AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AMS 315, Examination 1
October 18, 2018

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 250 points. There are 6 problems, and the value of each problem is 40 points except for problem 4, which is worth 50 points. 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 6:00 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.

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1. A research team conducted a pilot study of whether a training program improved the productivity of a firm's workers. They took a random sample of $n = 6$ workers and measured their productivity before and after the training program. The change in productivity was normally distributed with unknown mean and unknown variance. The average productivity increase was $\bar{y}_6 = 94.3$ units, and the variance of the productivity increase was 144.8 (the divisor in the variance was $n-1$). What is the 99 % confidence interval for the expected increase? This problem is worth 40 points.

99% CI FOR $E(Y)$:

$94.3 \pm 4.032\sqrt{\frac{144.8}{6}} = 94.3 \pm 4.032\sqrt{24.13}$

$= 94.3 \pm 4.032(4.913) = 94.3 \pm 19.81.$

$=$ 74.49 TO 114.11

−30 2.576 FOR 4.032

−10 INCORRECT T VALUE

−5 (−2) ARITHMETIC ERROR AFTER CORRECT FORMULA

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

1. A research team conducted a pilot study of whether a training program improved the productivity of a firm's workers. They took a random sample of $n = 4$ workers and measured their productivity before and after the training program. The change in productivity was normally distributed with unknown mean and unknown variance. The average productivity increase was $\bar{y}_4 = 20.7$ units, and the sample variance of the productivity increase was 148.3 (the divisor in the variance was $n-1$). Test the null hypothesis that $E(Y) = 0$ against the alternative that $E(Y) \neq 0$. Use levels of significance 0.10, 0.05, and 0.01. This problem is worth 40 points.

$$t_3 = \frac{20.7 - 0}{\sqrt{\frac{148.3}{4}}} = \frac{20.7}{\sqrt{37.075}} = \frac{20.7}{6.089}$$

$$= 3.400$$

α	Z_α	t_3	
.10	1.645	2.353	R
.05	1.960	3.182	R
.01	2.576	5.841	A.

REJECT H_0: $E(Y)=0$ VS H_1: $E(Y) \neq 0$ AT THE 0.05 LEVEL OF SIGNIFICANCE. ACCEPT H_0 AT THE 0.01 LEVEL.

-37 WRONG (INCONSISTENT) DECISION

-30 USE z BUT NOT t CRITICAL VALUES.

-10 WRONG T-VALUES

-20 ONE-SIDED TEST.

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

2. A research team took a random sample of 5 observations from a normally distributed random variable Y and observed that $\bar{y}_5 = 294.9$ and $s_Y^2 = 39.7$, where $\bar{y}_5$ was the average of the five observations sampled from Y and s_Y^2 was the unbiased estimate of $\text{var}(Y)$ (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 $\bar{x}_4 = 231.2$ and $s_X^2 = 35.1$, 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 $n-1$). Calculate the 95% confidence interval for $E(X) - E(Y)$ using the pooled variance estimator. This problem is worth 40 points.

$$S_p^2 = \frac{4(39.7) + 3(35.1)}{7} = \frac{158.8 + 105.3}{7}$$

$$= \frac{264.1}{7} = 37.73$$

$$t_{1.960,7} = 2.365$$ -30 FOR 1.960

95% CI FOR $E(X) - E(Y)$

$$231.2 - 294.9 \pm 2.365\sqrt{37.73}\sqrt{\frac{1}{5}+\frac{1}{4}}$$

$$= -63.7 \pm 2.365(6.142)(0.6708)$$

$$= -63.7 \pm 9.744 = -73.44 \text{ TO } -53.956$$

-5 ~~-10~~ FOR CENTERING ON 63.7

-30 NO $\sqrt{\frac{1}{n}+\frac{1}{m}}$

-30 1.960 FOR 2.365

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA •

2. A research team took a random sample of 2 observations from a normally distributed random variable Y and observed that $\bar{y}_2 = 105.4$ and $s_Y^2 = 220.7$, where $\bar{y}_2$ was the average of the two 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 3 observations from a normally distributed random variable X and observed that $\bar{x}_3 = 134.2$ and $s_X^2 = 230.8$, where $\bar{x}_3$ was the average of the three observations sampled from X and s_X^2 was the unbiased estimate of $\text{var}(X)$ (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.

$$S_p^2 = \frac{1(220.7) + 2(230.8)}{3} = \frac{682.3}{3} = 227.433$$

$$S_p = 15.08$$

$$t_3 = \frac{134.2 - 105.4 - 0}{15.08\sqrt{\frac{1}{2}+\frac{1}{3}}} = \frac{28.8}{15.08\sqrt{0.83333}}$$

$$= \frac{28.8}{13.77} = 2.092$$

α	Z_α	$t_{\alpha,3}$	
.10	1.645	2.353	A
.05	1.960	3.182	A
.01	2.576	5.841	A.

ACCEPT $H_0 : E(X) = E(Y)$ vs $H_1 : E(X) \neq E(Y)$

AT THE 0.10 LEVEL OF SIGNIFICANCE

−30 FOR 1.645, 1.960, 2.576 FOR 2.353, 3.182, 5.841.

−10 T VALUES OTHER THAN 2.353, ...

−37 INCONSISTENT OR NO DECISION

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

3. 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 $\bar{y}_6 = 267.1$, where $\bar{y}_6$ was the average of the six sampled observations and $s^2 = 643.7$ 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 99% confidence interval for σ^2. This problem is worth 40 points.

DF = 5.

$P(.4117 < \chi^2_5 < 16.75) = 0.99$

$P\left(.4117 < \frac{5s^2}{\sigma^2} < 16.75\right) = 0.99$

$P\left(\frac{1}{16.75} < \frac{\sigma^2}{5s^2} < \frac{1}{0.4117}\right) = 0.99.$

$\frac{5s^2}{16.75} < \sigma^2 < \frac{5s^2}{.4117}$ IS THE 99% CI FOR σ^2

$0.2985\, s^2$ TO $12.14\, s^2$

THE 99% CI FOR σ^2 IS

192.15 TO 7817.59

−10 FOR 6(643.7) INSTEAD OF 5(643.7).

−10 NO .4117

−10 NO 16.75

−40 ONE SAMPLE T-TEST

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

3. A research team took a random sample of 4 observations from a normally distributed random variable Y and observed that $\bar{y}_4 = 331.8$ and $s_Y^2 = 1740.8$, where $\bar{y}_4$ was the average of the four observations sampled from Y and s_Y^2 was the unbiased estimate of $\mathrm{var}(Y)$ (i.e., the divisor in the variance was $n-1$). A second research team took a random sample of 6 observations from a normally distributed random variable X and observed that $\bar{x}_6 = 982.5$ and $s_X^2 = 435.2$, where $\bar{x}_6$ was the average of the six observations sampled from X and s_X^2 was the unbiased estimate of $\mathrm{var}(X)$ (i.e., the divisor in the variance was $n-1$). Test the null hypothesis $H_0 : \mathrm{var}(Y) = \mathrm{var}(X)$ against the alternative $H_1 : \mathrm{var}(Y) > \mathrm{var}(X)$ at the 0.10, 0.05, and 0.01 levels of significance. This problem is worth 40 points.

$$TS = \frac{S_Y^2}{S_X^2} = \frac{1740.8}{435.2} = 4.00$$

α	$F_{3,5}$	
.10	3.62	R
.05	5.41	A
.01	12.06	A.

−40 TWO INDEPENDENT SAMPLE T TEST

−37 INCONSISTENT DECISION

−15 FOR EACH INCORRECT DF.

−20 WRONG SIDE (I.E. TS = $\widehat{VAR}X / \widehat{VAR}Y$).

REJECT H_0: VAR(Y) = VAR(X) VS H_1: $\widehat{VAR}(Y) > \widehat{VAR}(X)$ AT THE 0.10 LEVEL OF SIGNIFICANCE. ACCEPT H_0 AT THE 0.05 LEVEL

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

4. A research team collected data on $n = 253$ participants, who were between 8 and 10 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 per week reported was 19.6 hours, with an observed standard deviation of 4.61 hours (the divisor in the underlying variance calculation was $n-1$). Each participant also took a test of cognition. The average cognition score was 94.3, with an observed standard deviation of 12.7 (the divisor in the underlying variance calculation was $n-1$). The Pearson product moment correlation coefficient between the two variables was -0.24. The research team seeks to estimate the regression of cognition score on the average hours per week spent watching a screen.

a. Find the estimated regression equation of the cognition score on the hours spent watching a screen. Find the 99% confidence interval for the slope in this equation. (15 points).

b. Complete the analysis of variance table for this regression and test the null hypothesis that the slope is zero at levels of significance 0.10, 0.05, and 0.01. (15 points)

c. Use the least-squares prediction equation to estimate the expected cognition score for participants whose average screen watching time was 30 hours per week. Give the 99% confidence interval for the expected cognition score for the participants who spent 30 hours per week on average watching screens. (20 points)

$TSS = (252)(12.7)^2 = 40645.8$

$\Sigma(x_i - \bar{x}_m)^2 = 252(4.61)^2 = 5355.5292$

$REG\ SS = (-0.24)^2\ TSS = 2341.157$

$SSE = (1 - (-0.24)^2)\ TSS = 38303.923$

$MSE = SSE/(251) = 152.605 = (12.353)^2$

a $\hat{Y}(x) = 94.3 + (-0.24)\left(\frac{12.7}{4.61}\right)(x - 19.6)$

$= 94.3 - 0.6612(x - 19.6)$

$= 107.259 - 0.6612x$

99% CI FOR $\beta_1 = -0.6612 \pm 2.596\sqrt{\frac{152.605}{5355.5292}}$

$-0.6612 \pm 2.596(0.1688) = -1.099$ TO -0.2230

-0.6612 ± 0.4382

A +5 CORRECT OR CONSISTENT $\hat{\beta}_1$

-10 INCORRECT CI

315 F2018 4B CONT

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

b.

SOURCE	DF	SS	MS	F
REGRESSION	1	2341.157	2341.157	15.34
ERROR	251	38303.923	152.605	
TOTAL	252	40645.8		

α	F(1,251).		F(1,∞).
.10	2.726	R	2.71
.05	3.879.	R	3.84
.01	$6.737 = (2.596)^2$	R	6.63

REJECT $H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$ AT $\alpha = .01$ LEVEL.

c. $\hat{Y}(30) = 94.3 - 0.6612(30 - 19.6) = 94.3 - 0.6612(10.4)$

$= 87.42.$

$se(\hat{Y}(30)) = \sqrt{MSE\left(\frac{1}{253} + \frac{(10.4)^2}{5355.5292}\right)}$

$= \sqrt{152.605(0.003953 + 0.0202)}$

$= \sqrt{3.685} = 1.920.$

99% CI FOR $\beta_0 + 30\beta_1$: $87.42 \pm 2.596(1.920)$

$87.42 \pm 4.983 = 82.44$ TO 92.40

B -10 EACH INCORRECT ANOVA ENTRY.

-12 INCONSISTENT OR NO DECISION

C -20 USE PI FORMULA.

-5 COMP ERROR FOR $\hat{Y}(x)$

-10 FORMULA ERRORS

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

4. A research team collected data on $n = 178$ participants, who were between 8 and 10 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 15.4 hours, with an observed standard deviation of 3.98 hours (the divisor in the underlying variance calculation was $n-1$). Each participant also took a test of cognition. The average cognition score was 106.2, with an observed standard deviation of 16.4 (the divisor in the underlying variance calculation was $n-1$). The Pearson product moment correlation coefficient between the two variables was -0.28. The research team seeks to estimate the regression of cognition score on hours spent watching a screen.
   a. Find the estimated regression equation of cognition score on average time spent watching a screen. Find the 99% confidence interval for the slope in this equation. (15 points).
   b. Complete the analysis of variance table for this regression and test the null hypothesis that the slope is zero at levels of significance 0.10, 0.05, and 0.01. (15 points)
   c. Use the least-squares prediction equation to estimate the cognition score for a participant whose average time spent watching a screen was 22 hours. Give the 99% prediction interval for the cognition score for this participant (whose average time spent watching a screen was 22 hours). (20 points)

$TSS = 177(16.4)^2 = 47605.92$

$\Sigma(x_i - \bar{x}_n)^2 = 177(3.98)^2 = 2803.751$

$\hat{\beta}_1 = r \frac{s_{DV}}{s_{IV}} = -0.28\left(\frac{16.4}{3.98}\right) = -1.1538$

$REG\ SS = (-0.28)^2\ TSS = 3732.304$

$SSE = (1-(-0.28)^2)TSS = 43873.616$ ON 176 DF

$MSE = \frac{SSE}{176} = 249.282 = (15.79)^2$

a) 99% CI FOR β_1: $-1.1538 \pm 2.604\sqrt{\frac{249.282}{2803.751}}$

$-1.1538 \pm 2.604\sqrt{.0889} = -1.1538 \pm 0.7765$

$= -1.9302$ TO -0.3773

A) +5 CORRECT OR CONSISTENT $\hat{\beta}_1$

-10 INCORRECT CI.

AMS 315 F 2018 XI A4 CONTINUED
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA.

b. ANOVA TABLE

SOURCE	DF	SS	MS	F
REG	1	3732.304	3732.304	14.97
ERROR	176	43873.616	249.282	
TOTAL	177	47605.92		

α	$F(1,176)$	$F(1,\infty)$	
.10	2.734	2.71	R
.05	3.895	3.84	R
.01	6.781	6.63	R.

REJECT $H_0: \beta_1 = 0$ VS $H_1 \beta_1 \neq 0$ AT 0.01 LEVEL OF SIGNIFICANCE

c) $\hat{Y}(x) = 106.2 - 1.1538(x - 15.4)$

$= 106.2 + 1.1538(15.4) - 1.1538x$

$= 124.0 - 1.1538x$

$\hat{Y}(22) = 98.6$

$MSE\left(1 + \frac{1}{178} + \frac{(6.6)^2}{2803.751}\right) = MSE(1 + .005618 + 0.015536)$

$= MSE(1.02121) = (249.282)(1.02121) = 254.555$

$= (15.95)^2$

99% PI FOR $Y_F(22)$ IS: $98.6 \pm 2.604(15.95)$

$= 98.6 \pm 41.55 = 57.05$ TO 140.15

B) −10 EACH INCORRECT ANOVA ENTRY.

−12 INCONSISTENT OR NO DECISION.

C −20 USE CI FORMULA.

−5 COMP ERROR FOR $\hat{Y}(x)$

−10 FORMULA ERRORS

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

5. 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 $\rho = 0.20$ (that is, $\rho^2 = 0.04$). What is this value? (This question is worth 40 points).

$F(0.20) = \frac{1}{2} \ln\left(\frac{1.2}{0.8}\right) = .2027$

$$\sqrt{n-3} \geq \frac{1.645(1) + 1.960(1)}{.2027} = 17.782$$

$n-3 \geq 316.20$

$n \geq 320$

+10 CORRECT $F(0.20)$

−15 NO 1.645

−15 NO 1.960

−30 MAJOR COMP ERROR

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

5. 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 $\sigma_X = \sigma_B = 800$. The null hypothesis to be tested is that $E(X) - E(B) = 0$ against the alternative that $E(X) - E(B) > 0$ at the 0.025 level of significance. 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 0.05 when $E(X) - E(B) = 300$ and $\sigma_X = \sigma_B = 800$? What is the total number of subjects for this clinical trial? This problem is worth 40 points.

$$\sqrt{J} \geq \frac{|1.9601|\sqrt{2}(800) + |1.645|\sqrt{2}(800)}{|300 - 0|}$$

$$= \frac{(3.605)\sqrt{2}\,800}{300} = 13.595$$

$J \geq 185$ IN EACH GROUP.

THE STUDY NEEDS 370 PATIENTS.

−10 NO 2J

−15 NO 1.960

−15 NO 1.645.

−20 NO $\sqrt{2}$

−30 MAJOR COMP ERROR.

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

6. The random variables Y_i $i = 1, \cdots, n$ are independent and normally distributed; $E(Y_i) = \beta x_i$ with x_i a fixed and known constant; and $var(Y_i) = \sigma^2 > 0$. The random variable W is the linear function $W = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$.
   a. What is $E(W)$? Prove your answer (10 points).
   b. What is $var(W)$? Prove your answer (30 points).

End of Examination

a) $E(W) = E\left(\frac{\sum x_i Y_i}{\sum x_i^2}\right) = \frac{\sum x_i (E Y_i)}{\sum x_i^2} = \frac{\sum x_i (\beta x_i)}{\sum x_i^2} = \beta.$

b) $W = MY$ WHERE $M = \left[\frac{x_1}{\sum x_i^2} \;\; \frac{x_2}{\sum x_i^2} \;\; \cdots \;\; \frac{x_n}{\sum x_i^2}\right]$, $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

$var\, W = (M \, var\, Y (M^T) = M(\sigma^2 I_{n \times n}) M^T$

$= \sigma^2 M M^T = \sigma^2 \frac{\sum x_i^2}{(\sum x_i^2)^2} = \frac{\sigma^2}{(\sum x_i^2)}$

NOTES: $\frac{\sum x_i Y_i}{\sum x_i^2} \neq \frac{\sum Y_i}{\sum x_i}.$