Algorithm for bootstrapping a distribution of $\alpha$

Klaus Krippendorff
kkrippendorff@asc.upenn.edu
2006.08.16
Revised 2011.6.12

In the absence of a theoretically motivated distribution for $\alpha$, and especially because reliability data may be small and have various metrics (levels of measurement), the distribution of $\alpha$ is obtained by bootstrapping. It provides probabilities of the $\alpha$-values that can be expected when very many similar samples of reliability data were coded. This bootstrapping algorithm randomly draws a great number of samples from the cell contents of a matrix of observed coincidences, obtains a hypothetical disagreement $D_o$ for each, which together with the original expected disagreement $D_e$, gives rise to a probability distribution, $p_\alpha$ of likely $\alpha$-values.

Given:

- The square matrix of observed coincidences $o_{ck}$, which gave rise to the $\alpha$ as calculated, including the total number $n_\cdot$ of values contributing to pair comparisons $n_\cdot = \sum_{c=1}^{v} \sum_{k=1}^{r} o_{ck}$
- The expected disagreement $D_e$ in the denominator of the observed $\alpha = 1 - \frac{D_o}{D_e}$
- The applicable metric difference $\delta_{ck}^2$
- The number $X$ of resamples to be drawn – chosen by the analyst.

The bootstrapping algorithm is defined in four steps:

First. Define the function $\delta_{ck}^2 = f(R)$ where $R$ is a uniformly distributed random number between 0 and 1 within a continuum of adequate precision. That continuum is segmented by the probabilities $p_{ck} = \frac{o_{ck}}{n_\cdot}; \sum_{c=1}^{v} \sum_{k=1}^{r} p_{ck} = 1$ so that each segment $p_{ck}$ of $R$ is associated with its corresponding $\delta_{ck}^2$:

$$R = 0 \quad \delta_{11}^2 \quad \sum_{g<h} \sum_{c=1}^{v} p_{gh} \quad p_{ck} \quad \sum_{g>h} \sum_{k=1}^{r} p_{gh} \quad 1 = R$$
Second. Determine the number \( M \) of random draws with replacement from the data, capped by a practical limit.

Let \( Q = \) the number of non-zero \( c \cdot k \) coincidences, \( o_{ck} > 0 \),

\[
M = \min[25 \cdot Q, (m-1)n../2]
\]

Third. Bootstrap the distribution of \( \alpha \)

Set the array \( n_\alpha = 0 \); where \(-1 \leq \alpha \leq +1\), and \( \alpha \) has at least 4 significant digits.

Do \( X \) times \( X \) is chosen by the analyst, by default \( X = 20,000 \)

\[
\text{Do } M \text{ times}
\]

\[
\begin{cases} 
\text{Pick a random number } R \text{ between 0 and 1 (uniform distribution)} \\
\text{Determine } \delta^2_{ck} \text{ by means of the function } f(R) \\
\text{SUM } \leftarrow \text{SUM } + \delta^2_{ck} \\
\alpha = 1 - \frac{\text{SUM}}{M \cdot D_e} \\
\text{If } \alpha < -1.000, n_{\alpha-1} \leftarrow n_{\alpha-1} + 1 \\
\text{Otherwise: } n_\alpha \leftarrow n_\alpha + 1 
\end{cases}
\]

Forth. Correct the frequencies \( n_\alpha \) for situations in which the lack of variation should cause \( \alpha \) to be indeterminate (\( \alpha = 1 - 0/0 \)):

\[
n_x = 0 \\
\text{If the matrix of coincidences contains exactly one non-zero diagonal cell: } o_{cc} > 0: \\
n_x = n_{\alpha=1} \quad \text{and} \quad n_{\alpha=1} = 0 \\
\text{If the matrix of coincidences contains two or more non-zero diagonal cells: } o_{cc} > 0: \\
n_x = X \sum_{c=1}^{c=1} \left( \frac{o_{cc}}{n_{cc}} \right)^M \quad \text{and} \quad n_{\alpha=1} \leftarrow n_{\alpha=1} - n_x
\]

The resulting distribution of \( \alpha \) is expressed in terms of the probabilities \( p_\alpha = \frac{n_\alpha}{X - n_x} \).

This distribution offers two important statistical properties of \( \alpha \):

- The confidence interval for \( \alpha \) at a chosen level \( p \) of statistical significance (two-tailed):
  \[
  \begin{align*}
  \alpha_{\text{smallest}} &= \text{the smallest } \alpha \mid \sum_\alpha \frac{n_\alpha}{X - n_x} \geq \frac{p}{2} \\
  \alpha_{\text{largest}} &= \text{the largest } \alpha \mid \sum_\alpha \frac{n_\alpha}{X - n_x} \leq \left(1 - \frac{p}{2}\right) \\
  \alpha_{\text{smallest}} \leq \alpha \leq \alpha_{\text{largest}}
  \end{align*}
  \]

- The probability \( q \) that the reliability data fail to reach the smallest acceptable \( \alpha_{\text{min}} \):
  \[
  q = \sum_{\alpha < \alpha_{\text{min}}} \frac{n_\alpha}{X - n_x}
  \]