

Errata for Applied Linear Regression, Fourth Edition
May 18, 2021

We regret the existence of these errata. Please report any additional errata to sandy@umn.edu. ~~Strike out~~ text is to be deleted. Underlined text is to be added. Later printings of the book may have some of these errors corrected. I am grateful to all those who have taken the time to report the errors, particularly John Corbett and Michael Lavine.

1. (3/28/16) Page xv, 6, “Linear regression as an excellent...”
2. (3/28/16) Page 17, line 2, “...the response and ~~the~~ more than one...”
3. (9/10/14) Page 19, line 3 from the bottom, change 364 to 366.
4. (3/28/16) Page 21, line 3 from the bottom, change $E(e_i|x_i)$ to $E(e_i|X_i = x_i)$
5. (12/27/13) Page 26, line 1, change $-42.138^{\circ}\mathbf{F}$ to -42.138 .
6. (1/13/14) Page 26, eq. (2.10), left-side should be $\hat{\sigma}^2$ (the hat is missing)
7. (5/19/16) Page 27, line 2 from the bottom change $d_i = (1/n - c_i x_i)$ to $d_i = (1/n - c_i \bar{x})$. On the same line, change $\sum (d_i + c_i x_i) y_i$ to $\sum_j (d_j + c_j \bar{x}) y_j$.
8. (2/7/17) Page 33, eq. (2.17), replace σ by $\hat{\sigma}$ (the hat is missing)
9. (9/4/14) Page 35, equation (2.19), the enlarged bold-face character below needs to be substituted for Y:

$$\text{SSreg} = \text{SYY} - \left(\text{SYY} - \frac{(\text{SXY})^2}{\text{SXX}} \right) = \frac{(\text{SXY})^2}{\text{SXX}}$$

10. (9/19/17) page 57 and following. The fuel consumption data `fuel2001` contains a variable `income` that is defined to be the per capita income in 2000, in dollars. In several of the examples in the text, `income` was divided by 1,000, giving the income in thousands of dollars. This has no material effect on analyses because regression outcomes don't change when a predictor or regression is multiplied or divided by a constant. However, it will change the value of a regression coefficient by dividing or multiplying by that constant; see Problem 2.9.
11. (9/19/17) page 86, line 3 after (2.21), replace ~~sum of squares~~ by sum of squares
12. (9/19/17) page 38, line 10, replace ~~smaller~~ by larger.
13. (2/7/17) page 38, line 11 replace ~~interecept~~ by slope.
14. (9/17/14) Page 44, problem 2.10.3, line 2, change ~~RSS~~ to $\hat{\sigma}^2$.
15. (2/7/17) page 51, line 2 from the bottom, change ~~log(lifeExpF)~~ to lifeExpF.
16. (2/7/17) Page 52, Change the footnote to read: “There are a few localities with relatively large `log(ppgdp)` that have lower values of `lifeExpF` than would be expected by the overall trend in Figure 3.1a. Can you identify these localities and what they have in common (Problem 3.1)?”
17. (3/17/2014) Page 63, in the unnumbered equation between (3.16) and (3.17), a pair of parentheses is missing from the numerator:

$$\frac{(n - (p + 1)) \hat{\sigma}^2}{\sigma^2} \sim \chi^2(n - (p + 1))$$

18. (3/28/16) Page 67, two lines below (3.23), change “not ~~affected~~” to “not effected”
19. (9/10/14) Page 69, problem 3.1, change Figure 3.2 to Figure 3.1.
20. (10/12/17) Page 70, line 10 change “~~mass~~ in meters” to “height in meters.”

21. (1/13/14) Page 71, line 3, "... points in ~~in~~ an...".
22. (9/29/15) Page 72, replace the displayed equation that is part of Problem 3.7 at the top of the page by

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{12}\mathbf{E}_{22}^{-1}\mathbf{A}'_{12} & -\mathbf{A}_{12}\mathbf{E}_{22}^{-1} \\ -\mathbf{E}_{22}^{-1}\mathbf{A}'_{12} & \mathbf{E}_{22}^{-1} \end{pmatrix}$$

where $\mathbf{E}_{22} = \mathbf{A}_{22} - \mathbf{A}'_{12}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$. Thus each appearance of \mathbf{A}_{22} is replaced by \mathbf{E}_{22} .

23. (9/19/17) page 73, line 3 after (4.1), change gallons to gallons per person.
24. (9/19/17) Page 77, line -2 $\text{WT9} = \text{DW9} - \text{WT2}$ should be $\text{DW9} = \text{WT9} - \text{WT2}$.
25. (5/16/2021) Page 79, lines 2 and 3 of section 4.1.2, change "...vector of constants \mathbf{a} ..." to "...nonzero vector of constants \mathbf{a} ..."
26. (10/12/17) Page 79, 11 lines above the footnote, change "precipitation is lower" to "precipitation is higher".
27. (10.12.17) Page 82, line 4, change ~~effect~~ to effects
28. (9/12/17) Page 82, equation (4.4), In the last two lines of this equation change $\exp(\beta_1)$ to $\exp(\beta_j)$, and on the last line change X_2 to $X_{(j)}$ and \mathbf{x}_2 to $\mathbf{x}_{(j)}$.
29. (10/26/2015) Page 83, 2 lines after the second displayed equation, change " $= -2.8\%$ decrease" to " $= 2.8\%$ decrease".
30. (9/19/17) 3 lines above (4.5) change ~~this rarely makes sense to~~ this is meaningful only if both the regressor and the response are appropriately scaled.
31. (9/12/17) Page 83. The last line of equation (4.5) change $X_1 = x_1, X_2 = \mathbf{x}_2$ to $X_j = x_j, X_{(j)} = \mathbf{x}_{(j)}$.
32. (5/18/2021) Page 85, 8 lines from the bottom, change " η_0 and η_1 " to " β_0 and β_1 ".
33. (3/28/16) First line of Section 4.5.2 on page 92 should read "The sample squared multiple correlation coefficient R^2 can be shown to be the square of the correlation between ..."
34. (9/6/2016) Pare 99, line 4 of Section 5.1.1, the reference to Mosteller and Tukey (1977) should be to Tukey, J. (1977). EDA, Addison-Wesley.
35. (9/5/2016) Page 99, sentence beginning on line 7 from the bottom should read: The 'whiskers' extend to the most distant points from the median that are no more than 1.5 IQRs from the median"
36. (3/28/16) Page 104, equation (5.7), change $\beta_{22}U_2x$ to $\beta_{12}U_2x$ (subscript on β is incorrect). Also delete the label (5.7) from this equation, so it has no label
37. (3/28/16) Page 105, first displayed equation $\text{lifeExpF} \sim \text{group} + \log(\text{ppgdp}) + \text{group}:\log(\text{ppgdp})$ should be numbered as equation (5.7)
38. (3/28/16) In Table 5.3, the coefficient estimates incorrectly used logarithms to the base 2 rather than natural logarithms, so smoe of the estimates and standard errors are wrong. Here is the correct table, with the corrected values underlined:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept), $\widehat{\beta}_0$	59.2137	15.2203	3.89	0.0001
other, $\widehat{\beta}_2$	-11.1731	15.5948	-0.72	0.4746
africa, $\widehat{\beta}_3$	-22.9848	15.7838	-1.46	0.1470
$\log(\text{ppgdp}), \widehat{\beta}_1$	<u>2.2425</u>	<u>1.4664</u>	1.53	0.1278
other: $\log(\text{ppgdp}), \widehat{\beta}_{12}$	<u>0.9294</u>	<u>1.5177</u>	0.61	0.5410
africa: $\log(\text{ppgdp}), \widehat{\beta}_{13}$	<u>1.0950</u>	<u>1.5785</u>	0.69	0.4887

$$\widehat{\sigma} = 5.1293 \text{ with } 193 \text{ df, } R^2 = 0.74$$

39. (3/28/16) Page 107, line 4 from the bottom, “. . . one effects plots with. . .”
40. (3/28/16) Page 113, line 8 in section 5.4, change “doesn’t effect” to “doesn’t effect”
41. (9/23/14) Page 124, Table 5.8. Immediately before **year** on the third line insert a new line:
region A factor with levels giving the geographic names of six economic regions of Minnesota. Excluded economic regions had few farm sales.
42. (6/27/14) Page 124, Table 5.8. In the definition of **crp**, change ~~Enrolled of any~~ to Any.
43. (12/21/2016) Page 138, 3 lines above Section 6.1.1, change the first two words on this line from ~~in X_2~~ to in X_1 .
44. (3/28/16) Page 140, middle of the first bullet starting “Type I”, change the sentence beginning “One result” to “One result of this is that one of the interactions, **A:B** is adjusted for none of the other interactions, another, **A:C** is adjusted for **A:B**, and **B:C** is adjusted for both **A:B** and **A:C**.” Only the underlined text is changed.
45. (10/30/2015) Page 141, line 10 from the bottom, change “**len** and **load**” to “**len** and **amp**”
46. (10/20/14) Page 146, equation (6.21) should read

$$F = \frac{(\mathbf{L}\hat{\underline{\beta}} - \mathbf{c})'(\mathbf{L}\hat{\mathbf{V}}\mathbf{L}')^{-1}(\mathbf{L}\hat{\underline{\beta}} - \mathbf{c})}{q}$$

47. (4/7/2014) Page 147, line 5 from the bottom, change “graph ~~if~~” to “graph of”.
48. (7/25/14) Page 150, line 8 of Section 6.6.8, different outcomes
49. (3/28/16) Page 158, equation (7.6), replace β_2s by β_2s^{-1} .
50. (11/16/2015) Page 158, last line, ~~comparing~~
51. (3/28/16) Page 162, line 2, change “statum” to “stratum”
52. (3/17/2014) Page 166, at the bottom of the page, change **GasPres** to **TankTemp** in three places:
- 4 lines above the displayed equation, change ~~$Z = \mathbf{GasPres}$~~ to $Z = \mathbf{TankTemp}$.
 - Replace the displayed equation by
$$\text{Var}(Y|X, Z) = \sigma^2 \times \mathbf{TankTemp}$$
 - Last line on the page, change ~~1/**GasPres**~~ to 1/**TankTemp**.
53. (3/28/16) Page 169, line 14 in Section 7.4 change “db” to “dB”. Make the change in the horizontal axis labels to Figures 7.4 and 7.5.
54. (3/28/16) Page 169, line 7 from the bottom, change “likely to ~~the~~ correlated” to “likely to be correlated”
55. (9/6/2016) Page 170, Equation (7.14), change y to Y .
56. (9/6/2016) Page 170, Equation (7.15), change loudness_{ij} to x
57. (7/21/2014) Page 173, 3 lines above equation (7.20), change ~~j th~~ to i th.
58. (7/21/2014) Page 182, problem 7.8.2. Change $\text{Var}(\text{Weight}|\text{Age}) = \mathbf{n}\sigma^2/\mathbf{SD}^2$ to $\text{Var}(\text{Weight}|\text{Age}) = \mathbf{SD}^2/\mathbf{n}$.
59. (12/21/2016) Page 187, Figure 8.2(d), vertical axis label should read $\sqrt{\text{textttBrainWt}}$.

60. (11/17/2014) Page 200, Problem 8.2.3. The first sentence should read: Show that using $\lambda = 2$ to transform the predictor Speed as in Problem 8.2.2 does match ...”
61. (11/14/2014) Page 203, in the fourth line of Problem 8.6.2, change ~~amp by load~~ to len:amp.
62. (12/7/2015) Page 203, second to last line in 8.6.4, change horizontal axis to vertical axis.
63. (12/21/2016) Page 214, third line from the bottom, change $\mathbf{x}'_j\boldsymbol{\beta}$ to $\mathbf{x}'_i\boldsymbol{\beta}$
64. (9/6/2016) Page 220, line after (9.22) “. . . is the ordinary squared Euclidean. . .”
65. (9/6/2016) Page 225, Section 9.6. The unnumbered displayed equation, substitute $\boldsymbol{\beta}$ in place of $\hat{\boldsymbol{\beta}}$.
66. (9/6/2016) Page 228, Problem 9.5 should refer to Appendix A.9, not A.13.
67. (9/6/2016) Page 237, Section 10.2, line 1, “. . . is to discover which of ~~a~~ many. . .”
68. (12/3/2014) Page 238, first line of (10.5) change ~~\mathbf{x}_{p_C}~~ to bold face, \mathbf{x}_{p_C} , to read

$$E(Y|X_C = \mathbf{x}_{p_C}) = \beta_0 + \boldsymbol{\beta}'_{p_C}\mathbf{x}_{p_C}$$

69. (12/3/2014) Page 239, equation (10.7) ~~RSS $_{p_C}$~~ should read RSS_{p_C} :

$$BIC = n \log(RSS_{p_C}/n) + \log(n)p_C$$

70. (12/3/2014) Page 240, line 18, change ~~$k + (k - 1)/2$~~ to $k(k + 1)/2$. On the next line, change 45 to 55.
71. (12/3/2014), Page 240, line 12 from the bottom, insert a word, to read from the beginning of the line “. . . method, where at each step . . .”.
72. (12/21/2016) Page 244, line 7 of Section 10.2.3, change ~~spareity~~ to sparsity.
73. (9/6/2016) page 273, line 4 from the botton ($Y_i|X_i = x_i$) should be ($Y_i|X_i = \mathbf{x}_i$).
74. (10/13/2014) Page 274, 9 lines above the table at the bottom of the page, “. . . samples sizes. . .”. Two lines later, change “the the” to “the”
75. (10/13/2014) Page 274, 5 lines above the table at the bottom of the page, change ~~$m = 3$~~ to $m = 2$.
76. (9/6/2016) Page 274, 2 lines above the table at the bottom of the page, “The results from fitting. . .”
77. (12/29/2015) Page 278, replace the last 6 lines on the page by:

Figure 12.4 is a summary effects plot for the blowdown data. The model fit suggests that effects of ~~\mathbf{s}~~ and ~~\mathbf{d}~~ are needed for most tree species. We see there are interesting differences between species. For ~~black ash red pine~~, the probability of blowdown appears to decrease with ~~\mathbf{d}~~ , while for ~~paper birch jack pine~~ the probability of blowdown may be independent of ~~\mathbf{d}~~ . ~~Cedar trees were relatively immune to blowdown except in areas of very high severity.~~ Further analysis of these data would require more work, and quite likely a separate analysis for each species ~~separately~~ could be enlightening.

78. (10/13/2014) Page 280, Table 12.4, change the heading ~~Level~~ to Type.
79. (10/21/2014) Page 280, Equation (12.11) has a missing y_i and should read

$$G^2 = 2 \sum_{i=1}^n \left[y_i \log(y_i/\hat{y}_i) - (y_i - \hat{y}_i) \right]$$

80. (9/6/2016) Page 283, 4 lines above Section 12.4, change “smaller samples χ^2 . . .” to “smaller samples X^2 . . .”.

81. (9/6/2016) Page 285, Problem 12.1.1, “Create a table that gives the number...”
82. (10/21/2014) Problem 12.2, line 5, change to read “... is physically attractive. The predictors are `gender, discipline, numYears,`”
83. (10/14/2014) Page 288, problem 12.5.4, replace `age + age^2` by `poly(age, 2)`