

# Errata for *Computational Statistics*, 1st Edition, 2nd Printing

Geof H. Givens and Jennifer A. Hoeting

March 25, 2014

The final paragraph of the preface indicates if you own a 2nd printing copy.

Here is a list of corrections and other notes. We appreciate comments from our careful readers, including Jim Albert, Shan Ba, Jim Brennan, Shoja'eddin Chenouri, Hugh Chipman, Mark Delorey, Stephanie Fitchett, Doug Gorman, Andrew Hill, Michael Höhle, Quiming Huang, Mori Jamshidian, Yueyang Jiang, Wentao Li, Duncan Murdoch and Jason Song.

Corrections are offered in later printings now available from Wiley.

## Website:

- Dataset `alzheimers.dat`: as of 11/21/05 this is no longer missing.
- Dataset `baseball.dat` (9/29/05) and `mars.dat` (11/29/05): the final few lines were mysteriously truncated. The datasets and the overall zipfile were updated on the dates indicated.

## Chapter 1:

- Page 8, the multivariate normal density needs a minus sign inside the curly braces.
- Page 9, first line below (1.25). Replace “is a convex function” with “is a strictly monotonic function”.
- Page 11, two lines below equation (1.30), clearly  $c = 1 / \int f(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathbf{x}) d\boldsymbol{\theta}$ .
- Page 12, last paragraph above Example 1.2. The word *influence* is misspelled.

## Chapter 2:

- Page 39 first paragraph of section 2.2.2.3 and page 40 first paragraph. Information is gained about the curvature of  $\mathbf{g}$  not  $\mathbf{g}'$ .

## Chapter 3:

- The AIC values in this chapter are 2 units too low.
- Section 3.2, last sentence of third paragraph. It is slightly clearer to say “If the neighborhood is defined by allowing as many as  $k$  changes to the current candidate solution in order to produce the next candidate, then it is a *k-neighborhood*, and the alteration of those features is called a *k-change*.”
- Section 3.5.1.1, third sentence. Replace “a individual” with “an individual”.

- Page 76, Figure 3.6: graphical convention would normally have the vertical arrow pointing up, in the direction of increasing fitness.
- Section 3.5.2.2, last sentence of third paragraph. Replace “Such an...” with “Such a...”
- Exercise 3.4. The steady state GA should have  $G = 1/P$ .

#### Chapter 4:

- Page 95, the first equation *below* (4.18), there is a log missing in the last term on the right hand side. In other words, the correct equation is  $E\{\log f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta})|\mathbf{x}, \boldsymbol{\theta}^{(t)}\} = E\{\log f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\theta})|\mathbf{x}, \boldsymbol{\theta}^{(t)}\} - E\{\log f_{\mathbf{Z}|\mathbf{X}}(\mathbf{z}|\mathbf{x}, \boldsymbol{\theta})|\mathbf{x}, \boldsymbol{\theta}^{(t)}\}$
- Page 98, fifth line up from the bottom of example 4.4, the end of the line should be  $p_C$ , not  $p_c$ .
- Section 4.2.3. We have received the following email from Mori Jamshidian expressing his view of the SEM algorithm.

I'm using your text for my computational stats class, and it's been very good, especially in terms of the topics covered. When covering Chapter 4, Section 4.2.3 on EM variance estimation, I noticed that you cover SEM algorithm as one of the main algorithms for EM standard error estimation. In a paper that you have also cited in your book (Jamshidian and Jennrich 2000, JRSS-B) we have noted that SEM does not have a solid theoretical foundation, and have explained why it's prone to all sorts of numerical inaccuracies. Thus, we recommend that the SEM method not be used at all. You mention the method in Jamshidian and Jennrich (2000) as a “more sophisticated numerical differentiation strategy.” It turns out that implementation of the methods in Jamshidian and Jennrich (2000) are much simpler than that of SEM, and as we show in our paper they result in highly accurate results. In our view, SEM is a somewhat unsuccessful attempt in using numerical differentiation in the context of EM, as we explain in our paper. Just thought to bring it up, in case you may find this useful for your future editions of the book.

- Page 102, third line of example 4.6 should be  $\hat{p}_T = 0.0132$  not 0.132.
- Page 104, equation (4.49): Omit the  $\delta_i$  from the denominator of this expression.
- Page 104, the line above equation (4.51) should begin “for  $k = 1, \dots, C$ ”.
- Page 112, the line above equation (4.78) should begin “Finally, note that  $\mathbf{b}^{(t)}, \dots$ ”.

#### Chapter 6:

- Page 150, example 6.2. Replace  $\log \lambda \sim N(4, 0.5^2)$  with  $\log \lambda \sim N(\log 4, 0.5^2)$ .
- Page 171, the sentence on lines 5–6 should read “Hence,  $h_2(1 - \mathbf{U}_{i1}, \dots, 1 - \mathbf{U}_{im}) = h_1(F_1^{-1}(1 - U_{i1}), \dots, F_m^{-1}(1 - U_{im}))$  is monotone in each argument and has the...”.
- Page 178, problem 6.3. There is a portion missing here. Part (a $\frac{1}{2}$ ) is: Repeat the estimation using rejection sampling. Then the last line of part (b) should read “...to the output obtained in part (a $\frac{1}{2}$ )” and part (c) should ask you to compare (a $\frac{1}{2}$ ) to (b).

## Chapter 8:

- Page 224, paragraph 2, line 5 “parameters” is misspelled.
- Page 232, step 2, line 2 should read “. . .and run each chain. . .”, and line 4 should read  $X_k^{(0)} = q\left(X_k^{(-1)}, \mathbf{U}^{(0)}\right)$ .
- page 232, step 3, third line from end replace “much” with “must”.
- pge 237, equation (8.17), there should be a minus sign directly preceding the summation symbol.
- In the first bullet of exercise 8.5, the matrix should be  $22 \times 22$ , not  $42 \times 42$ .

## Chapter 9:

- Section 9.2.4, last sentence of first paragraph. The bootstrap estimate of the bias is  $\sum_{i=1}^B (\hat{\theta}_i^* - \hat{\theta})/B = \bar{\theta}^* - \hat{\theta}$ .
- Page 271, bottom paragraph. Delete the sentence beginning “For two-sided intervals. . .” and the clause after the colon in the subsequent sentence. The improvement offered by the nested bootstrap depends on the accuracy of the original interval and the type of interval. In general, nested bootstrapping can reduce the rate of convergence of coverage probabilities by an additional multiple of  $n^{-1/2}$  or  $n^{-1}$ . See the cited references.

## Chapter 11:

- Page 336-7, Figure 11.13: The line types don’t match the text or the previous figure. We’ll fix the figure in the third printing, but for now, the caption and the third line on page 337 should indicate  $j = 1, h_1 = 0.05n$  (solid);  $j = 2, h_2 = 0.2n$  (dotted); and  $j = 3, h_3 = 0.5n$  (dashed).
- Page 337, fourth line from the bottom should read “ $\mathbf{Y}^* = \hat{\mathbf{Y}} + \mathbf{e}^*$ ”.

## Chapter 12:

- Page 366, item 3, two instances of “training-set” should be replaced by “validation-set”.