AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Common Information for Questions 1 and 2

A research team sought to estimate the model $E(Y) = \beta_0 + \beta_1 x + \beta_2 w$. The variable Y was the measure of the task proficiency of an employee observed 2 years after initial training; the variable w was the measure of the employee's achievement level with supplemental training given 1 year after initial training; and the variable x was the measure of the task proficiency of the employee observed immediately after training. They observed values of y, x, w on $n = 312$ employees. They found that the standard deviation of x, where the variance estimator used division by $n - 1$ was 55.6; the standard deviation of w was 33.1; and the standard deviation of Y was 41.7. The correlation between Y and x was 0.58; the correlation between Y and w was 0.69; and the correlation between x and w was 0.82. (Total 100 points)

1. Compute the partial correlation coefficient $r_{Yx \cdot w}$ and the partial correlation coefficient $r_{Yw \cdot x}$. Is a mediation model or an explanation model a better explanation of the observed results? (50 points)
2. Compute the analysis of variance table for the multiple regression analysis of Y. Include the sum of squares due to the regression on x, the sum of squares due to the regression on w after including x, the sum of squared errors, the total sum of squares, and each degree of freedom. What is your decision for the test of the null hypothesis that $\beta_2 = 0$? Use levels of significance 0.10, 0.05, and 0.01. (50 points)

End of application of common information

1A. $$r_{Yx \cdot w} = \frac{r_{Yx} - r_{Yw} r_{xw}}{\sqrt{(1 - r_{Yw}^2)(1 - r_{xw}^2)}} = \frac{0.58 - (0.69)(0.82)}{\sqrt{(1 - .69^2)(1 - .82^2)}}$$

$$= \frac{.0142}{\sqrt{(.5239)(.3276)}} = \frac{.0142}{\sqrt{.17163}} = 0.03427$$

$$r_{Yw \cdot x} = \frac{0.69 - (0.58)(0.82)}{\sqrt{(1 - .58^2)(1 - .82^2)}} = \frac{0.2144}{\sqrt{(.6636)(.3276)}}$$

$$= \frac{0.2144}{\sqrt{.21740}} = 0.45983.$$

SINCE $se(r_{Yx \cdot w}) \cong \frac{1}{\sqrt{n-3}} = 0.057$, $r_{Yx \cdot w} \cong 0$. KEY VARIABLE IS w

$$\underset{H}{x} \to \underset{H+1}{w} \to \underset{H+2}{Y}$$

MEDIATION MODEL. DESCRIBES THIS DATA.

AMS 315 F2020 EXAMINATION 3 SOLUTION

2A. $TSS = (n-1)(SD_{DV}^2) = 311(41.7)^2 = 540,794.79.$

$SS\,REG(x) = (.58)^2 TSS = 181,923.37$

$TSS - SSREG(x) = 358,871.42$

$SS\,REG(w|x) = (r_{YW\cdot X})^2\,(TSS - SS\,REG(x))$

$= (0.45983)^2\,(358,871.42)$

$= 75,881.08$

$SSE = TSS - SS\,REG(x) - SS\,REG(w|x)$

$= 282,990.34.$

ANOVA TABLE

SOURCE	DF	SS	MS
REG (x)	1	181,923.37	
REG (w|x)	1	75,881.08	75,881.08
ERROR	309	282,990.34	915.83
TOTAL	311	540,794.79	

$F_{w|x} = \dfrac{MS\,REG(w|x)}{MSE} = \dfrac{75,881.08}{915.83} = 82.86.$

α	$F(1,309)$	$F(1,\infty)$	
.10	2.722	2.71	REJECT
.05	3.872	3.84	REJECT
.01	6.718	6.64	REJECT

REJECT $H_0: \beta_2 = 0$ VS $H_1\ \beta_2 \neq 0$ AT $\alpha = .01$ (AND .05 AND .10).

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

Common Information for Questions 1 and 2

A research team sought to estimate the model $E(Y) = \beta_0 + \beta_1 x + \beta_2 w$, using data from $n = 471$ participants. The variable Y was the measure of the participant's criminality at age 20 (a higher number reflecting more criminal behavior); the variable w was the participant's education attainment at age 18 (higher number reflecting more attainment); and the variable x was the participant's delinquency at age 15 (higher number reflecting more delinquency). They found that the standard deviation of Y, where the variance estimator used division by $n - 1$, was 34.9; the standard deviation of x was 15.7; and the standard deviation of w was 41.6. The correlation between Y and w was -0.21; the correlation between Y and x was 0.61; and the correlation between x and w was -0.32.

1. Compute the partial correlation coefficient $r_{Yx \cdot w}$ and the partial correlation coefficient $r_{Yw \cdot x}$. Is a mediation model or an explanation model a better explanation of the observed results? (50 points)
2. Compute the analysis of variance table for the multiple regression analysis of Y. Include the sum of squares due to the regression on x, the sum of squares due to the regression on w after including x, the sum of squared errors, the total sum of squares, and each degree of freedom. What is your decision for the test of the null hypothesis that $\beta_2 = 0$? Use levels of significance 0.10, 0.05, and 0.01. (50 points)

End of application of common information

1B $$r_{Yx \cdot w} = \frac{r_{Yx} - r_{Yw} r_{xw}}{\sqrt{(1-r_{Yw}^2)(1-r_{xw}^2)}} = \frac{0.61-(-0.21)(-0.32)}{\sqrt{(1-(-.21)^2)(1-(-.32)^2)}}$$

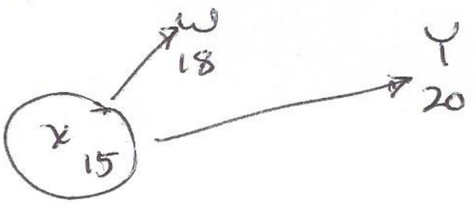
$$= \frac{.5428}{\sqrt{(.9559)(.8976)}} = \frac{.5428}{\sqrt{.85802}} = .58600$$

$$r_{Yw \cdot x} = \frac{-.21-(0.61)(-0.32)}{\sqrt{(1-.61^2)(1-(-.32)^2)}} = \frac{-0.0148}{\sqrt{(.6279)(.8976)}}$$

$$= \frac{-0.0148}{\sqrt{.56360}} = \frac{-0.0148}{0.75073} = -0.01971$$

SINCE $se(r_{Yw.x}) \approx \frac{1}{\sqrt{468}} = .046$, $r_{Yw.x} \approx 0$, AND

x IS THE KEY VARIABLE.



THIS IS AN EXPLANATION MODEL.

AMS 315 F 2020 EXAMINATION 3 SOLUTION

2B. $TSS = (n-1) SD_{DV}^2 = 470 (34.9)^2 = 572,464.7$

$SSREG(x) = r_{Yx}^2 \; TSS = (.61)^2 \, TSS = 213,014.1$

$TSS - SSREG(x) = 359,450.6$

$SSREG(w|x) = r_{Yw \cdot x}^2 (TSS - SSREG(x))$

$= (-0.01971)^2 (TSS - SSREG(x))$

$= 139.6$

$SSERR = TSS - SSREG(x) - SSREG(w|x)$

$= 359,311.0$

ANOVA TABLE

SOURCE	DF	SS	MS
REG (x)	1	213,014.1	
REG(w|x)	1	139.6	139.6
ERROR	468	359,311.0	767.8
TOTAL	470	572,464.7	

$$F_{w|x} = \frac{139.6}{767.8} = 0.18$$

α	$F(1, 468)$	
.10	2.716	ACCEPT
.05	3.861	ACCEPT
.01	6.689	ACCEPT

ACCEPT $H_0: \beta_2 = 0$ VS $H_1: \beta_2 \neq 0$ AT $\alpha = .10$ (AND $\alpha = .05$ AND $\alpha = .01$). EDUCATIONAL ATTAINMENT AT AGE 18 IS NOT ASSOCIATED WITH CRIMINALITY AT AGE 20 AFTER CONTROLLING FOR DELINQUENCY AT AGE 15.

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

Common Information for Questions 3 and 4

A research team studied Y, the protein production of a laboratory animal, and how Y was affected by the dose of medicine. The research team sought to set the dose of the medicine so that $E(Y)$ is maximized. They used four doses of the medicine: 0, 1, 2, and 3 units respectively. They randomly assigned 10 animals to dosage 0, 10 to dosage 1, 10 to dosage 2, and 10 to dosage 3. They observed that the average values of Y at each dosage were $y_{0\bullet} = 57.2, y_{1\bullet} = 82.4,\ y_{2\bullet} = 41.6,\ and\ y_{3\bullet} = 62.8,$ where $y_{i.}$ was the average of the observations taken with dosage $i = 0,1,2,3$, respectively. They also observed that $s_0^2 = 2{,}582, s_1^2 = 3{,}649,\ s_2^2 = 1{,}877,$ and $s_3^2 = 2{,}264,$ where s_i^2 was the unbiased estimate of the variance for the observations taken with dosage $i = 0,1,2,3$, respectively.

3. Complete the analysis of variance table for these results; that is, be sure to specify the degrees of freedom, sums of squares, mean squares, F-test, and your conclusion. Test the null hypothesis that all treatment means are equal using significance levels 0.10, 0.05, and 0.01. This question is worth 40 points.
4. Find the estimated linear contrast, the sum of squares due to the linear contrast and the 99% Scheffe confidence interval for the linear contrast. The coefficients of the linear contrast are $-3, -1, 1, 3$. What is the optimal setting of dosage, and how do you document it? This question is worth 40 points.

End of Application of Common Information

3C DATA

DOSE	J_i	$y_{i\bullet}$	$y_{i\bullet}-y_{\bullet\bullet}$	s_i^2
0	10	57.2	-3.8	2582
1	10	82.4	21.4	3649
2	10	41.6	-19.4	1877
3	10	62.8	1.8	2264
		$\Sigma = 244$	0	10372

$y_{\bullet\bullet} = 61$

MSE = 2593.

$\Sigma(y_{i\bullet}-y_{\bullet\bullet})^2 = [(-3.8)^2 + (21.4)^2 + (-19.4)^2 + (1.8)^2] = 852.$

$$MSE = \frac{s_1^2 + s_2^2 + s_3^2 + s_4^2}{4} = 2593 \text{ ON } 36\text{DF}$$

$SSE = 93{,}348.$

$SSTREATMENT = J\,\Sigma(y_{i\bullet}-y_{\bullet\bullet})^2 = 10\,(852) = 8{,}520$ ON 3DF

$$MS\ TREATMENT = \frac{SS\ TREATMENT}{I-1} = \frac{8520}{3} = 2840$$

ANOVA TABLE

SOURCE	DF	SS	MS	F
TREATMENT	3	8520	2840	1.095
(PURE) ERROR	36	93,348	2593	
TOTAL	39	101,868		

α	F(3,36)	
.10	2.243	ACCEPT
.05	3.025	ACCEPT
.01	4.377	ACCEPT

AMS 315 F2020 EXAMINATION 3 SOLUTION.

3C CONTINUED

ACCEPT H_0: ALL FOUR DOSE MEANS ARE EQUAL

VS H_1: AT LEAST TWO DOSES HAVE DIFFERENT

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

4C $\hat{\lambda}_{LINEAR} = -3(57.2) - 82.4 + 41.6 + 3(62.8)$

$= -24.$

$$SS_{LINEAR} = \frac{(-24)^2}{[(-3)^2 + (-1)^2 + (1)^2 + 3^2]/10} = 288$$

99% SCHEFFE CI FOR λ_2 IS.

$$-24 \pm \sqrt{3(4.377)}\sqrt{2593\left(\frac{20}{10}\right)}$$

$$-24 \pm (3.624)(72.0)$$

$$-24 \pm 260.98$$

$$= -284.98 \text{ TO } 236.98.$$

THERE IS NO OPTIMAL SETTING!

ALL DOSES APPEAR EQUAL.

DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD

Common Information for Questions 3 and 4

A research team studied Y, the protein production of a laboratory animal, and how Y was affected by the dose of medicine. The research team sought to set the dose of the medicine so that $E(Y)$ is minimized. They used four doses of the medicine: 0, 1, 2, and 3 units respectively. They randomly assigned 16 animals to dosage 0, 16 to dosage 1, 16 to dosage 2, and 16 to dosage 3. They observed that the average values of Y at each dosage were $y_{0\bullet} = 132.8, y_{1\bullet} = 97.4, y_{2\bullet} = 61.6,$ *and* $y_{3\bullet} = 84.2$, where $y_{i.}$ was the average of the observations taken with dosage $i = 0,1,2,3$, respectively. They also observed that $s_0^2 = 1{,}564, s_1^2 = 3{,}256, s_2^2 = 2{,}030,$ and $s_3^2 = 1{,}456$, where s_i^2 was the unbiased estimate of the variance for the observations taken with dosage $i = 0,1,2,3$, respectively.

3. Complete the analysis of variance table for these results; that is, be sure to specify the degrees of freedom, sums of squares, mean squares, F-test, and your conclusion. Test the null hypothesis that all treatment means are equal using significance levels 0.10, 0.05, and 0.01. This question is worth 40 points.
4. Find the estimated quadratic contrast, the sum of squares due to the quadratic contrast and the 99% Scheffe confidence interval for the quadratic contrast. The coefficients of the quadratic contrast are $1, -1, -1, 1$. What is the optimal setting of dosage, and how do you document it? This question is worth 40 points.

End of Application of Common Information

3D DATA.

DOSE	J_i	$y_{i\circ}$	$y_{i\circ} - y_{\circ\circ}$	s_i^2
0	16	132.8	38.8	1,564
1	16	97.4	3.4	3,256
2	16	61.6	-32.4	2,030
3	16	84.2	-9.8	1,456
		$\Sigma = 376.0$	0	8306

$y_{\circ\circ} = 94.0$

$\sum (y_{i\circ} - y_{\circ\circ})^2 = [38.8^2 + 3.4^2 + (-32.4)^2 + (-9.8)^2] = 2,662.8$

$$MSE = \frac{s_0^2 + s_1^2 + s_2^2 + s_3^2}{4} = \frac{8306}{4} = 2076.5.$$

$SSE = 60(MSE) = 124,590$ ON 60 DF

$SSTREATMENT = 16(\sum (y_{i\circ} - y_{\circ\circ})^2) = 42,604.8$

$$MS\ TREATMENT = \frac{SSTREATMENT}{(I-1)} = 14,201.6.$$

ANOVA TABLE

SOURCE	DF	SS	MS	F
TREATMENT	3	42,604.8	14,201.6	6.84.
(PURE) ERROR	60	124,590.0	2,076.5	
TOTAL	63	167,194.8		

AMS 315 F2020 EXAMINATION 3 SOLUTION.

α	$F(3,60)$	DECISION
.10	2.177	REJECT
.05	2.758	REJECT
.01	4.126.	REJECT

REJECT H_0: ALL DOSE MEANS EQUAL VS H_1: AT LEAST TWO DOSE MEANS ARE DIFFERENT AT $\alpha = .01$ (AND $\alpha = .05$ AND $\alpha = .10$).

4.D $\hat{\lambda}_{QUADRATIC} = 1(132.8) - 97.4 - 61.6 + 84.2 = 58.0$

$$SS_{QUADRATIC} = \frac{(\hat{\lambda}_{QUADRATIC})^2}{[1^2 + (-1)^2 + (-1)^2 + 1^2]/16} = 13{,}456.$$

99% SCHEFFE CI FOR $\lambda_{QUADRATIC}$ IS

$$58.0 \pm \sqrt{3(4.126)}\sqrt{2076.5\left(\tfrac{4}{16}\right)}$$

$$58.0 \pm (3.518)\ (22.78).$$

$$58.0 \pm 80.15 = -22.15 \text{ TO } 138.15$$

CI INCLUDES 0.

DOSE WITH LOWEST AVERAGE IS DOSE 2.

99% LSD IS $t_{2.576,60}\sqrt{MSE\left(\frac{2}{J}\right)}$

$$= (2.660)\sqrt{2076.5\left(\tfrac{2}{16}\right)}$$

$$= (2.660)(16.11) = 42.86.$$

NOTE $61.6 + (99\%\ LSD) = 104.46$.

ONLY DOSE 0 APPEARS DIFFERENT FROM DOSE 2.

DOSES 1, 2, AND 3 HAVE NEARLY EQUAL MEANS BY LSD.

EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE

Common Information for Questions 5 and 6

A research team randomly assigned animals to four settings of a dosage of an experimental medicine and observed the response Y. The research team sought to find the dosage that minimized the response variable. Thirty five animals were given one unit of dosage with observed average and sample variance (unbiased estimate) $y_{1\cdot} = 124$ and $s_1^2 = 35{,}040$; 35 were given two units of dosage with $y_{2\cdot} = 156$ and $s_2^2 = 26{,}492$; 35 were given three units of dosage with $y_{3\cdot} = 278$ and $s_3^2 = 36{,}820$; and 35 were given four units of dosage with $y_{4\cdot} = 274$ and $s_4^2 = 24{,}846$. The total sum of squares is 4,854,292. With regard to the orthogonal polynomials, the estimated linear contrast is 572, and its coefficients are $-3,-1,1,3$. The estimated quadratic contrast is ***-36***, and its coefficients are $1,-1,-1,1$. The estimated cubic contrast is ***-216***, and its coefficients are $-1,3,-3,1$.

5. Complete the analysis of variance table for the linear regression of the dependent variable on the dosage level by using the sum of squares for the linear contrast as the regression sum of squares. Test the null hypothesis that the average response is not linearly associated with the dosage given. Use the 0.10, 0.05, and 0.01 levels of significance.
6. Test the null hypothesis that the linear model is adequate at the 0.10, 0.05, and 0.01 levels. Report the analysis of variance table that includes the sum of squares for lack of fit of the linear regression and the sum of squares due to pure error. What is your recommendation for the optimum setting of the dosage?

End of Application of Common Information

5E $$SS\ LINEAR = \frac{(\hat{\lambda}_{LINEAR})^2}{[(-3)^2+(-1)^2+1^2+3^2]/35}$$

$$= \frac{(572)^2}{20/35} = 572{,}572 \text{ . ON 1 DF}$$

ANOVA TABLE

SOURCE	DF	SS	MS	F
LINEAR	1	572,572	572,572	18.45
ERROR	138	4,281,720	31,026.96	
TOTAL	139	4,854,292		

$SSE = SS\ TOTAL - SS\ LINEAR = 4{,}854{,}292 - 572{,}572$
$= 4{,}281{,}720$

$$F_{LINEAR} = \frac{MS\ REG}{MSE} = \frac{572{,}572}{31{,}026.96} = 18.45$$

α	$F(1,138)$	
.10	2.742	REJECT
.05	3.910	REJECT
.01	6.822	REJECT

REJECT H_0: NO LINEAR ASSOCIATION VS H_1: LINEAR ASSOCIATION AT $\alpha = .01$ (AND $\alpha = .05$ AND $\alpha = .01$).

AMS 315 F2020 EXAMINATION 3 SOLUTION.

6E: $SS_{LOF} = SS_{QUAD} + SS_{CUB}$.

$$SS_{QUAD} = \frac{(\hat{\lambda}_{QUAD})^2}{[(1)^2+(-1)^2+(-1)^2+1^2]/35} = \frac{(-36)^2}{4/35} = 11{,}340$$

$$SS_{CUBIC} = \frac{(\hat{\lambda}_{CUBIC})^2}{[(-1)^2+(3)^2+(-3)^2+1^2]/35]}$$

$$= \frac{(-216)^2}{20/35} = 81{,}648$$

$SS_{LOF} = 11{,}340 + 81{,}648 = 92{,}988$ ON 2 DF

$MS_{LOF} = SS_{LOF}/2 = 46{,}494$

$$MS_{PE} = \frac{S_1^2+S_2^2+S_3^2+S_4^2}{4} = \frac{35{,}040+26{,}492+36{,}820+24{,}846}{4}$$

$$= \frac{123{,}198}{4} = 30{,}799.5 \text{ ON 136 DF}$$

$SSPE = 136(MSPE) = 4{,}188{,}732$

ANOVA TABLE LACK OF FIT.

SOURCE	DF	SS	MS
LINEAR	1	572,572	572,572
LACK OF FIT	2	92,988	46,494
PURE ERROR	136	4,188,732	30,799.5
TOTAL	139	4,854,292	

$$F_{LOF} = \frac{MS_{LOF}}{MS_{PE}} = \frac{46{,}494}{30{,}799.5} = 1.51 \text{ ON } (2, 136) \text{ DF.}$$

α	$F(2,136)$	
.10	2.342	ACCEPT
.05	3.063	ACCEPT
.01	4.765	ACCEPT

ACCEPT H_0 LINEAR MODEL ADEQUATE AT $\alpha = .10$

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

DOSAGE 1 HAS SMALLEST MEAN: POSITIVE SLOPE

SUGGESTS DOSAGE OF THIS MEDICINE DOES NOT HELP.

99% LSD = 109.6 DOSE 1 HAS LOWER MEAN THAN DOSE 3 OR DOSE 4.

FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF

Common Information for Questions 5 and 6

A research team randomly assigned animals to four settings of a dosage of an experimental medicine and observed the response *Y*. The research team sought to find the dosage that maximized the response variable. Twenty eight animals were given one unit of dosage with observed average and sample variance (unbiased estimate) $y_{1\cdot} = 220$ and $s_1^2 = 44{,}912$; 28 were given two units of dosage with $y_{2\cdot} = 196$ and $s_2^2 = 51{,}016$; 28 were given three units of dosage with $y_{3\cdot} = 280$ and $s_3^2 = 55{,}384$; and 28 were given four units of dosage with $y_{4\cdot} = 444$ and $s_4^2 = 23{,}408$. The total sum of squares is 5,766,096. With regard to the orthogonal polynomials, the estimated linear contrast is 756, and its coefficients are $-3,-1,1,3$. The estimated quadratic contrast is 188, and its coefficients are $1,-1,-1,1$. The estimated cubic contrast is *-28*, and its coefficients are $-1,3,-3,1$.

5. Complete the analysis of variance table for the linear regression of the dependent variable on the dosage level by using the sum of squares for the linear contrast as the regression sum of squares. Test the null hypothesis that the average response is not linearly associated with the dosage given. Use the 0.10, 0.05, and 0.01 levels of significance. (40 points)
6. Test the null hypothesis that the linear model is adequate at the 0.10, 0.05, and 0.01 levels. Report the analysis of variance table that includes the sum of squares for lack of fit of the linear regression and the sum of squares due to pure error. What is your recommendation for the optimum setting of the dosage? (50 points)

End of Application of Common Information

$$5F.\ SS\ LINEAR = \frac{(\hat{\lambda}_{LINEAR})^2}{[(-3)^2+(-1)^2+(1)^2+3^2]/28} = \frac{(756)^2}{20/28} = 800{,}150.4.$$

SS ERR = SS TOTAL − SS LINEAR

= 5,766,096 − 800150.4. = 4,965,945.6 ON 110 DF.

ANOVA TABLE PROBLEM 5F

SOURCE	DF	SS	MS	F
LINEAR	1	800,150.4	800,150.4.	17.72.
ERROR	110	4,965,945.6	45,144.96	
TOTAL	111	5,766,096		

α	$F(1,110)$.	
.10	2.752	REJECT
.05	3.927	REJECT
.01.	6.871	REJECT

REJECT H_0: NO LINEAR ASSOCIATION VS. H_1: LINEAR ASSOCIATION AT $\alpha=.01$ (AND $\alpha=.05$ AND $\alpha=.01$).

AMS 315 F 2020 EXAMINATION 3 SOLUTION.

$$MSPE = \frac{\Delta_1^2 + \Delta_2^2 + \Delta_3^2 + \Delta_4^2}{4} = \frac{44,912 + 51,016 + 55,384 + 23,408}{4}$$

6F:

$$= \frac{174,720}{4} = 43,680 \text{ ON } 108 \text{ DF}$$

$$SSPE = 4,717,440$$

I $SS_{LOF} = TSS - SS_{LIN} - SSPE$

$$= 5,766,096 - 800,150.4 - 4,717,404 = 248,505.6 \text{ ON 2DF.}$$

II $SS_{LOF} = SS_{QUAD} + SS_{CUB} = 247,408 + 1097.6 = 248,505.6$

$$SS_{QUAD} = 247,408 = \frac{(188)^2}{4/28} \qquad SS_{CUB} = 1097.6 = \frac{(-28)^2}{20/28}$$

ANOVA TABLE LACK OF FIT.

SOURCE	DF	SS	MS
LINEAR	1	800,150.4	
LACK OF FIT	2	248,505.6	124,252.8
PURE ERROR	108	4,717,440	43,680.
TOTAL.	111	5,766,096	

$$F_{LOF} = \frac{MS \text{ LACK OF FIT}}{MS\ PE} = \frac{124,252.8}{43,680} = 2.845 \text{ ON } (2, 108) \text{ DF.}$$

α	$F(2,108)$	
.10	2.352	REJECT
.05	3.080	ACCEPT
.01	4.807	ACCEPT

ACCEPT H_0 LINEAR MODEL ADEQUATE AT $\alpha = .01$ AND $\alpha = .05$. REJECT H_0 LINEAR MODEL ADEQUATE AT $\alpha = .10$.

DOSE 4 HAS MAXIMUM MEAN. SINCE SLOPE IS POSITIVE LARGER DOSES MAY HAVE LARGER MEANS.

99% LSD = 146.5; 444 − 146.5 = 297.5 DOSE 4 HAS GREATER MEAN THAN DOSES 1, 2, 3 USING 99% LSD.

GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG

7. Consider the usual regression model. The random vector Y is $n \times 1$, with $Y = X\beta + \varepsilon$, where β is a $p \times 1$ vector of (unknown) constants, X is an $n \times p$ matrix of known constants with $rank(X) = p$ (so that $(X^T X)^{-1}$exists), ε is an $n \times 1$vector of random variables with $E(\varepsilon) = 0$ and $vcv(\varepsilon) = \sigma^2 I_{n\times n}$ where $I_{n\times n}$ is the $n \times n$ identity matrix. Let $W = [I_{n\times n} - X(X^T X)^{-1} X^T]Y$, where $X^T$is the transpose of X.
   a. Find $E(W)$. (10 points)
   b. Find the variance-covariance matrix of W, $vcv(W)$. (40 points).

End of the Examination

a $E(W) = [I - X(X^T X)^{-1} X^T] EY$

$= [I - X(X^T X)^{-1} X^T] X\beta$.

$= X\beta - X(X^T X)^{-1}(X^T X)\beta$

$= X\beta - X[(X^T X)^{-1}(X^T X)]\beta = 0_{n\times 1}$

b. LET $M = [I - X(X^T X)^{-1} X^T]$

$vcv(W) = vcv(MY) = M\, vcv(Y) M^T$

$= [I - X(X^T X)^{-1} X^T]\sigma^2 I [I - X(X^T X)^{-1} X^T]^T$

$= \sigma^2 [I - X(X^T X)^{-1} X^T][I^T - (X^T)^T ((X^T X)^{-1})^T X^T$

$= \sigma^2 [I - X(X^T X)^{-1} X^T][I - X(X^T X)^{-1} X^T]$

$= \sigma^2 [I \times I - I X(X^T X)^{-1} X^T - X(X^T X)^{-1} X^T I$

$+ X(X^T X)^{-1} X^T X (X^T X)^{-1} X^T]$

$= \sigma^2 [I - 2X(X^T X)^{-1} X^T + X(X^T X)^{-1}[X^T X(X^T X)^{-1}]X^T]$

$= \sigma^2 [I - X(X^T X)^{-1} X^T]$