Krippendorff's alpha formulation in the calculation of Conformity Score 1

Supporting Material 7 to:
Kolumbo Volcano, Greece: expert elicitation findings supporting volcanic hazard and risk assessment

Andrea Bevilacqua[1✉] · Mattia de'Michieli Vitturi[1] · Alessandro Tadini[1] · Augusto Neri[1] · Willy P Aspinall[2,3] · R Stephen J Sparks[2]

[1] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, Pisa, Italy.

[2] School of Earth Sciences, University of Bristol, Bristol, UK.

[3] Aspinall & Associates, Tisbury, UK.

In this Supporting File we report the technical details of the algorithm used to calculate the Krippendorff's alpha in our specific application to expert elicitation. The description is based on the formulation provided by Charles Zaiontz on the website https://real-statistics.com/.

For every target item Q_n, n = 1, …, 64, we considered the agreement matrix $(R_{ij})_{i = 1,2,3;\, j=1, \dots, 10}$, where R_{ij} defined by counting the number of percentile i^{th} values falling in the j^{th} bin in which the intrinsic range is split. In this notation, percentile n.1 is the 5^{th}, percentile n.2 is the median, and percentile n.3 is the 95^{th}. The following example is related to target item Q1 and the m = 13 experts that responded in both elicitations:

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10
5%ile	2	7	2	2	0	0	0	0	0	0
median	0	1	3	4	4	1	0	0	0	0
95%ile	0	0	1	2	3	2	0	2	0	3

We weighted the bin pairs following an interval scale, based on their relative distance:

$$w(h,k) = 1- [(h - k) / 9]^2.$$

Note that $0 \leq w(h,k) \leq 1$; $w(h,h) = 1$; for each h,k in 1, …, 10; $w(1,10) = w(10,1) = 0$.

Krippendorff alpha is defined as:

$$\alpha = 1 - (1 - p_a) / (1 - p_e),$$

where $(1 - p_a)$ is called the Observed Disagreement, and $(1 - p_e)$ the Expected Disagreement (Krippendorff, 1970; 2004). In this formula, the values p_a and p_e are obtained from the agreement matrix (R_{ij}) and the weights w(h,k), as a function of the number m of experts:

$$p_a = C_1 \cdot \Sigma_i \{\Sigma_k R_{ik} \cdot [(\Sigma_h w(h,k) \cdot R_{ih}) -1] \} + C_2,$$

$$p_e = (C_2)^2 \cdot \Sigma_{h,k} [w(h,k) \cdot (\Sigma_i R_{ih}) \cdot (\Sigma_i R_{ik})],$$

where $C_1 \cong (3 \cdot m - 1) / [9 \cdot m^2 \cdot (m - 1)]$ and $C_2 \cong 1 / (3 \cdot m)$.

Note that the Classical Model elicitation procedure requires the triplets of percentile judgments to be given as ascending values (Cooke 1991). Therefore, the 5^{th} percentile judgments generally cluster in lower value bins than those for the 50^{th} percentile judgements which, in turn, typically occupy bins generally lower than those for the 95^{th} percentile judgments. This situation generally produces a greater agreement than a random distribution of values in the ten bins, and therefore our α values are mostly positive (see Supporting Material S7).

The standard deviation of the Krippendorff's alpha is the square root of the variance calculated as follows, as a function of the number m of experts:

$\mathrm{Var}[\alpha] = C_3 \cdot \Sigma_i \left\{(p_i - p_e) / (1 - p_e) - \left[2 \cdot (1 - \alpha) \cdot (e_i - p_e)\right] / (1 - p_e) - \alpha \right\}^2$,

where:

$p_i = C_4 \cdot \Sigma_k \left\{R_{ik} \cdot \left[(\Sigma_h w(h,k) \cdot R_{ih}) -1\right]\right\}$,

$e_i = C_5 \cdot \Sigma_k \left\{\Sigma_h \left[w(h,k) \cdot (\Sigma_i R_{ih})\right] \cdot R_{ik}\right\}$,

and $C_3 \cong 1/6$, $C_4 \cong 1 / [m^2 \cdot (m - 1)]$, $C_4 \cong 1 / (3 \cdot m^2)$.