The Technical Whitepaper Series

7

KR20 & Coefficient Alpha

Their equivalence for binary-scored items

pbarrett.net
Advanced Projects R&D

June, 2007

Internal Consistency Reliability for Dichotomous Items – KR-20 & Alpha

There is apparent confusion within some individuals concerning the calculation of internal consistency reliability for dichotomous response, items whose cumulative sum form the scale score for a test. I've heard some individuals tell others that they MUST use a Kuder-Richardson KR-20 coefficient instead of a Cronbach alpha.

This is completely incorrect – in that the KR-20 is mathematically equivalent to the formula for coefficient alpha! What seems to be confusing these individuals is that the formula for the variance of a binary item variable looks quite different to that of a continuous-valued variable.

The KR-20 formula is:

$$KR_{20} = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} p_i q_i}{\sigma_{test}^2}\right) \qquad \text{where:}$$

$k =$ number of items in the test

$p_i =$ the proportion of respondents answering an item i in the keyed direction [1]

$q_i =$ the proportion of respondents answering an item i in the non-keyed direction [0]

$\sigma_{test}^2 =$ the test-score variance

The Cronbach alpha formula is:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_{test}^2}\right) \qquad \text{where:}$$

$k =$ number of items in the test

$\sigma_i^2 =$ the variance of item i

$\sigma_{test}^2 =$ the test-score variance

The Equivalency

$$\sum_{i=1}^{k} p_i q_i = \sum_{i=1}^{k} \sigma_i^2$$

Note – we are using population variance formulae here – hence the use of σ^2

A Worked Example

6 items and 10 respondents

	KR-20 and alpha test spreadsheet						
	1 item1	2 item2	3 item3	4 item4	5 item5	6 item6	7 Test Score
1	0	0	0	0	0	0	0
2	0	0	0	1	0	0	1
3	1	0	1	0	1	0	3
4	1	1	1	0	0	1	4
5	1	1	0	1	0	1	4
6	0	0	1	0	0	0	1
7	0	0	0	0	1	0	1
8	0	0	0	1	0	0	1
9	0	0	0	0	1	0	1
10	0	1	0	0	0	1	2

For every item, p_i = 0.3, q_i = 0.7 = 0.21, so $\sum_{i=1}^{k} p_i q_i = 10(0.3 \cdot 0.7) = 1.26$

The Individual item variances – calculated using the population variance rather than sample variance formula …

$$\sigma_i^2 = \frac{\sum_{j=1}^{n} (x_{ij} - \bar{x})^2}{n} \quad \text{where}$$

$$n = \text{ number of respondents}$$

$$x_i = \text{ score in item } i \text{ for person } j$$

note the division in the equation by *n* and not (*n*-1) – remember, it is the population variance which is computed in the KR-20 and alpha formulae

The individual item means are all equal to 0.3, with variances equal to 0.21, and

$$\sum_{i=1}^{k} \sigma_i^2 = 6 \cdot 0.21 = 1.26$$

The total test score variance is: **1.76**

Thus, using the formula for either the KR-20 or alpha …

$$\alpha \text{ or } KR_{20} = \frac{k}{k-1}\left(1-\frac{\sum_{i=1}^{k}\sigma_i^2}{\sigma_{test}^2}\right) = \frac{6}{6-1}\left(1-\frac{1.26}{1.76}\right) = 0.340909$$

So far so good … until you use STATISTICA or SPSS … where the item and test score variances are not the same as computed above, and both report coefficient alpha.

STATISTICA v.8 : Items descriptives

	Means and Standard Deviations (KR-20 & Alpha data.sta)		
variable	Mean	Variance	Std.Dev.
item1	0.300000	0.233333333	0.483046
item2	0.300000	0.233333333	0.483046
item3	0.300000	0.233333333	0.483046
item4	0.300000	0.233333333	0.483046
item5	0.300000	0.233333333	0.483046
item6	0.300000	0.233333333	0.483046

But, as noted on page 2, using the variance calculated for a binary variable, every items variance was $p_i = 0.3$, $q_i = 0.7$ = **0.21**, with the sum as **1.26** – rather than this new sum of (6*2.333333 = **1.4**)

Then, look at the test score variance …

Reliability Results: KR-20 & Alpha data.sta

```
          Number of items in scale: 6

             Number of valid cases: 10
 Number of cases with missing data: 0
         Missing data were deleted: casewise

                    SUMMARY STATISTICS FOR SCALE
              Mean:  1.800000000                 Sum: 18.000000000
Standard Deviation:  1.398411798            Variance:  1.955555556
          Skewness:   .743538298            Kurtosis:  -.830965909
           Minimum:  0.000000000             Maximum:  4.000000000
  Cronbach's alpha:   .340909091   Standardized alpha:   .340909091
                      Average Inter-Item Correlation:   .462120251
```

This is **1.955555556** – whereas above we computed it as **1.76**.

Putting these values into the formula -

$$\alpha \text{ or } KR_{20} = \frac{k}{k-1}\left(1-\frac{\sum_{i=1}^{k}\sigma_i^2}{\sigma_{test}^2}\right) = \frac{6}{6-1}\left(1-\frac{1.4}{1.955555556}\right) = 0.340909$$

So, the same alpha is achieved. Why, because the item variances and test score variances are both computed using the sample rather than population formulae divisor of *n-1* rather than *n*. So, the relative ratio between these two quantities is exactly the same. The sum of item variances using either formula:

$$\sum_{i=1}^{k} p_i q_i = \sum_{i=1}^{k}\sigma_i^2 = 1.26$$

$$\sum_{i=1}^{k}\left(p_i q_i \cdot \left(\frac{n}{n-1}\right)\right) = \sum_{i=1}^{k}\left((0.3\cdot 0.7)\cdot\left(\frac{10}{9}\right) = 0.233333\right) = \sum_{i=1}^{k} s_i^2 = 1.4$$

The calculation of the item variances using the usual sum of squared deviations formula is:

$$\sigma_i^2 = \frac{\sum_{j=1}^{n}(x_{ij}-\overline{x})^2}{n} \qquad \text{where}$$

$n =$ number of respondents

$x_i =$ score in item i for person j

$$\sigma_i^2 = \frac{\sum_{j=1}^{n}(x_{ij}-\overline{x})^2}{n} = \frac{2.1}{10} = 0.21 \qquad \text{(population variance)}$$

$$s_i^2 = \frac{\sum_{j=1}^{n}(x_{ij}-\overline{x})^2}{n-1} = \frac{2.1}{9} = 0.23333333 \qquad \text{(sample variance)}$$

The calculation of the text score variance is :

$$\sigma_{test}^2 = \frac{\sum_{j=1}^{n}(x_j - \overline{x})^2}{n} \qquad \text{where}$$

$n =$ number of respondents

$x_j =$ test score for person j

$$\sigma_{test}^2 = \frac{\sum_{j=1}^{n}(x_j - \overline{x})^2}{n} = \frac{17.6}{10} = 1.76 \qquad \text{(population variance)}$$

$$s_i^2 = \frac{\sum_{j=1}^{n}(x_j - \overline{x})^2}{n-1} = \frac{17.6}{9} = 1.95555556 \qquad \text{(sample variance)}$$

The ratios between the respective sample or population variances are equal:

$$\left[\frac{\sum_{i=1}^{k}\sigma_i^2}{\sum_{j=1}^{n}\sigma_j^2}\right] \equiv \left[\frac{\sum_{i=1}^{k}s_i^2}{\sum_{j=1}^{n}s_j^2}\right] \ldots \left[\frac{1.26}{1.76}\right] = \left[\frac{1.4}{1.9555556}\right] \ldots \left[0.715909\right]$$

In SPSS 15 – the same sample statistics are computed as in STATISTICA :

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.340909	.340909	6

Item Statistics

	Mean	Std. Deviation	N
item1	.300000	.483046	10
item2	.300000	.483046	10
item3	.300000	.483046	10
item4	.300000	.483046	10
item5	.300000	.483046	10
item6	.300000	.483046	10

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	.300	.300	.300	.000	1.000	.000	6
Item Variances	.233	.233	.233	.000	1.000	.000	6

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
1.800000000	1.955556	1.398411798	6

Interestingly, Nunnally, J.C. & Bernstein, I.H. (1994) *Psychometric Theory 3rd. Edition*. McGraw-Hill. ISBN: 0-0704-7849-X., p. 234 and 235 define alpha and KR-20 in terms of population variances.

So do Crocker, L. & Algina, J. (1986) *Introduction to Classical and Modern Test Theory*. Harcourt Brace Jovanovich Publishers. ISBN: 0-03-061634-4., p. 138 and 139.

Anyway, I hope the above puts to rest the old chestnut that KR-20 must be calculated for dichotomous response items and alpha for Likert/continuous valued response items. KR-20 was always just a convenient way of simplifying the calculations of reliability for binary-response items. That's all.

Likewise that other old chestnut – you must use a point-biserial correlation when one variable is dichotomous and the other continuous. Take a look at a nice statement from Denis Roberts (1992) http://www.personal.psu.edu/users/d/m/dmr/papers/formula.PDF

David Howell in his series of textbooks "*Statistical Methods for Psychology*" (now 6th Edition, 2006, Wasdworth Publishing, ISBN: 0495012874) has always stated and demonstrated this. Yet, still I hear some people talk about a "point-biserial" correlation. Why? The point-biserial **is** a Pearson correlation.

End of story