON } 454 \text{ DF.} = (29.077)^2$$

ANOVA TABLE				
SOURCE	DF	SS	MS	F
REG	1	64,779.5439	64,779.5439	76.62
ERROR	454	383,832.2561	845.4455	
TOTAL	455	448,611.8		

α	$F(1, \infty)$	$F(1, 454)$	
.10	2.71	2.717	R
.05	3.84	3.862	R
.01	6.63	6.691	R

REJECT $H_0: \beta_1 = 0$ VS $H_1: \beta_1 \neq 0$
AT $\alpha = .01$ (+ .05 + .10).

B FORM

$$3B: \hat{\beta}_1 = (-0.38) \frac{31.4}{1.98} = -6.026$$

$$\hat{y}(x) = 128.7 - 6.026(x - 8.7)$$

$$= 181.13 - 6.026x$$

$$\widehat{\text{VAR}}(\hat{\beta}_1) = \frac{845.4455}{1783.782} = 0.474 \quad \text{SE}(\hat{\beta}_1) = \sqrt{0.474} = 0.6885$$

$$\begin{aligned} 99\% \text{ CI FOR } \beta_1: -6.026 \pm 2.587(0.6885) &= -6.026 \pm 1.781 \\ &= -7.807 \text{ TO } -4.245 \end{aligned}$$

$$\begin{aligned} 3C: \hat{y}(15) &= 128.7 - 6.026(15 - 8.7) = 128.7 - 6.026(6.3) \\ &= 90.7362 \end{aligned}$$

99% CI FOR $\beta_0 + 15\beta_1$

$$\begin{aligned} &90.7362 \pm 2.587 \sqrt{845.4455 \left(\frac{1}{456} + \frac{(6.3)^2}{1783.782} \right)} \\ &= 90.7362 \pm 2.587(29.077) \sqrt{.002193 + .0225} \\ &= 90.7362 \pm 2.587(29.077)(0.1563) = 90.7362 \pm 11.76 \\ &= 78.97 \text{ TO } 11.76 \end{aligned}$$

DEDUCTIONS

B NO POINTS OFF IF B+C CALCULATIONS ARE CONSISTENT WITH A.

-10 WRONG $\hat{\beta}_1, \hat{\beta}_0$

-10 WRONG SE($\hat{\beta}_1$)

-10 WRONG CI FOR $\hat{\beta}_1$

C -20 USE AT RATHER THAN CI

-10 SUBSTANTIVE FORMULA ERRORS

AA

4. In a clinical trial, $2J$ patients suffering from an illness will be randomly assigned to one of two groups so that J will receive an experimental treatment and J will receive the best available treatment. The random variable X is the response of a patient to the experimental medicine, and the random variable B is the response of a patient to the best currently available treatment. Both X and B are normally distributed with $\sigma_X = \sigma_B = 500$. The null hypothesis to be tested is that $E(X) - E(B) = 0$ against the alternative that $E(X) - E(B) > 0$ at the 0.005 level of significance.

- a. What is the number J in each group that would have to be taken so that the probability of a Type II error for the test of the null hypothesis specified in the common section is 0.01 when $E(X) - E(B) = 350$ and $\sigma_X = \sigma_B = 500$?

This part is worth 45 points.

- b. What is the total number of subjects for this clinical trial? This part is worth 5 points.

$$\sqrt{J} \geq \frac{2.576 \sqrt{2} 500 + 2.326 \sqrt{2} 500}{350}$$

$$= \frac{4.902 (1.414) 500}{350} = 9.903$$

$$\sqrt{J} \geq 9.903$$

$$J \geq 98.08$$

$J \geq 99$ IN EACH GROUP

B). $2J = 198$.

- 35 FORGET TO SQUARE $\sqrt{2}$
- 35 FORGET $\sqrt{2}$
- 20 NO 2.576
- 20 NO 2.326

AA

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 $W_2 = 2Y_1 - 2Y_2 - 2Y_3 + 2Y_4$. Find the variance covariance matrix of (W_1, W_2) . (50 points)

$$M = \begin{bmatrix} -3 & -1 & 1 & 3 \\ 2 & -2 & -2 & 2 \end{bmatrix}$$

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

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

$$= \sigma^2 \begin{bmatrix} 20-20\rho & 0 \\ 0 & 16-16\rho \end{bmatrix}$$

+15 CORRECT M.

-20 WRONG VAR W_1

-20 WRONG VAR W_2

-20 WRONG COV(W_1, W_2); -10 ONE COV = 0, OTHER COV $\neq 0$.

-10 NO σ^2 FACTOR.

[illegible]

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

$W_1 = -6Y_1 - 2Y_2 + 2Y_3 + 6Y_4$, and let $W_2 = Y_1 - Y_2 - Y_3 + Y_4$. Find the variance covariance matrix of (W_1, W_2) . (50 points)

$$M = \begin{bmatrix} -6 & -2 & 2 & 6 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

$$VCV(W) = VCV(MY) = \frac{1}{\sigma^2} \begin{bmatrix} -6 & -2 & 2 & 6 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & P & P & P \\ P & 1 & P & P \\ P & P & 1 & P \\ P & P & P & 1 \end{bmatrix} M^T$$

$$= 6^2 \begin{bmatrix} -6+6p & -2+2p & 2-2p & 6-6p \\ 1-p & -1+p & -1+p & 1-p \end{bmatrix} \begin{bmatrix} -6 & 1 \\ -2 & -1 \\ 2 & -1 \\ 6 & 1 \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 80 - 80\rho & 0 \\ 0 & 4 - 4\rho \end{bmatrix}$$

+15 CORRECT M

-20 WRONG VAR W_1

-20 WRONG VAR W2

-20 WRONG cov (w_1, w_2)

-20
-10
0 MIT σ^2 FACTOR

-10

AA

6. Let $(x_i, y_i), i = 1, \dots, n$ be the n observations used to fit the linear regression of y on

x , and let $\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x}_n)(y_i - \bar{y}_n)}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}$ be the usual ordinary least squares estimate of

the slope of the regression line, where \bar{x}_n and \bar{y}_n are the sample means of the x and

y values respectively. Evaluate $\sum_{i=1}^n [(y_i - \bar{y}_n)(x_i - \bar{x}_n) - \hat{\beta}_1(x_i - \bar{x}_n)^2]$. This problem

is worth 60 points.

End of Examination

$$\begin{aligned} & \sum [(y_i - \bar{y}_n)(x_i - \bar{x}_n) - \hat{\beta}_1(x_i - \bar{x}_n)^2] \\ &= \left[n \sqrt{\sum (x_i - \bar{x}_n)^2 \sum (y_i - \bar{y}_n)^2} \right] - n \frac{\sqrt{\sum (y_i - \bar{y}_n)^2}}{\sqrt{\sum (x_i - \bar{x}_n)^2}} \sum (x_i - \bar{x}_n)^2 \end{aligned}$$

$$= 0$$

+10 EACH CORRECT STEP

+10 FOR 0 ANSWER WITH INCORRECT PROOF

63 SAME QUESTION