SOLUTION

AMS 315 M2 SPRING 2020.

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Common Information for Questions 1, 2, and 3

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 35 animals to dosage 0, 35 to dosage 1, 35 to dosage 2, and 35 to dosage 3. They observed that the average values of Y at each dosage were $y_{0\bullet} = 217, y_{1\bullet} = 181,\ y_{2\bullet} = 118,\ and\ y_{3\bullet} = 68$, where $y_{i.}$ was the average of the observations taken with dosage $i = 0,1,2,3$, respectively. They also observed that $s_0^2 = 26{,}425, s_1^2 = 19{,}564,\ s_2^2 = 31{,}950$, and $s_3^2 = 22{,}917$, where s_i^2 was the unbiased estimate of the variance for the observations taken with dosage $i = 0,1,2,3$, respectively.

1. 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 50 points.

2. 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 $-3, -1, 1, 3$. This question is worth 30 points.

3. What is the optimal setting of dosage, and how do you document it? This question is worth 10 points.

End of Application of Common Information

DATA

1A.

i	DOSE	J_i	$y_{i\cdot}$	$\hat{\alpha}_i = y_{i\cdot} - y_{\cdot\cdot}$	s_i^2
1	0	35	217	71	26,425
2	1	35	181	35	19,564
3	2	35	118	-28	31,950
4	3	35	68	-78	22,917
			584.	0	100,856
			$y_{\cdot\cdot} = 146$		25,214. = MSE

$\sum(\hat{\alpha}_i)^2 = (71)^2 + (35)^2 + (-28)^2 + (-78)^2 = 13134$

$SSTR = 35 \sum(\hat{\alpha}_i)^2 = 35[13,134] = 459,690$

$MSTR = 153,230$

$SSE = 136(MSE) = 3,429,104.$

ANOVA TABLE.

SOURCE	DF	SS	MS	F
TREATMENTS	3	459,690	153,230	6.077.
ERROR	136	3,429,104	25,214	
TOTAL.	139	3,888,794.		

α	$F(3,136)$.	
.10	2.124	REJECT
.05	2.671	REJECT
.01	3.929	REJECT

REJECT $H_0: E(Y_{1j}) = \ldots = E(Y_{4j})$ AGAINST $H_1: E(Y_{ij}) \neq E(Y_{i'j})$ AT $\alpha = .01$.

2A $\hat{\lambda}_{LIN} = -3(217) - 1(181) + 1(118) + 3(68) = -510$

$SS_{LIN} = \dfrac{(\hat{\lambda}_{LIN})^2}{20/35} = 455,175.0$

99% SCHEFFE CI FOR λ_{LIN}: $-510 \pm \sqrt{3(3.929)}\sqrt{25,214\left(\frac{20}{35}\right)}$

$= -510 \pm \sqrt{11.787}\sqrt{14,408} = -510 \pm 3.433(120.03)$

$= -510 \pm 412.1 = -922.1$ TO -97.9

3A. DOSE 3 OR HIGHER MINIMIZES $E(Y_{ij})$ AS SHOWN BY THE NEGATIVE $\hat{\lambda}_{LIN}$ WHICH IS SIGNIFICANT AT $\alpha = .01$ BY SCHEFFE CONFIDENCE INTERVAL FOR λ_L.

BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

Common Information for Questions 1, 2, and 3

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 $y_{0\bullet} = 92, y_{1\bullet} = 116,\ y_{2\bullet} = 110,\ and\ y_{3\bullet} = 86,$ where $y_{i.}$ was the average of the observations taken with dosage $i = 0,1,2,3,$ respectively. They also observed that $s_0^2 = 1{,}130, s_1^2 = 1{,}225,\ s_2^2 = 958,$ and $s_3^2 = 1{,}419,$ where s_i^2 was the unbiased estimate of the variance for the observations taken with dosage $i = 0,1,2,3,$ respectively.

1. 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 50 points.

2. 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 $1, -1, -1, 1$. This question is worth 30 points.

3. What is the optimal setting of dosage, and how do you document it? This question is worth 10 points.

End of Application of Common Information

DATA

1B

DOSE	i	J_i	$y_{i\cdot}$	$\hat{\alpha}_i = y_{i\cdot} - y_{\cdot\cdot}$	Δ_i^2
0	1	10	92	−9	1130
1	2	10	116	15	1225
2	3	10	110	9	958
3	4	10	86	−15	1419
		40	404	0	4732

$y_{\cdot\cdot} = 101.$ MSE = 1183

SSE = 42,588 = 36(1183) ON 36 DF.

$\Sigma(\hat{\alpha}_i)^2 = 612$ SS TREATMENT = 10(612) = 6,120 ON 3 DF.

MSTREATMENT = 2040

ANOVA TABLE

SOURCE	DF	SS	MS	F
TREATMENTS	3	6,120	2040	1.724.
ERROR	36	42,588	1183	
TOTAL	39	48,708		

α	F(3,36).	
.10	2.243	ACCEPT
.05	2.866	ACCEPT
.01	4.377	ACCEPT

ACCEPT $H_0: E(Y_{1j}) = \cdots = E(Y_{4j})$ VS $H_1: E(Y_{ij}) \neq E(Y_{i'j})$ AT $\alpha = .10$ LEVEL.

2B. $\hat{\lambda}_Q = 92 - 116 - 110 + 86 = -48$

$$SS_Q = \frac{(\hat{\lambda}_Q)^2}{4/10} = 5760.0$$

99% SCHEFFE CI FOR λ_Q IS $-48 \pm \sqrt{3(4.377)}\sqrt{1183(\frac{4}{10})}$

$= -48 \pm \sqrt{13.131}\sqrt{473.2} = -48 \pm 3.624(21.75)$

$= -48 \pm 78.8 = -126.8$ TO 30.8

3B. THERE IS NO OPTIMAL DOSE AS WE ACCEPT THE OVERALL NULL HYPOTHESIS AT $\alpha = .10$

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

Common Information for Problems 4, 5, and 6

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 minimize $E(Y)$. They used four doses of 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 $y_{0\bullet} = 388, y_{1\bullet} = 324,\ y_{2\bullet} = 222,$ and $y_{3\bullet} = 174$, where $y_{i.}$ was the average of the observations taken with dosage $i = 0,1,2,3$, respectively. They also observed that $s_0^2 = 25{,}040.5, s_1^2 = 19{,}918.3,\ s_2^2 = 31{,}125.6,$ and $s_3^2 = 24{,}815.6,$ where s_i^2 was the unbiased estimate of the variance for the observations taken with dosage $i = 0,1,2,3$, respectively. They correctly calculated that $y_{\bullet\bullet} = 277$ and that the average s_i^2 was 25,225. The total sum of squares was 1,964,124.0. They also correctly calculated the values of the three contrasts. The coefficients of the linear contrast were $-3,-1,1,3$;, and $\hat{\lambda}_{Lin} = -744$. The coefficients of the quadratic contrast were $1,-1,-1,1$; and $\hat{\lambda}_{Quad} = 16$. The coefficients of the cubic contrast were $-1,3,-3,1$; and $\hat{\lambda}_{Cubic} = 92$.

4. What are the values of the sum of squares due to the linear contrast, the sum of squares dues to the quadratic contrast, and the sum of squares due to the cubic contrast? This question is worth 20 points.

5. Report the analysis of variance table for the linear regression of Y on dosage. Use the sum of squares due to the linear contrast as the sum of squares for the regression of Y on dosage. Test the null hypothesis that there is no linear association at the 0.10, 0.05, and 0.01 levels of significance. What is the optimal setting of the dosage? This question is worth 40 points.

6. Test the null hypothesis that the linear model is adequate at the 0.10, 0.05, and 0.01 levels of significance. Report the analysis of variance table including the sum of squares due to lack of fit. This question is worth 50 points.

End of application of common information

4C $\hat{\lambda}_{LIN} = -744.$ $SS_{LIN} = \frac{(-744)^2}{20/16} = 442,828.8.$

$\hat{\lambda}_{QUAD} = 16$ $SS_{QUAD} = \frac{(16)^2}{4/16} = 1,024.0$

$\hat{\lambda}_{CUBIC} = 92$ $SS_{CUBIC} = \frac{(92)^2}{20/16} = 6,771.2$

NOTE $SS_{LIN} + SS_{QUAD} + SS_{CUBIC} = 450,624.0 = SS_{TREATMENT}$

5C ANOVA TABLE, CHI LINEAR REGRESSION.

SOURCE	DF	SS	MS	F
LINEAR	1	442,828.8	442,828.8	18.05
ERROR	62	1,521,295.2	24,537.0	
TOTAL	63	1,964,124.0		

α	$F(\alpha,1,62)$	
.10	2.788	REJECT
.05	3.996	REJECT
.01	7.062	REJECT.

REJECT H_0: NO LINEAR ASSOCIATION VS H_1: LINEAR ASSOCIATION AT $\alpha = .01$.

OPTIMAL (MINIMIZING) DOSAGE IS 3 UNITS OR HIGHER AS DOCUMENTED BY HIGHLY SIGNIFICANT AND NEGATIVE $\hat{\lambda}_{LIN}$.

SSE CAN BE CALCULATED BY SUBTRACTION OR $SSPE + SS_{QUAD} + SS_{CUBIC}$.

6C. ANOVA TABLE LACK OF FIT OF LINEAR MODEL.

SOURCE	DF	SS	MS
LINEAR	1	442,828.8	442,828.8
LACK OF FIT	2	7,795.2	3,897.6
PURE ERROR	60	1,513,500	25,225.0
TOTAL	63	1,964,124.0	

$F_{LOF} = \frac{MS_{LOF}}{MS_{PE}} = \frac{3,897.6}{25,225.0} = 0.155$

α	$F(2,60)$.	
.10	2.393.	ACCEPT
.05	3.150	ACCEPT
.01	4.977	ACCEPT

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

DDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD

Common Information for Problems 4, 5, and 6

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 maximize $E(Y)$. They used four doses of medicine: 0, 1, 2, and 3 units respectively. They randomly assigned 25 animals to dosage 0, 25 to dosage 1, 25 to dosage 2, and 25 to dosage 3. They observed that the average values of *Y* at each dosage were $y_{0\bullet} = 24, y_{1\bullet} = 136,\ y_{2\bullet} = 184,$ and $y_{3\bullet} = 152$, where $y_{i.}$ was the average of the observations taken with dosage $i = 0{,}1{,}2{,}3$, respectively. They also observed that $s_0^2 = 15{,}040.5, s_1^2 = 19{,}918.3,\ s_2^2 = 11{,}125.6,$ and $s_3^2 = 16{,}815.6,$ where s_i^2 was the unbiased estimate of the variance for the observations taken with dosage $i = 0{,}1{,}2{,}3$, respectively. They correctly calculated that $y_{\bullet\bullet} = 124$ and that the average s_i^2 was $15{,}725$. The total sum of squares was 1,872,800.0. They also correctly calculated the values of the three contrasts. The coefficients of the linear contrast were $-3, -1{,}1{,}3;$, and $\hat{\lambda}_{Lin} = 432$. The coefficients of the quadratic contrast were $1, -1, -1{,}1;$ and $\hat{\lambda}_{Quad} = -144$. The coefficients of the cubic contrast were $-1{,}3, -3{,}1;$ and $\hat{\lambda}_{Cubic} = -16$.

4. What are the values of the sum of squares due to the linear contrast, the sum of squares dues to the quadratic contrast, and the sum of squares due to the cubic contrast? This question is worth 20 points.

5. Report the analysis of variance table for the linear regression of *Y* on dosage. Use the sum of squares due to the linear contrast as the sum of squares for the regression of *Y* on dosage. Test the null hypothesis that there is no linear association at the 0.10, 0.05, and 0.01 levels of significance. What is the optimal setting of the dosage? This question is worth 40 points.

6. Test the null hypothesis that the linear model is adequate at the 0.10, 0.05, and 0.01 levels of significance. Report the analysis of variance table including the sum of squares due to lack of fit. This question is worth 50 points.

End of application of common information

4D $\hat{\lambda}_{LIN} = 432, \quad SS_{LIN} = \dfrac{(432)^2}{20/25} = 233{,}280.0$

$\hat{\lambda}_{QUAD} = -144., \quad SS_{QUAD} = \dfrac{(-144)^2}{4/25} = 129{,}600.0$

$\hat{\lambda}_{CUBIC} = -16, \quad SS_{CUBIC} = \dfrac{(-16)^2}{20/25} = 320.0.$

NOTE $SS_{LIN} + SS_{QUAD} + SS_{CUBIC} = 363{,}200 = SS_{TREATMENTS}$.

5D. ANOVA TABLE, CH11 LINEAR REGRESSION ANALYSIS

SOURCE	DF	SS	MS	F
LINEAR	1	233,280.0	233,280.0	13.94
ERROR	98	1,639,520.0	16,729.8	
TOTAL	99	1,872,800		

α	$F(\alpha,1,98)$.	
.10	2.757.	REJECT
.05	3.938	REJECT
.01	6.901.	REJECT

REJECT H_0: NO LINEAR ASSOCIATION VS H_1: LINEAR ASSOCIATION AT $\alpha = .01$.

THE LARGEST OBSERVED RESPONSE IS FOR DOSAGE 2.

FISHER'S LSD IS $t_{2.576,96}\sqrt{MSPE\left(\frac{2}{25}\right)}$

$= 2.628\sqrt{MSPE\left(\frac{2}{25}\right)} = 93.2.$ BY LSD, DOSE 0 IS ONLY.

REPORTING DOSE 2 AS

LESS THAN THE OTHERS. THE SIGNIFICANT LINEAR

OPTIMAL IS SUPPORTED BY THE

MOST ANALYSTS WOULD

TREND AND QUADRATIC TREND.

FIT A QUADRATIC MODEL TO THIS DATA.

6D ANOVA TABLE, LACK OF FIT FOR LINEAR MODEL.

SOURCE	DF	SS	MS.	F
LINEAR	1	233,2800		
LACK OF FIT	2	129,920.0	64,960	4.131
PURE ERROR	96	1,509,600.0	15,725	
TOTAL	99	1,872,800.0		

α	$F(2,96)$
.10	2.359.
.05	3.091
.01	4.833.

REJECT H_0 LINEAR MODEL ADEQUATE VS H_1 LINEAR MODEL NOT ADEQUATE AT $\alpha = .10$ AND $\alpha = .05$ BUT NOT AT $\alpha = .01$.

EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE

7. The random variable $Y, Y > 0$, has $E(Y) = \theta$ and $var(Y) = \theta^4, \theta > 0$. Find the approximate mean and variance of $W = Y^p, p \neq 0$. For what value of p is the approximate variance of W constant? (50 points).

$E(w) \cong \theta^p \quad p \neq 0; \ w = y^p. \ f'(w) = p(y^{p-1})$.

$VAR(w) = VAR(Y^p)$

$VAR(Y^p) \cong (f'(\theta))^2 \ VAR(Y) = p^2 (\theta^{p-1})^2 \theta^4$

$VAR(Y^p) \cong p^2 \theta^{2p-2} \theta^4 = p^2 \theta^{2p+2}$.

$\theta^{2p+2} = k$ WHEN $2p + 2 = 0, \ p = -1$

THAT IS, USE Y^{-1}.

ALSO

$VAR(Y) = \theta^4; \ \sigma_Y = \theta^2, \ \ln\sigma_Y = 2\ln\theta = 2\ln(\mu_Y)$

VARIANCE STABILIZING TRANSFORMATION IS

Y^{1-m} WHERE $m = 2$. USE Y^{-1}

EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE

8. 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 $vcv(Y) = V$ where V is an $n \times n$ positive definite matrix. Let $W = (X^T X)^{-1} X^T Y$, where $X^T$is the transpose of X.
    a. Find $E(W)$. (10 points)
    b. Find the variance-covariance matrix of W, $vcv(W)$. (50 points).

End of the Examination

a. $E(W) = E\left((X^T X)^{-1} X^T Y\right) = (X^T X)^{-1} X^T E(Y)$

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

b. $vcv(W) = vcv\left[(X^T X)^{-1} X^T Y\right] = vcv(MY)$

WHERE $M = (X^T X)^{-1} X^T$.

$vcv(W) = M\, vcv(Y)\, M^T = (X^T X)^{-1} (X^T V X)(X^T X)^{-1}$.