Instructor Solutions Manual

to accompany

Applied Linear Statistical Models

Fifth Edition

Michael H. Kutner
Emory University
Christopher J. Nachtsheim
University of Minnesota
John Neter
University of Georgia
William Li
University of Minnesota

2005
McGraw-Hill/Irwin
Chicago, IL
Boston, MA

PREFACE

This Solutions Manual gives intermediate and final numerical results for all end-of-chapter Problems, Exercises, and Projects with computational elements contained in *Applied Linear Statistical Models*, 5th edition. This Solutions Manual also contains proofs for all Exercises that require derivations. No solutions are provided for the Case Studies.

In presenting calculational results we frequently show, for ease in checking, more digits than are significant for the original data. Students and other users may obtain slightly different answers than those presented here, because of different rounding procedures. When a problem requires a percentile (e.g. of the t or F distributions) not included in the Appendix B Tables, users may either interpolate in the table or employ an available computer program for finding the needed value. Again, slightly different values may be obtained than the ones shown here.

We have included many more Problems, Exercises, and Projects at the ends of chapters than can be used in a term, in order to provide choice and flexibility to instructors in assigning problem material. For all major topics, three or more problem settings are presented, and the instructor can select different ones from term to term. Another option is to supply students with a computer printout for one of the problem settings for study and class discussion and to select one or more of the other problem settings for individual computation and solution. By drawing on the basic numerical results in this Manual, the instructor also can easily design additional questions to supplement those given in the text for a given problem setting.

The data sets for all Problems, Exercises, Projects and Case Studies are contained in the compact disk provided with the text to facilitate data entry. It is expected that the student will use a computer or have access to computer output for all but the simplest data sets, where use of a basic calculator would be adequate. For most students, hands-on experience in obtaining the computations by computer will be an important part of the educational experience in the course.

While we have checked the solutions very carefully, it is possible that some errors are still present. We would be most grateful to have any errors called to our attention. Errata can be reported via the website for the book: http://www.mhhe.com/KutnerALSM5e. We acknowledge with thanks the assistance of Lexin Li and Yingwen Dong in the checking of Chapters 1-14 of this manual. We, of course, are responsible for any errors or omissions that remain.

Michael H. Kutner
Christopher J. Nachtsheim
John Neter
William Li

Contents

X-ray: $\hat{Y} = 6.5664 + .0378X$

c. Infection risk: $MSE = 2.638$

Facilities: $MSE = 3.221$

X-ray: $MSE = 3.147$

1.46. a. Region 1: $\hat{Y} = 4.5379 + 1.3478X$

Region 2: $\hat{Y} = 7.5605 + .4832X$

Region 3: $\hat{Y} = 7.1293 + .5251X$

Region 4: $\hat{Y} = 8.0381 + .0173X$

c. Region 1: $MSE = 4.353$

Region 2: $MSE = 1.038$

Region 3: $MSE = .940$

Region 4: $MSE = 1.078$

1.47. a. $L(\beta_0, \beta_1) = \prod_{i=1}^{6} \frac{1}{\sqrt{32\pi}} \exp[-\frac{1}{32}(Y_i - \beta_0 - \beta_1 X_i)^2]$

b. $b_0 = 1.5969$, $b_1 = 17.8524$

c. Yes

Chapter 7

MULTIPLE REGRESSION – II

7.1. (1) 1 (2) 1 (3) 2 (4) 3

7.3. a. $SSR(X_1) = 1,566.45$, $SSR(X_2|X_1) = 306.25$, $SSE(X_1, X_2) = 94.30$, df: 1, 1, 13.

b. H_0: $\beta_2 = 0$, H_a: $\beta_2 \neq 0$. $SSR(X_2|X_1) = 306.25$, $SSE(X_1, X_2) = 94.30$, $F^* = (306.25/1) \div (94.30/13) = 42.219$, $F(.99; 1, 13) = 9.07$. If $F^* \leq 9.07$ conclude H_0, otherwise H_a. Conclude H_a. P-value = 0+.

7.4. a. $SSR(X_1) = 136,366$, $SSR(X_3|X_1) = 2,033,566$, $SSR(X_2|X_1, X_3) = 6,674$, $SSE(X_1, X_2, X_3) = 985,530$, df: 1, 1, 1,48.

b. H_0: $\beta_2 = 0$, H_a: $\beta_2 \neq 0$. $SSR(X_2|X_1, X_3) = 6,674$, $SSE(X_1, X_2, X_3) = 985,530$, $F^* = (6,674/1) \div (985,530/48) = 0.32491$, $F(.95; 1, 17) = 4.04265$. If $F^* \leq 4.04265$ conclude H_0, otherwise H_a. Conclude H_0. P-value = 0.5713.

c. Yes, $SSR(X_1)+SSR(X_2|X_1) = 136,366+5,726 = 142,092$, $SSR(X_2)+SSR(X_1|X_2) = 11,394.9 + 130,697.1 = 142,092$.

Yes.

7.5. a. $SSR(X_2) = 4,860.26$, $SSR(X_1|X_2) = 3,896.04$, $SSR(X_3|X_2, X_1) = 364.16$, $SSE(X_1, X_2, X_3) = 4,248.84$, df: 1, 1, 1, 42

b. H_0: $\beta_3 = 0$, H_a: $\beta_3 \neq 0$. $SSR(X_3|X_1, X_2) = 364.16$, $SSE(X_1, X_2, X_3) = 4,248.84$, $F^* = (364.16/1) \div (4,248.84/42) = 3.5997$, $F(.975; 1, 42) = 5.4039$. If $F^* \leq 5.4039$ conclude H_0, otherwise H_a . Conclude H_0. P-value = 0.065.

7.6. H_0: $\beta_2 = \beta_3 = 0$, H_a: not both β_2 and $\beta_3 = 0$. $SSR(X_2, X_3|X_1) = 845.07$, $SSE(X_1, X_2, X_3) = 4,248.84$, $F^* = (845.07/2) \div (4,248.84/42) = 4.1768$, $F(.975; 2, 42) = 4.0327$. If $F^* \leq 4.0327$ conclude H_0, otherwise H_a. Conclude H_a. P-value = 0.022.

7.7. a. $SSR(X_4) = 40.5033$, $SSR(X_1|X_4) = 42.2746$, $SSR(X_2|X_1, X_4) = 27.8575$, $SSR(X_3|X_1, X_2, X_4) = 0.4195$, $SSE(X_1, X_2, X_3, X_4) = 98.2306$, df: 1, 1, 1, 1, 76.

b. H_0: $\beta_3 = 0$, H_a: $\beta_3 \neq 0$. $F^* = (0.42/1) \div (98.2306/76) = 0.3249$, $F(.99; 1, 76) = 6.9806$. If $F^* \leq 6.9806$ conclude H_0, otherwise H_a. Conclude H_0. P-value = .5704.

b.

Cutoff	Renewers	Nonrenewers	Total
.40	18.8	50.0	33.3
.45	25.0	50.0	36.7
.50	25.0	35.7	30.0
.55	43.8	28.6	36.7
.60	43.8	21.4	33.3

c. Cutoff = .50. Area = .70089.

14.34. a. $z(.975) = 1.960$, $s^2\{b_0\} = 19.1581$, $s^2\{b_1\} = .00006205$, $s\{b_0, b_1\} = -.034293$

X_h	$\hat{\pi}_h'$	$s\{\hat{\pi}_h'\}$	
550	.0971	.4538	$.312 \leq \pi_h \leq .728$
625	1.5161	.7281	$.522 \leq \pi_h \leq .950$

b.

Cutoff	Able	Unable	Total
.325	14.3	46.2	29.6
.425	14.3	38.5	25.9
.525	21.4	30.8	25.9
.625	42.9	30.8	37.0

c. Cutoff = .525. Area = .79670.

14.35. a. $z(.975) = 1.960$, $\hat{\pi}_h' = -.04281$, $s^2\{b_0\} = .021824$, $s^2\{b_1\} = .000072174$, $s\{b_0, b_1\} = -.0010644$, $s\{\hat{\pi}_h'\} = .0783$, $.451 \leq \pi_h \leq .528$

b.

Cutoff	Purchasers	Nonpurchasers	Total
.15	4.81	71.54	76.36
.30	11.70	45.15	56.84
.45	23.06	23.30	46.27
.60	23.06	23.30	46.27
.75	48.85	9.64	52.49

c. Cutoff = .45 (or .60). Area = .82445.

14.36. a. $\hat{\pi}_h' = -1.3953$, $s^2\{\hat{\pi}_h'\} = .1613$, $s\{\pi_h'\} = .4016$, $z(.95) = 1.645$. $L = -1.3953 - 1.645(.4016) = -2.05597$, $U = -1.3953 + 1.645(.4016) = -.73463$.
$L^* = [1 + \exp(2.05597)]^{-1} = .11345$, $U^* = [1 + \exp(.73463)]^{-1} = .32418$.

b.

Cutoff	Received	Not receive	Total
.05	4.35	62.20	66.55
.10	13.04	39.37	52.41
.15	17.39	26.77	44.16
.20	39.13	15.75	54.88

c. Cutoff = .15. Area = .82222.

14.38. a. $b_0 = 2.3529$, $b_1 = .2638$, $s\{b_0\} = .1317$, $s\{b_1\} = .0792$, $\hat{\mu} = \exp(2.3529 + .2638X)$.

18.35. a. H_0: all μ_i are equal $(i = 1, ..., 4)$, H_a: not all μ_i are equal.

$MSTR = 955.5$, $MSE = 31.8$, $F_R^* = 955.5/31.8 = 30.1$,

$F(.95; 3, 32) = 2.90$. If $F_R^* \leq 2.90$ conclude H_0, otherwise H_a. Conclude H_a.

P-value = 0+

c. $\bar{R}_{1.} = 7.938$, $\bar{R}_{2.} = 22.375$, $\bar{R}_{3.} = 8.571$, $\bar{R}_{4.} = 27.962$,

$n_1 = 8, n_2 = 8, n_3 = 7, n_4 = 13$, $B = z(.99583) = 2.638$

Comparison	Testing Limits	
1 and 2	$-14.437 \pm 2.638(5.268)$	-28.334 and $-.541$
1 and 3	$-.633 \pm 2.638(5.453)$	-15.017 and 13.751
1 and 4	$-20.024 \pm 2.638(4.734)$	-32.513 and -7.535
2 and 3	$13.804 \pm 2.638(5.453)$	$-.580$ and 28.188
2 and 4	$-5.587 \pm 2.638(4.734)$	-18.076 and 6.902
3 and 4	$-19.391 \pm 2.638(4.939)$	-32.421 and -6.361

Group 1	**Group 2**	**Group 3**
Region 1	Region 2	Region 2
Region 3	Region 3	Region 4

18.36. Under H_0, each arrangement of the ranks 1, ..., 4 into groups of size 2 are equally likely and occur with probability $2!2!/4! = 1/6$. The values of F_R^* computed for the six arrangements are 0, .5, and 8, each occurring twice. Therefore the probability function $f(x)$ is:

x	$f(x) = P(F_R^* = x)$
0	1/3
.5	1/3
8	1/3

18.37. c. For the F distribution with $\nu_1 = 2$ degrees of freedom and $\nu_2 = 27$ degrees of freedom, the mean is:

$$\frac{\nu_2}{\nu_2 - 2} = 1.08$$

and the standard deviation is:

$$\frac{\nu_2}{\nu_2 - 2}\left[\frac{2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)}\right]^{1/2} = 1.17.$$

d. Expect 90% less than 2.51 and 99% less than 5.49.

$\hat{Y} = 2.75121 - .34885X_4 - .19441X_5 + .19925X_1X_4 + .19481X_1X_5$
$\quad -.01178X_2X_4 - .08433X_2X_5 - .22413X_3X_4 + .00731X_3X_5,$

$SSE(R) = 4.9506$

H_0: all α_i equal zero $(i = 1, 2, 3)$, H_a: not all α_i equal zero.

$F^* = (4.2326/3) \div (.7180/33) = 64.845$, $F(.99; 3, 33) = 4.437$.

If $F^* \leq 4.437$ conclude H_0, otherwise H_a. Conclude H_a. P-value = 0+

Degree:

$Y_{ijk} = \mu_{..} + \alpha_1 X_{ijk1} + \alpha_2 X_{ijk2} + \alpha_3 X_{ijk3} + (\alpha\beta)_{11} X_{ijk1} X_{ijk4}$
$\quad + (\alpha\beta)_{12} X_{ijk1} X_{ijk5} + (\alpha\beta)_{21} X_{ijk2} X_{ijk4} + (\alpha\beta)_{22} X_{ijk2} X_{ijk5}$
$\quad + (\alpha\beta)_{31} X_{ijk3} X_{ijk4} + (\alpha\beta)_{32} X_{ijk3} X_{ijk5} + \epsilon_{ijk}$

$\hat{Y} = 2.88451 - .44871X_1 - .09702X_2 + .36160X_3 - .06779X_1X_4 + .08939X_1X_5$
$\quad -.05349X_2X_4 - .03742X_2X_5 + .07190X_3X_4 - .10851X_3X_5,$

$SSE(R) = 8.9467$

H_0: $\beta_1 = \beta_2 = 0$, H_a: not both β_1 and β_2 equal zero.

$F^* = (8.2287/2) \div (.7180/33) = 189.10$, $F(.99; 2, 33) = 5.321$.

If $F^* \leq 5.321$ conclude H_0, otherwise H_a. Conclude H_a. P-value = 0+

e. $\hat{D}_1 = \hat{\mu}_{1.} - \hat{\mu}_{2.} = 2.1500 - 2.8067 = -.6567$, $\hat{D}_2 = \hat{\mu}_{1.} - \hat{\mu}_{3.} = 2.1500 - 3.1133 = -.9633$, $\hat{D}_3 = \hat{\mu}_{1.} - \hat{\mu}_{4.} = 2.1500 - 2.8400 = -.6900$, $\hat{D}_4 = \hat{\mu}_{2.} - \hat{\mu}_{3.} = -.3066$, $\hat{D}_5 = \hat{\mu}_{2.} - \hat{\mu}_{4.} = -.0333$, $\hat{D}_6 = \hat{\mu}_{3.} - \hat{\mu}_{4.} = .2733$, $s\{\hat{D}_1\} = .06642$, $s\{\hat{D}_2\} = .07083$, $s\{\hat{D}_3\} = .07497$, $s\{\hat{D}_4\} = .06316$, $s\{\hat{D}_5\} = .06777$, $s\{\hat{D}_6\} = .07209$, $q(.95; 4, 33) = 3.825$, $T = 2.705$

$$\begin{array}{rr}
-.6567 \pm 2.705(.06642) & -.836 \leq D_1 \leq -.477 \\
-.9633 \pm 2.705(.07083) & -1.155 \leq D_2 \leq -.772 \\
-.6900 \pm 2.705(.07497) & -.893 \leq D_3 \leq -.487 \\
-.3066 \pm 2.705(.06316) & -.477 \leq D_4 \leq -.136 \\
-.0333 \pm 2.705(.06777) & -.217 \leq D_5 \leq .150 \\
.2733 \pm 2.705(.07209) & .078 \leq D_6 \leq .468
\end{array}$$

f. $\hat{D}_1 = \hat{\mu}_{.1} - \hat{\mu}_{.2} = 2.3875 - 2.4675 = -.0800$, $\hat{D}_2 = \hat{\mu}_{.1} - \hat{\mu}_{.3} = 2.3875 - 3.3350 = -.9475$, $\hat{D}_3 = \hat{\mu}_{.2} - \hat{\mu}_{.3} = -.8675$, $s\{\hat{D}_1\} = .06597$, $s\{\hat{D}_2\} = .05860$, $s\{\hat{D}_3\} = .05501$, $q(.95; 3, 33) = 3.470$, $T = 2.4537$

$$\begin{array}{rr}
-.0800 \pm 2.4537(.06597) & -.242 \leq D_1 \leq .082 \\
-.9475 \pm 2.4537(.05860) & -1.091 \leq D_2 \leq -.804 \\
-.8675 \pm 2.4537(.05501) & -1.002 \leq D_3 \leq -.733
\end{array}$$

23.11. a. Full model:

$Y_{ijk} = \mu_{..} + \alpha_1 X_{ijk1} + \alpha_2 X_{ijk2} + \alpha_3 X_{ijk3} + \beta_1 X_{ijk4} + \beta_2 X_{ijk5} + \epsilon_{ijk}$

X_{ijk1}, $X_{ijk2} X_{ijk3}$, X_{ijk4}, X_{ijk5} defined same as in Problem 23.9a

Reduced models:

Factor A: $Y_{ijk} = \mu_{..} + \beta_1 X_{ijk4} + \beta_2 X_{ijk5} + \epsilon_{ijk}$