DESCRIPTIVES

Notation

The following notation is used throughout this chapter unless otherwise stated:

X_i	Value of the variable for case i
w_i	Weight for case i
N	Number of cases
W_i	Sum of the weights for the first i cases
$\overline{X}_i$	Mean for the first i cases

Moments

Moments about the mean are calculated recursively using a provisional means algorithm (Spicer, 1972):

$$W_j = \sum_{i=1}^{j} w_i$$

$$v_j = \frac{w_j}{W_j}\left(X_j - \overline{X}_{j-1}\right)$$

$$M_j^4 = M_{j-1}^4 - 4v_j M_{j-1}^3 + 6v_j^2 M_{j-1}^2 + \left(\frac{W_j^2 - 3w_j W_{j-1}}{w_j^3}\right) v_j^4 W_{j-1} W_j$$

$$M_j^3 = M_{j-1}^3 - 3v_j M_{j-1}^2 + \frac{W_j W_{j-1}}{w_j^2}\left(W_j - 2w_j\right) v_j^3$$

$$M_j^2 = M_{j-1}^2 + \frac{W_j W_{j-1}}{w_j} v_j^2$$

$$\overline{X}_j = \overline{X}_{j-1} + v_j$$

$$W_0 = \overline{X}_0 = M_0^2 = M_0^3 = M_0^4 = 0$$

After the last observation has been processed,

W_N = sum of weights for all cases

$\overline{X}_N$ = mean

$$M_N^r = \sum_{i=1}^{N} w_i \left(X_i - \overline{X} \right)^r$$

Basic Statistics

Mean

$$\overline{X}_N$$

Variance

$$S^2 = M_N^2 \Big/ (W_N - 1)$$

Standard Deviation

$$S = \sqrt{S^2}$$

Standard Error

$$S_{\overline{X}} = \frac{S}{\sqrt{W_N}}$$

Minimum

$$\min_j X_j$$

Maximum

$$\max_j X_j$$

Sum

$$\overline{X}_N W_N$$

Skewness and Standard Error of Skewness

$$g_1 = \frac{W_N M_N^3}{(W_N - 1)(W_N - 2)S^3} \qquad \text{se}(g_1) = \sqrt{\frac{6W_N(W_N - 1)}{(W_N - 2)(W_N + 1)(W_N + 3)}}$$

If $W_N \leq 2$ or $S^2 < 10^{-20}$, g_1 and its standard error are not calculated.

Kurtosis (Bliss, 1967, p. 144) and Standard Error of Kurtosis

$$g_2 = \frac{W_N(W_N + 1)M_N^4 - 3M_N^2 M_N^2(W_N - 1)}{(W_N - 1)(W_N - 2)(W_N - 3)S^4} \qquad \text{se}(g_2) = \sqrt{\frac{4(W_N^2 - 1)(SE(g_1))^2}{(W_N - 3)(W_N + 5)}}$$

If $W_N \leq 3$ or $S^2 < 10^{-20}$, g_2 and its standard error are not calculated.

Z-Scores

$$Z_i = \frac{X_i - \overline{X}_N}{S}$$

If X_i is missing or $S \leq 0$, Z_i is set to the system missing value.

References

Bliss, C. I. 1967. *Statistics in biology*, Volume 1. New York: McGraw-Hill.

Spicer, C. C. 1972. Algorithm AS 52: Calculation of power sums of deviations about the mean. *Applied Statistics*, 21: 226–227.