Appendix to Content Explorer: Recommending Novel Entities for a Document Writer

Michal Lukasik
Google AI
mlukasik@google.com

Richard Zens
Google AI
zens@google.com

1 Intraclass Coefficient

Let us consider the following random effects model: $Y_{ijk} = \mu + a_i + d_{ij} + r_k + v_{ik} + e_{ijk}$. Here, Y_{ijk} corresponds to the score for suggestion (entity) j recommended to item (document) i by rater k, μ corresponds to the average score for the entire annotation, a_i corresponds the item specific effect, d_{ij} corresponds to the suggestion-item specific effect, r_k corresponds to the reviewer specific effect, v_{ik} corresponds to the item-reviewer specific effect, and e_{ijk} is the residual. Then, the Intraclass Coefficient for consistency under k raters takes the following form:

$$ICC(C,k) = \frac{v}{v + (1 - v - r)/3}, \qquad (1)$$

where $v = Var(d_{ij}) + Var(a_i)$ is the variance explained by the underlying score, and $r = Var(r_k) + Var(v_{ik})$ is the variance corresponding to preferences from reviewers which are discounted (McGraw and Wong, 1996).

References

K. O. McGraw and S. P. Wong. 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1):30–46.