Algorithm for bootstrapping a distribution of α

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

In the absence of a theoretically motivated distribution for α, and especially because reliability data may be small and have various metrics (levels of measurement), the **distribution of** α is obtained by **bootstrapping**. It provides probabilities of the α-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_α, of likely α-values.

Given:

- The square **matrix of observed coincidences** o_{ck}**,** which gave rise to the calculated α, including the total number ***n..*** of values contributing to pair comparisons $n.. = \sum_{c=1}^{v}\sum_{k=1}^{v} o_{ck}$
- The **number** ***m*** **of observers or coders** generating the reliability data
- The **expected disagreement** D_e in the denominator of the observed $\alpha = 1 - \dfrac{D_o}{D_e}$
- The applicable **metric difference** ${}_{metric}\delta_{ck}^2$
- The **number** ***X*** of resamples to be drawn – chosen by the analyst – ***X*** = 20,000 by default.

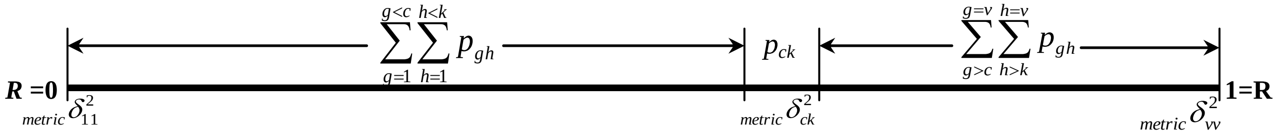
The bootstrapping algorithm is defined in four steps:

First. Define the function ${}_{metric}\delta_{ck}^2 = f(R)$ 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

$$p_{ck} = \frac{o_{ck}}{n..}; \quad \sum_{c=1}^{v}\sum_{k=1}^{v} p_{ck} = 1$$

so that each segment p_{ck} of ***R*** is associated with its corresponding ${}_{metric}\delta_{ck}^2$:



Second. Determine the number $\boldsymbol{M}$ of random draws with replacement from the data, capped by a practical limit.

Let Q = the number of non-zero $\boldsymbol{c}$-$\boldsymbol{k}$ coincidences, $\boldsymbol{o}_{ck} > 0$,

$$\boldsymbol{M} = \min[25\cdot\boldsymbol{Q}, (\boldsymbol{m}-1)\cdot\boldsymbol{n}_{..}/2]$$

Third. Bootstrap the distribution of $\boldsymbol{\alpha}$:

Set the array $\boldsymbol{n}_\alpha = 0$; where $-1 \leq \boldsymbol{\alpha} \leq +1$, and $\boldsymbol{\alpha}$ has at least 4 significant digits.

Do $\boldsymbol{X}$ times

- $\boldsymbol{SUM} = 0$
- Do $\boldsymbol{M}$ times
  - Pick a random number $\boldsymbol{R}$ between 0 and 1 (uniform distribution)
  - Find ${}_{metric}\delta^2_{ck} = \boldsymbol{f}(\boldsymbol{R})$
  - $\boldsymbol{SUM} \Leftarrow \boldsymbol{SUM} + {}_{metric}\delta^2_{ck}$
- $\boldsymbol{\alpha} = 1 - \dfrac{SUM}{M \cdot D_e}$
- If $\boldsymbol{\alpha} < -1.000$, $\boldsymbol{n}_{\alpha=-1} \Leftarrow \boldsymbol{n}_{\alpha=-1} + 1$
- Otherwise: $\boldsymbol{n}_\alpha \Leftarrow \boldsymbol{n}_\alpha + 1$

Forth. Correct the frequency $\boldsymbol{n}_{\alpha=1}$ for situations in which the lack of variation should cause $\boldsymbol{\alpha}$ to be indeterminate ($\boldsymbol{\alpha} = 1 - 0/0$):

$\boldsymbol{n}_x = 0$

If the matrix of coincidences contains **exactly one non-zero diagonal cell:** $\boldsymbol{o}_{cc} > \mathbf{0}$:

$\boldsymbol{n}_x = \boldsymbol{n}_{\alpha=1}$ and $\boldsymbol{n}_{\alpha=1} = 0$

If the matrix of coincidences contains **two or more non-zero diagonal cells:** $\boldsymbol{o}_{cc} > \mathbf{0}$:

$$\boldsymbol{n}_x = X\sum_{c=1}^{c=v}\left(\frac{o_{cc}}{n_{..}}\right)^M \quad \text{and} \quad \boldsymbol{n}_{\alpha=1} \Leftarrow \boldsymbol{n}_{\alpha=1} - \boldsymbol{n}_x$$

The resulting distribution of $\boldsymbol{\alpha}$ is expressed in terms of the probabilities $p_\alpha = \dfrac{n_\alpha}{X - n_x}$.

The resulting distribution offers **two** important **statistical properties of α:**

1. The **confidence interval** for $\boldsymbol{\alpha}$ at a chosen level $\boldsymbol{p}$ of statistical significance (two-tailed):

$$\alpha_{smallest} = \textit{the smallest } \alpha \,\Big|\, \sum_\alpha \frac{n_\alpha}{X - n_x} \geq \frac{p}{2}$$

$$\alpha_{largest} = \textit{the largest } \alpha \,\Big|\, \sum_\alpha \frac{n_\alpha}{X - n_x} \leq \left(1 - \frac{p}{2}\right)$$

$$-1 \leq \alpha_{smallest} \leq \alpha \leq \alpha_{largest} \leq 1$$

2. The **probability** $\boldsymbol{q}$ that the reliability data fail to reach the smallest acceptable $\boldsymbol{\alpha}_{min}$:

$$q = \sum_{\alpha < \alpha_{min}} \frac{n_\alpha}{X - n_x}$$