Errata

An R and S-PLUS Companion to Applied Regression

John Fox

10 September 2008

Please note that some (or all) of these errors may be corrected in your printing of the book.

1. General: Many of the examples use `T` for `TRUE` and `F` for `FALSE`. This is a perfectly safe practice in S-PLUS, but it can occasionally cause problems in R, where `T` and `F` are not reserved symbols. That is, the user can redefine the values of `T` and `F`, including such a perverse assignment as `T <- FALSE`. It is therefore good practice in R to use `TRUE` and `FALSE` in preference to `T` and `F`.

2. Pages 15 and 48 (Boxes on `remove` in S-PLUS): I should have pointed out that an alternative is to use the `rm` function: e.g., `rm(x, y, z)` works in both R and S-PLUS.

3. Page 30: In the code creating the graph of hat values in Figure 1.7(a), the call to `identify` misspells the name of the `hatvalues` function as `hat.values`; the line should read:

   ```
   > identify(1:45, hatvalues(duncan.model), row.names(Duncan))
   ```

4. Page 44, footnote 3: I imply that specifying `header=T` is *necessary* in S-PLUS in this example; this is not the case.

5. Page 46: The example meant to illustrate the use of `read.fwf` (near the top of the page) inadvertently duplicates the material in the box immediately below (showing how to use `scan` to read fixed-format data). The intended command is as follows:

   ```
   > Prestige <- read.fwf('d:/data/Prestige-fixed.txt',
   +     row.names='occupation',
   +     col.names=c('occupation', 'education', 'income', 'women',
   +         'prestige', 'census', 'type'),
   +     widths=c(25, 5, 5, 5, 4, 4, 4))
   >
   ```

6. Pages 59 and 151: The text misleadingly states that there is no `na.action` option in S-PLUS. There is an `na.action` option, but it is initially unset, in which case (as stated) `na.fail` is invoked by statistical-modeling functions such as `lm`. You can set the `na.action` option — e.g., `options(na.action="na.omit")`. If you *always* want to set the `na.action` option at the start of a session, you may do so by defining a `.First` function, which is automatically executed at session start-up; for example:

   ```
   .First <- function() {
       options(na.action = "na.omit")
       }
   ```

   Other options may be set in `.First` as well.

   It is also possible to associate an `na.action` with a specific data frame, by setting an `na.action` attribute, as in `attr(df, "na.action") <- na.omit`. The `na.action` attribute will be invoked when the data frame `df` is passed via the `data` argument to a statistical-modeling function such as `lm`.

7. Page 68: The entry under the bulleted item *Vectors* should read: "One-dimensional arrays of numbers, character strings, or logical values (i.e., `TRUE` or `FALSE`). Single numbers, character strings, or logical values in S are treated as vectors of length 1."

8. Page 98: The parenthetical reference at the end of the first paragraph should read: "(a topic described in detail in Chapter 7)."

9. Page 156: The line for the probit link in Table 5.1 reverses the link and inverse-link (mean) functions; these should read $\Phi^{-1}(\mu_i)$ for the link function in column 2, and $\Phi(\eta_i)$ for the inverse-link in column 3 of the table.

10. Page 159, Figure 5.1: To conform to the text, the vertical axis should be labelled $\mu = g^{-1}(\eta)$; i.e., the link function is g, not L.

11. Pages 162, 221, 306: "Akaike" (in the Akaike Information Criterion, or AIC) is consistently misspelled as "Akiake."

12. Page 173: A command, defining the response variable `participation`, is omitted from the example:

```
> participation <- ordered(partic, # reorder levels
+     levels=c('not.work', 'parttime', 'fulltime'))
```

    [This command does appear in the chapter script file, and must be placed somewhere before the call to `multinom()`.] The object is to reorder the levels of the response variable so that `not.work` is the first, and hence baseline, level. Although it isn't strictly necessary to make `participation` an *ordered* factor, it is convenient to do so for the subsequent example (on the proportional-odds model).

13. Page 185: The sentence, "The gamma distribution requires a nonnegative, continuous response variable," is misleading, and should read, "Gamma GLMs require a positive, continuous response variable." Likewise, footnote 10, which currently begins, "The inverse-Gaussian family is also appropriate for continuous nonnegative data," should read, "The inverse-Gaussian family is also appropriate for continuous positive data."

14. Page 189: The text seems to imply that the `offset` function is unavailable in S-PLUS; although there is no `offset` *argument* to `glm` in S-PLUS, there is an `offset` *function* for use in model formulas.

15. Page 209: An error in the `hccm` function produced incorrect values in the heteroscedasticity-adjusted ANOVA table; the corrected table is

```
> Anova(mod.ornstein, white.adjust="hc3")
Anova Table (Type II tests)
Response: interlocks + 1
          Sum Sq  Df F value  Pr(>F)
assets      4496   1   46.56 7.5e-11
nation      3764   3   12.99 7.0e-08
sector      3377   9    3.89 0.00013
Residuals  22595 234
```

16. Page 213 (Box on `Ask` in S-PLUS): The "cautionary note" at the end of the box is inaccurate: Assigning in frame 1 *does not* assign to a global variable. The example is therefore unproblematic in S-PLUS.

17. Page 221: The text states that "`step` attempts to maximize the AIC"; as the subsequent discussion makes clear, the AIC is *minimized* not maximized.

18. Page 255 (Box on errors bars in S-PLUS): There is an `error.bar` function in S-PLUS. It is therefore unnecessary to use `segments` or to write one's own function.

19. Page 267: The `Matrix` library *is* available for S4 (though at present not under Windows).

20. Page 289: There are two errors in the code for the function `lreg.4`:

    (a) Immediately below the `while` loop, the condition should read
    `if(it > max.iter) warning('maximum iterations exceeded')`

    (b) In building the `result` object, the `coefficients` slot should be specified as `coefficients=as.vector(b)`.