Errata for *Regression Diagnostics, Second Edition* (Sage, 2020)

John Fox

updated: 2022-06-29

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

1. Page 47: In the following sentence,

   > Because we have picked the biggest of n test statistics, however, it is no longer legitimate simply to use t_{n-k-2} to find the p-value for t_m^*: For example, even if our model is wholly adequate, and disregarding for the moment the dependence among the e^*s, we would expect to observe about 5% of e^*s beyond $\pm t_{0.975} \approx \pm 2$, about 1% beyond $\pm t_{0.995} \approx \pm 2.6$, and so forth.

   the reference to t_m^* should be to e_m^* (i.e., the largest absolute studentized residual).

2. Page 48: The reference to "(deleted) estimates of the coefficient standard errors" [i.e., the $\mathrm{SE}_{(-i)}(b_j)$] in the denominator of d_{ij}^* is ambiguous. Think of the following as a footnote added immediately before the equation for d_{ij}^*:

   > To compute $\mathrm{SE}_{(-i)}(b_j)$, we'd ideally like to do the equivalent of removing the ith case from the data. The statistical software of which I'm aware, following Belsley, Kuh, and Welsch (1980), takes a simpler approximate approach, in effect replacing the residual standard deviation s in the formula for the coefficient standard error with the deleted version of this quantity, $s_{(-i)}$. Thus, if $\mathrm{SE}(b_j)$ is the standard error of b_j for the full data set (i.e., the square-root of $\widehat{V}(b_j)$ from Equation 2.2 on page 9), then
   >
   > $$\mathrm{SE}_{(-i)}(b_j) = \frac{s_{(-i)}}{s}\mathrm{SE}(b_j)$$
   >
   > Again referring to Equation 2.2, to compute an exact version of $\mathrm{SE}_{(-i)}(b_j)$, we would also have to do the equivalent of recomputing the variance of x_j and its multiple correlation R_j with the other xs after deleting the ith case.

   I'm grateful to Rachel Gordon for pointing out this ambiguity.

3. Page 75: The score statistic for the Breusch-Pagan test, given as $S_0^2 = \sum(\widehat{u}_i - \overline{u})/2$, should be $S_0^2 = \sum(\widehat{u}_i - \overline{u})^2/2$.

   I thank a participant in the May 2022 SORA-TABA workshop on regression diagnostics for alerting me to this error.