Implementation of `lm.beta`

Stefan Behrendt

January 01, 2023

The package `lm.beta` is based on equation (1) to estimate the standardized regression coefficients.

$$\hat{\beta}_i = \hat{b}_i \cdot \frac{s(X_i)}{s(Y)} \tag{1}$$

using

$$s(A) = \sqrt{\frac{\sum_j w_j \cdot (A_j - m(A) \cdot I)^2}{(n_w - 1)/n_w \cdot \sum_j w_j}}$$

$$m(A) = \frac{\sum_j w_j \cdot A_j}{\sum_j w_j}$$

with

- $\hat{\beta}_i$ the i-th standardized regression coefficient
- $\hat{b}_i$ the i-th unstandardized regression coefficient
- $I = \begin{cases} 0/1 & \text{for models without intercept*} \\ 1 & \text{for models with intercept} \end{cases}$
  - * argument `complete.standardization` chooses the factor: `complete.standardization = FALSE` ⇒ $I = 0$ / `complete.standardization = TRUE` ⇒ $I = 1$
  - * IBM® SPSS Statistics®, e.g., always uses $I = 0$ for models without intercept
  - * see e.g. `https://online.stat.psu.edu/~ajw13/stat501/SpecialTopics/Reg_thru_origin.pdf`[1] for further information on which I to choose
- Y the dependent variable
- X_i the i-th independent variable
- w the case weights
- n_w the number of non-zero weights

[1] Eisenhauer J.G. (2003). Regression through the Origin. *Teaching Statistics*, 25(3), p. 76-80.

A simplification for $I = 1$ is shown in equation (2) and for $I = 0$ in equation (3).

$$\hat{\beta}_i = \hat{b}_i \cdot \frac{s_{X_i}}{s_Y} \tag{2}$$

$$\hat{\beta}_i = \hat{b}_i \cdot \frac{\sigma_{X_i}}{\sigma_Y} \tag{3}$$

with (additionally to above)

- s_A the standard deviation of A (*)
- $\sigma_A = \sqrt{\sum_j A_j^2}$ an estimate of the uncentered second moment of A (*)

  * The sample size—and the different methods for correcting it—doesn't have to be considered when estimating the moments, because the factors would be similar in numerater and denominater, and therefore would be reduced.

Simplifications of non-weighted cases are

$$s(A) = \sqrt{\frac{\sum_j (A_j - m(A) \cdot I)^2}{n-1}}$$

$$m(A) = \frac{\sum_j A_j}{n}$$

with (additionally to above)

- n the number of non-empty cases