20 Lack of Fit Test

Common Information for Questions 1 and 2

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 animals were given one unit of dosage with observed average and sample variance (unbiased estimate) $y_{1.} = 1438$ and $s_1^2 = 635{,}849$; thirty were given two units of dosage with $y_{2.} = 612$ and $s_2^2 = 745{,}621$; thirty were given three units of dosage with $y_{3.} = 596$ and $s_3^2 = 546{,}237$; and thirty were given four units of dosage with $y_{4.} = 1390$ and $s_4^2 = 657{,}435$. The grand average of the outcome variables was $y_{..} = 1009$. The total sum of squares is 94,690,518. With regard to the orthogonal polynomials, the estimated linear contrast is -160, and its coefficients are $-3,-1,1,3$. The estimated quadratic contrast is 1620, and its coefficients are $1,-1,-1,1$. The estimated cubic contrast is 0, and its coefficients are $-1,3,-3,1$.

1. 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.
2. 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

1. $TSS = 94{,}690{,}518$ (GIVEN).

$$SS_{LINEAR} = \frac{(\hat{\lambda}_{LINEAR})^2}{[(-3)^2+(-1)^2+1^2+3^2]/30}$$

$$= \frac{(-160)^2}{[20/30]} = 38{,}400 \text{ ON 1 DF.}$$

$$SSERR = TSS - SS_{LINEAR}$$

$$= 94{,}690{,}518 - 38{,}400 = 94{,}652{,}118 \text{ ON } N-2 \text{ DF}$$

$$MSE = \frac{94{,}652{,}118}{(120-2)} = \frac{94{,}652{,}118}{118} = 802{,}136.59$$

$$\text{NOTE } R^2 = \frac{SS_{LINEAR}}{TSS} = \frac{38{,}400}{94{,}690{,}518} = 4.1\times 10^{-4}$$

ANOVA TABLE

LINEAR REGRESSION.

SOURCE	DF	SS	MS	F
LINEAR DOSE	1	38,400	38,400	0.048
ERROR	118	94,652,118	802,136.59	
TOTAL	119	94,690,518		

CRITICAL VALUES

α	$F(1, 118)$	DECISION
.10	2.749	ACCEPT
.05	3.921	ACCEPT
.01	6.855	ACCEPT

ACCEPT H_0: NO LINEAR ASSOCIATION BETWEEN DOSE AND RESPONSE AT $\alpha = .10$ (AND $\alpha = .05$ AND $\alpha = .01$).

2. LACK OF FIT TEST.

FIRST FIND MSPE, THE MEAN SQUARE FOR PURE ERROR.

$$MSPE = \frac{S_1^2 + S_2^2 + S_3^2 + S_4^2}{4}$$

$$= \frac{635,849 + 745,621 + 546,237 + 657,435}{4}$$

$$= \frac{2,585,142}{4} = 646,285.5 \text{ ON } 116 \text{ DF.}$$

$$SSPE = DFPE \times MSPE$$

$$= 116 \times (646,285.5) = 74,969,118 \text{ ON } 116 \text{ DF.}$$

THERE ARE TWO WAYS TO GET SS LOF

$$SSLOF = TSS - SSLIN - SSPE$$

$$= 94,690,518 - 38,400 - 74,969,118$$

$$= 19,683,000 \quad \text{ON } 2 \text{ DF}$$

$$DF\ SSLOF = DF\ TSS - DF\ SSLIN - DF\ SSPE$$

$$= 119 - 1 - 116 = 2.$$

SECOND WAY:

$$SS_Q = \frac{(\hat{\lambda}_Q)^2}{[(1)^2 + (-1)^2 + (-1)^2 + (1)^2]/J} = \frac{(1620)^2}{4/30}$$

$$= 19,683,000$$

$$SS_C = \frac{(\hat{\lambda}_C)^2}{[(-1)^2 + 3^2 + (-3)^2 + 1^2]/J} = \frac{0^2}{20/30} = 0.$$

$$SS_{LOF} = SS_Q + SS_C = 19,683,000 \quad \text{ON } 2 \text{ DF}$$

ANOVA TABLE

LACK OF FIT OF LINEAR MODEL.

SOURCE	DF	SS	MS
LINEAR DOSE	1	38,400	38,400
LACK OF FIT	2	19,683,000	9,841,500
PURE ERROR	116	74,969,118	646,285.5
TOTAL	119	94,690,118	

$$F_{LOF} = \frac{MS\,LOF}{MSPE} = \frac{9,841,500}{646,285.5} = 15.23$$

α	$F(2,116)$	DECISION
.10	2.349	REJECT
.05	3.074	REJECT
.01	4.793	REJECT

REJECT H_0: LINEAR MODEL ADEQUATE
AT $\alpha = .01$ (AND .05 AND .01)

ONE CAN CONCLUDE THAT A QUADRATIC MODEL WOULD BE ADEQUATE. SINCE $\hat{\lambda}_Q > 0$, THIS MODEL WILL HAVE A MINIMUM.

SINCE $y_{2\cdot} = 612$ AND $y_{3\cdot} = 596$, THE OPTIMAL DOSE IS BETWEEN 2 + 3 UNITS.