

Data Standardization

z-scores

$$z_{rj} = \frac{x_{rj} - \hat{\mu}_j}{\hat{\sigma}_j} \quad \begin{array}{l} r = 1, \dots, N \\ j = 1, \dots, p \end{array}$$

$$\mathbf{Z} = (z_{rj})$$

General Standardized Scores

$$w_{rj} = \frac{x_{rj} - l_j}{s_j}$$

where

l_j is a location measure for the j^{th} variable

s_j is a scale measure for the j^{th} variable

Standardization Methods available in PROC STDIZE

Method Keyword	Location (l)	Scale (s)
MEAN	mean	1
MEDIAN	median	1
STD	mean	st. dev.
RANGE	minimum	range
MIDRANGE	midrange =(min+max)/2	range/2
IQR	median	IQR
MAD	median	median abs. dev. from median