Errata

This document contains corrections to errors in the printing of *Regression Analysis and Linear Models* that slipped through editorial screening. The correction is in **bold and underlined**. The book is currently in its second printing. To determine your printing number, look at the copyright page. You will see a string of numbers such as "10 9 8 7 6 . . . ". If this string of numbers ends in "1", for example, you have a first printing. The errors listed in the first printing were fixed in the second printing. This document was produced **February 27, 2020**.

First printing

Page 24, bottom of the page

The formula at the bottom of the page is for Var(X) rather than the nonexistent Var(XY). It should read

$$\text{Var}(\underline{\boldsymbol{X}}) = \frac{N\sum_{i=1}^{N} X_i^2 - (\sum_{i=1}^{N} X_i)^2}{N^2}$$

Page 61, bottom of the page

> "The same procedure can be used to generate b_1. In a crosswise regression, estimate X_1 from X_2 and construct the residuals $X_{1.2} = \underline{\boldsymbol{X_1 - \hat{X}_1}}$."

Page 74, middle of the page

> "Many people use $\tilde{b}_j$ as a measure of the relative importance of a regressor in a model. That is, if $\underline{\tilde{\boldsymbol{b}}_{\boldsymbol{1}}}$ is larger than $\underline{\tilde{\boldsymbol{b}}_{\boldsymbol{2}}}$, then some would say that X_1 is more important in a statistical sense than is X_2 in estimating Y. "

Page 106, toward the bottom of the page

The tabled critical value used in the computation of the confidence interval is different than the critical value provided in the line above. It should read

> "For instance, for a two-tailed 95% confidence interval, the critical value of t when $df_{residual} = 6$ is 2.447. Thus, in this example, a 95% confidence limit for ${}_{T}b_1$ is 1.045 ± **2.447** × 0.422 = 0.012 to 2.078."

Page 185, middle of the page

The two formulas on this page contain a couple of errors. The two formulas should read

$$\rho^4 = (R_a^2)^2 - \frac{2k(1 - R_a^2)^2}{(N-1)(N-k+1)}$$

and

$$RS = \sqrt{\frac{(N-k-3)\rho^4 + R_a^2}{R_a^2(N-2k-2)+k}}$$

Page 233, toward the top of the page

> "We can think of $\boldsymbol{b_1 - b_2}$ as a weighted linear combination of regression coefficients, just as is $\hat{Y}$"

Page 237, Table 8.2

The second column heading should start "Adding $\boldsymbol{j}$" and the third column heading should start "Adding $\boldsymbol{i}$".

Page 246, just under Table 9.1

The first occurrence of the word "married" in the text below the table should be "divorced." So the sentence should read

> "This model would lead to the claim that **divorced** people (X_1 = 2) are estimated to have a life satisfaction of $\hat{Y} = 89.104 - 3.725 \times 2 = 81.654$ units on average, single people ($X_1 = 3$) are estimated to have a life satisfaction of $\hat{Y} = 89.104 - 3.725 \times 3 = 77.929$ units on average, and the correlation between actual and estimated life satisfaction is 0.536. "

Pages 259-268

On pages 259 and 265, all instances of the regression constant should be **44.178** rather than 44.032.

At the bottom of page 265 and the top of page 266, there are two instances of 3.828 that should be **3.628**.

On page 267, the adjusted means for the single and divorced groups were reversed. So two of the equations on this page should read

adj. $\overline{Y}_2 = 44.178 + 9.000(0) + 3.628(1) + 8.329(0) + 0.536(56.9) =$ **78.304**

adj. $\overline{Y}_3 = 44.178 + 9.000(0) + 3.628(0) + 8.329(1) + 0.536(56.9) =$ **83.005**

and on page 268, the last sentence in the paragraph at the top should say

> "Similarly, $b_2 = 3.628$ is the distance between the **divorced** line and the widowed line, or the difference between the adjusted means of these two groups; b_3 is interpreted equivalently for the comparison between single and widowed."

Page 271, near the bottom

> "It is equal to $SS_{Mstatus+Income}$, from the model with both income and marital status (A + B + C = 910.843), minus $SS_{Mstatus}$, which is $SS_{regression}$ from a model with just marital status as a predictor (A + B =**500.050**)."

Page 381, toward the top

> "So we have a linear model of the conditional effect of $\boldsymbol{X_1}$ and hence *simple linear interaction*."

Page 382, toward the top

The sign for the weight for the crossproduct is reversed. The equation should be

$$\hat{Y} = 10 - 0.3X_1 + 2X_2 \mathbf{+} 0.1X_1X_2$$

Page 389, toward the bottom

"As is apparent in Figure 13.5, it seems that exhaustion has a bigger effect on the use of safety protocol work-arounds among those with less experience working on the job, as reflected in the steeper slope of the line when job tenure X_2 is **smaller**."

Pages 403-404

"This change in the difference with a 1 unit change in X_1 is $b_4 = -0.071$, and it is invariant to where you start on $\boldsymbol{X_1}$."

Page 420, just above equation 14.10

"So for **three** groups,"

Page 430, at the bottom

There is a subscript missing on the squared term in the equation at the bottom. It should read

$$\hat{Y} = 15.228 - 4.060X_1 + 0.340\boldsymbol{X_1}^2$$

Page 440, near the bottom

X_2 and its coefficient is missing from the equation. It should read

$$\hat{Y} = b_0 + b_1X_1 + b_2X_1^2 + b_3X_1X_2 + b_4X_1^2X_2 \boldsymbol{+ b_5X_2}$$

Page 462, near the bottom

The PROCESS commands on page 462 are outdated because they are for PROCESS version 2. This version of PROCESS is no longer available. The proper PROCESS command for version 3 of PROCESS is

```
process y=injury/x=exhaust/m=safety/cov=injuryb/model=4/total=1/
    seed=15456.
```

The equivalent command in the SAS version of PROCESS is

```
%process (data=hospital,y=injury,x=exhaust,m=safety,cov=injuryb,model=4,
    total=1,seed=15456);
```

Page 605, toward the top of the page

There is a small error in the R code for producing standardized regression coefficients. The code should be

```
lm(scale(health$wtloss)~scale(health$exercise)+scale(health$food))
```

Second printing

Page 185, middle of the page

The two formulas on this page contain a couple of errors. The two formulas should read

$$\rho^4 = (R_a^2)^2 - \frac{2k(1 - R_a^2)^2}{(N-1)(N-k+1)}$$

and

$$RS = \sqrt{\frac{(N-k-3)\rho^4 + R_a^2}{R_a^2(N-2k-2)+k}}$$