

Errata for *Applied Linear Regression*, Third Edition
June 19, 2009

We regret the existence of these errata. Please report any additional errata to sandy@stat.umn.edu. Words that are underlined are to be deleted. Words in **bold face** are to be added. I am particularly grateful to Miroslav Karamanov and Pierre Duchesne for finding many of these errors. Errata found after February 2006 are dated.

1. Page 2, line 8, change E. S. Pearson to **Karl** Pearson.
2. Page 2, point 2 near the middle of the page. Change "... scatterplot was..." to "... scatterplot **were**...".
3. Page 4, line 7. Indent "While".
4. (3/29/2006) Page 6, line 2 from the bottom: change seven to **eight**.
5. Page 21, eq. (2.3) change $(\hat{\beta}_0 + \hat{\beta}_1)$ to $(\hat{\beta}_0 + \hat{\beta}_1 x_i)$ (insert missing x_i).
6. Page 24, eq. (2.5), change the exponent to $1/2$ rather than 2:

$$\hat{\beta}_1 = \frac{SXY}{SXX} = r_{xy} \frac{sd_y}{sd_x} = r_{xy} \left(\frac{SYY}{SXX} \right)^{1/2}$$

7. Page 27 seven lines before the beginning of Section 2.5: The phrase "Under the assumption... in symbols as" is to be replaced by **The assumption that errors are independent of each other and of the predictors, and are normally distributed with common variance, is written in symbols as**
8. (1/22/2007) Page 33, line 3, change $-42.13\bar{6}$ to -42.138 .
9. (8/23/07) Page 35, unnumbered displayed equation before Section 2.8.4: The df given in the display are incorrect. The equation should read

$$\hat{\beta}_0 + \hat{\beta}_1 x_* \pm t(.025, \mathbf{1373}) \text{sepred}(D\tilde{h}eight_* | M\tilde{h}eight_*)$$

10. (1/22/2007) Page 35, line 5 from the bottom, change $E(Y|X = \underline{x})$ to $E(Y|X = x_*)$. Three lines later, change $\hat{y} = \beta_0 + \beta_1 x$ to $\hat{y} = \beta_0 + \beta_1 x_*$. In both cases, a subscript "*" is added to x .
11. (9/7/2007) The displayed equation on the last line of the page should be

$$\text{sefit}(\hat{y}|x) = \hat{\sigma} \left(\frac{1}{n} + \frac{(x - \bar{x})^2}{SXX} \right)^{1/2}$$

This changes x_* to x and \tilde{y}_* to \hat{y} .

12. (6/19/2009) Page 45, problem 2.12.2, change confidence interval to **prediction interval**.
13. Page 45, problem 2.13.4, change $x = \underline{CSpd}$ to $x = \mathbf{RSpd}$.
14. (9/28/2006) Page 47, line 9 from the bottom. change the base-two log fertility to **the natural log of fertility**.
15. (1/31/2007) Page 49, line 5 in Section 3.1.1: change 91% to **81%**.
16. Page 52, line 2 from the bottom (excluding the footnote), change 626 to **317**.
17. Page 52, line 11, change given by height divided by weight squared to **given by weight divided by height squared**.
18. Page 58, line 4 of Section 3.4.5. The equation for $\mathbf{X}'\mathbf{Y}$ is incorrect and should be:

$$\mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

That is, change $\sum \underline{y_i^2}$ to $\sum x_i y_i$.

19. Page 58, line 8 of Section 3.4.5. The displayed equation for $\hat{\beta}$ has two errors, indicated by the bold face entries, one in the two by two matrix, and one in the last vector below:

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \frac{1}{SXX} \begin{pmatrix} \sum x_i^2/n & -\bar{x} \\ -\bar{x} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \sum y_i \\ \sum \mathbf{x}_i\mathbf{y}_i \end{pmatrix}$$

20. (3/12/2007) Page 60, third equation from the top for RSS , should read

$$RSS = \mathcal{Y}'\mathcal{Y} - \hat{\beta}^{*'}(\mathcal{X}'\mathcal{X})\hat{\beta}^* = 193700$$

21. Page 61, Table 3.4 has errors in the df column and in the MS column:

Source	df	SS	MS	F	p-value
Regression	p	$SSreg$	$MSreg = SSreg/p$	$MSreg/\hat{\sigma}^2$	
Residual	$n - p'$	RSS	$\hat{\sigma}^2 = RSS/(n - p')$		
Total	$n - 1$	SYY			

22. Page 65, last line of Problem 3.1.1. Change "... between the height variables" to "... between **these same variables and compare to the scatterplot matrix**".

23. (1/22/2007) Page 69, line 3 in Section 4.1, change 18.54 to 18.55.

24. (5/11/2007) Page 74, equation (4.5), change η_0 and η_1 to β_0 and β_1 , respectively, to read:

$$E(\log(\text{Area}) | \log(\text{Length})) = \beta_0 + \beta_1 \log(\text{Length}) + \beta_2 E(\log(\text{Width}) | \log(\text{Length}))$$

25. Page 76, line before Section 4.1.7. Change (4.2) to **(4.1)**.

26. (9/24/2007) Page 81, equation (4.14) is missing a prime ('), and should read

$$\begin{pmatrix} \mathbf{x}_i \\ y_i \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma'_{xy} & \sigma_y^2 \end{pmatrix} \right)$$

27. Page 81, displayed equation after (4.15) should read:

$$\beta^* = \Sigma_{xx}^{-1}\Sigma_{xy}; \quad \sigma^2 = \sigma_y^2 - \Sigma'_{xy}\Sigma_{xx}^{-1}\Sigma_{xy} = \sigma_y^2(1 - \mathcal{R}^2)$$

(A minus sign has been inserted after the first σ_y^2 .)

28. (9/24/2007) Page 84, line 2 of Section 4.4.3, change $1 - SSreg / \sum y_i^2$ to $1 - RSS / \sum y_i^2$.

29. Page 85, line 8 from the bottom. Delete a blank space after "Rubin".

30. Page 88, line 7. Change $\beta_j T_j$ to $\beta_j T_j$ (second j should be a subscript).

31. (1/22/2007) Page 91, line 21, change 0.0399 to **0.399**.

32. (1/22/2007) Page 92, heading in Table 4.4, change Point Estimate to **Slope Estimate**.

33. Page 98, line -15, change μb to **mb**.

34. Page 101, line 2, change Table 5.4 to Table **5.2**.

35. (1/30/2007) Page 103, line 5 after Table 5.4, change $\hat{\sigma}_{pe}$ to $\hat{\sigma}_{pe}^2$.

36. Page 115, line 8 from the bottom (excluding the footnote but including the equation), change this line from "increasing but not linear, while in Figure 6.1b it is decreasing but not linear" to "**decreasing** but not linear, while in Figure 6.1b it is **increasing** but not linear".

37. Page 117, second and third sentences after Section 6.1.1 should be changed as follows:

With k predictors, the second-order model includes an intercept, k linear terms, k quadratic terms, and $k(k-1)/2$ interaction terms. If $k = 5$, the second-order mean function has **21** terms, and with $k = 10$, it has **66** terms.

38. (10/4/2007) Page 121, equation (6.13). The left-side of this equation should be $\text{Var}\left(g(\hat{\beta})\right)$, there is a hat missing from β .

39. Page 129, equation (6.21) should read

$$F_\ell = \frac{(RSS_\ell - RSS_1)/(df_\ell - df_1)}{RSS_1/df_1}$$

The division symbol is missing from the numerator of the fraction.

40. Page 132, last five lines of the page: Many of the values of sums of squares, degrees of freedom and test statistics are incorrect. Here is the correct version of the last five lines, with all changes shown in bold face:

that has a separate mean function for each of the two sexes. For the first of these two, we find $RSS = \mathbf{1457}$ with $202 - \mathbf{5} = \mathbf{197}$ df. For the second mean function $RSS = \mathbf{1168}$ with $202 - 8 = 194$ df. The value of the test statistic is $F = \mathbf{16.00}$ with $(\mathbf{3}, 194)$ df, for a corresponding p -value that is zero to three decimal places. We have strong evidence that the smaller of these mean functions is inadequate.

41. (10/12/2007) Page 133, equation (6.26), delete the term $\beta_1 \text{Sex}$.

42. Page 133, line 8 from the bottom: Change 1144.2 to **1186**.

43. Page 133, line 6 from the bottom: Change $F = 0.63$ to $F = \mathbf{1.47}$. Also, change "... p -value of about 0.53" to "... p -value of about **0.23**".

44. (2/6/2007) Page 151, In the legend for Figure 7.3, the value corresponding to the dotted line is -1 , not 1.

45. Page 154, line 4 after the table, change Acpts to **Acpt**.

46. (2/6/2007) Page 159, In the legend for Figure 7.7, the value corresponding to the dotted line is -1 , not 1.

47. (3/30/2007) Page 160, equation (7.10) should read:

$$\psi_{YJ}(U, \lambda) = \begin{cases} \psi_M(U + 1, \lambda) & \text{if } U \geq 0 \\ -\psi_M(-U + 1, 2 - \lambda) & \text{if } U < 0 \end{cases}$$

That is, insert a minus sign for the case $U < 0$.

48. Page 162, line 3 of problem 7.1.3, change log log to **log**.

49. Page 167, equation (8.1) change $\hat{\sigma}^2$ to σ^2 .

50. (3/12/2007) Page 168, equation (8.6), remove the "hat" on σ^2 , to read $\text{Var}(\hat{e}_i) = \sigma^2(1 - h_{ii})$

51. (3/12/2007) Page 169, last two lines should read, "where $\mathbf{x}'_i = (1, \mathbf{x}'_i)$ ", and $\bar{\mathbf{x}}$ is the mean of the \mathbf{x}'_i . The second term on the right-hand side of (8.11) is the equation of an ellipsoid centered at $\bar{\mathbf{x}}$.

52. Page 170, line 4, change 0.04, 0.06, 0.08 and 0.10 to **0.04, 0.06 and 0.08**.

53. Page 172, Figure 8.2a, the label for the vertical axis is missing and should be r_i .

54. Page 184, Table 8.4. The labels for "choice for Z " in the first two rows of the table are reversed: the first row is for *GasPres* and the second row is for *TankTemp*. The correct table is:

Choice for Z	df	S	p -value
<i>GasPres</i>	1	5.50	.019
<i>TankTemp</i>	1	9.71	.002
<i>TankTemp, GasPres</i>	2	11.78	.003
<i>TankTemp, GasTemp TankPres, GasPres</i>	4	13.76	.008
Fitted values	1	4.80	.028

55. (3/12/2007) Page 191, Problem 8.1.1, change to “Prove the results given by (8.8) and (8.9).”
56. (1/14/2008) Page 191, Problem 8.3.3 change $\sigma^2/Field$ to $\sigma^2 \times Field$.
57. (1/14/2008) Page 191, Problem 8.3.4 change $\sigma^2/Field^2$ to $\sigma^2 \times Field^2$.
58. (10/28/2007) Page 191, Problem 8.3, Change the third sentence to “In this analysis, the lab measurement is the response variable and the field measurement is the predictor variable.”
59. Page 197, line 7 change 1.36 to **1.359**. The next line, a displayed and unnumbered equation, should read:

$$R_{12} = \frac{\mathbf{1.359}}{0.379\sqrt{1 - .0639}} = \mathbf{3.708}$$

The second displayed equation should read

$$t_i = 3.708 \left(\frac{17 - 2 - 1}{17 - 2 - 3.708^2} \right)^{1/2} = 12.41$$

60. Page 197, second line of Section 9.1.4, change (2004) to (1994).
61. Page 213, line 2 after the table, change $\beta_3 \neq 0$ to $\beta_4 \neq 0$ (subscript is wrong).
62. Page 215, line 2 from the bottom, change with at least one $c_j \neq 0$ to **with at least two** $c_j \neq 0$. Also, change For that j to **For one of the j with** $c_j \neq 0$.
63. (1/22/2007) Page 216, line before (10.4), and the line before (10.5) change 10.7 to **10.6**.
64. Page 217, Equation (10.7), multiply p_C by 2:

$$AIC = n \log(RSS_C/n) + \mathbf{2}p_C$$

65. (6/19/2009) Page 239 line 14 from top change $D = 0.16$ to $A = 0.16$. Last displayed equation on the page, change $-\theta_3 D$ to $-\theta_3 A$. Page 240, line 2, change $-\theta_3 D$ to $-\theta_3 A$. Three lines later, change $-D$ to $-A$.
66. (2/8/2006) Page 242, last displayed equation, change the numbers in bold face:

$$\hat{\sigma}_{pe}^2 = \frac{SS_{pe}}{df_{pe}} = \frac{\sum(m-1)sd^2}{\sum(m-1)} = \frac{19916}{\mathbf{57}} = \mathbf{349.40}$$

67. (2/8/2006) Page 243, first displayed equation and the 2 lines following it, change the numbers shown in bold:

$$F = \frac{4326.1/10}{\mathbf{349.40}} = \mathbf{1.24}$$

which, when compared to the $F(10, \mathbf{57})$ distribution, gives a p -value of about **0.29**...

68. (2/8/2006) Page 243, Table 11.3, change 70 in the df column to **57**.
69. (2/8/2006) Page 244, first displayed equation and the line following, change the numbers as shown below in bold:

$$F = \frac{528.4/2}{\hat{\sigma}_{pe}^2} = \mathbf{0.75} \sim F(2, \mathbf{57})$$

for which $p = \mathbf{0.47}$...

70. Page 287, equation (A.26) has an extra **X** and should be

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{R}^{-1}(\mathbf{Q}'\mathbf{Y})$$

71. Page 293. The correct year for Barnett and Lewis is (1994).
72. Page 297, Pearson, E. S. and Lee, S. (1903) is out of alphabetical order, and should appear before Pinheiro and Bates. Also, the first author should be Pearson, **K.**, not Pearson, E. S. E. S. Pearson was Karl Pearson's son.
73. (8/23/07) Remove Pearson, E. S. from the index. Change the index entry for Pearson, K. to pages 2, 110, 297.