Optimal Data Analysis
Vol. 4 (November 28, 2015), 135-136

2155-0182/10/$3.00

UniODA *vs.* Cochran's *Q* Test for Related Proportions: Measures of Effect Size

Paul R. Yarnold, Ph.D.
Optimal Data Analysis, LLC

Maximum-corrected (η^2_Q) and chance-corrected ($\mathcal{R}$ and *ESS*) measures of effect size for the *Q* test are compared.

Berry, et al.[1] present a hypothetical example having $c = 2$ treatments (time 1 and time 2) and $n = 10$ subjects: at each time each subject voted "Pro" or "Con" (Table 1). Note this design has both inter-rater (data from multiple raters are pooled) and temporal (the tabled outcome is the result of a temporal comparison) characteristics.

Table 1: Hypothetical 2 x 2 Data Table

	Time 2	
Time 1	Pro	Con
Pro	5	2
Con	0	3

For these data the maximum-corrected measure of effect size that standardizes *Q* by its theoretical maximum value (Q_{max})—a function of *n* and *c*—is $\eta^2_Q = 0.20$.[1] This erroneous underestimate occurs because Q_{max} cannot be obtained in this example.[1]

For these data the chance-corrected measure of effect size—an exact measure derived for *Q*—is $\mathcal{R} = 1.0$, indicating perfect inter-rater agreement.[1]

While results in Table 1 are consistent with the definition of agreement by the *Q* statistic, they are inconsistent with the conceptualization of agreement as defined in the ODA paradigm.[2,3] In ODA the effect size corresponding to observed inter-rater or inter-device agreement, *ESS*, "…is zero under chance conditions, unity when agreement among the *n* subjects (raters, devices) is perfect, and negative under conditions of disagreement".[1] These conditions correspond to *ESS* = 0 (chance), *ESS* = 100 (perfect agreement) and *ESS* = -100 (perfect disagreement).[2-7] The UniODA model assessing level of agreement was obtained using the following UniODA[2] and MegaODA[8-10] software syntax:

```
OPEN DATA;
OUTPUT qtest.out;
CATEGORICAL ON;
TABLE 2;
DIRECTIONAL < 1 2;
MCARLO ITER 25000;
DATA;
5 2
0 3
END DATA;
GO;
```

Here the chance-corrected measure of inter-rater agreement—an exact measure of agreement as defined above—is *ESS* = 0.60: a relatively strong effect.[2,3]

2155-0182/10/$3.00

References

[1]Berry KJ, Johnston JE, Mielke PW (2007). An alternative measure of effect size for Cochran's *Q* test for related proportions. *Perceptual and Motor Skills*, *104*, 1236-1242. DOI: 10.2466/PMS.104.3.1236-1242

[2]Yarnold PR, Soltysik RC (2005). *Optimal data analysis: A guidebook with software for Windows*. Washington, DC: APA Books.

[3]Yarnold PR, Soltysik RC (In Review). *Maximizing predictive accuracy*. Chicago, IL: ODA Books.

[4]Yarnold PR (2014). UniODA *vs*. kappa: Evaluating the long-term (27-year) test-retest reliability of the Type A Behavior Pattern. *Optimal Data Analysis*, *3*, 14-16. URL: http://optimalprediction.com/files/pdf/V3A5.pdf

[5]Yarnold PR (2014). How to assess inter-observer reliability of ratings made on ordinal scales: Evaluating and comparing the Emergency Severity Index (Version 3) and Canadian Triage Acuity Scale. *Optimal Data Analysis*, *3*, 42-49. URL: http://optimalprediction.com/files/pdf/V3A15.pdf

[6]Yarnold, PR (2014). How to assess the inter-method (parallel-forms) reliability of ratings made on ordinal scales: Evaluating and comparing the Emergency Severity Index (Version 3) and Canadian Triage Acuity Scale. *Optimal Data Analysis*, *3*, 50-54. URL: http://optimalprediction.com/files/pdf/V3A16.pdf

[7]Yarnold PR (2015). Estimating inter-rater reliability using pooled data induces paradoxical confounding: An example involving Emergency Severity Index triage ratings. *Optimal Data Analysis*, *4*, 21-23. URL: http://optimalprediction.com/files/pdf/V4A6.pdf

[8]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Separating the chaff. *Optimal Data Analysis*, 2, 194-197. URL: http://optimalprediction.com/files/pdf/V2A29.pdf

[9]Soltysik RC, Yarnold PR (2013). MegaODA large sample and BIG DATA time trials: Harvesting the Wheat. *Optimal Data Analysis*, 2, 202-205. URL: http://optimalprediction.com/files/pdf/V2A31.pdf

[10]Yarnold PR, Soltysik RC (2013). MegaODA large sample and BIG DATA time trials: Maximum velocity analysis. *Optimal Data Analysis*, 2, 220-221. URL: http://optimalprediction.com/files/pdf/V2A35.pdf

[11]Yarnold PR, Soltysik RC (2014). Discrete 95% confidence intervals for ODA model- and chance-based classifications. *Optimal Data Analysis*, *3*, 110-112. URL: http://optimalprediction.com/files/pdf/V3A26.pdf

Author Notes

The study analyzed artificial data and was exempt from Institutional Review Board review. No conflict of interest was reported.

Mail: Optimal Data Analysis, LLC
6348 N. Milwaukee Ave., #163
Chicago, IL 60646
USA