The Relationship Between Spearman Brown and Cronbach's Alpha

Steven W. Nydick

May 25, 2012

In Chapter 6 of Lattin, Carroll, and Green (2003), the authors claim that

$$\alpha = \frac{k\bar{r}}{1 + (k-1)\bar{r}} \tag{1}$$

where $\bar{r}$ designates the *average* inter-item correlation. Several professors and graduate students are under the impression that Equation (1) (listed on p. 188 of Lattin et al., 2003) is misnamed. The left-hand side of Equation (1) designates Cronbach's alpha (Cronbach, 1951), whereas the right-hand side of Equation (1) is the Spearman-Brown prediction formula (Brown, 1910; Spearman, 1910). However, Lattin et al. (2003) set the item variances to 1 so that the results could be interpreted in terms of standardization (and relationships as correlations).

An assumption slightly less strict than variances of 1.0 is equal variances across all items. As long as "equal item variances" is assumed, then

$$\bar{r} = \frac{\sum\limits_{j \neq i}\sum\limits_{i} r_{ij}}{k(k-1)} = \frac{\sum\limits_{j \neq i}\sum\limits_{i} \frac{\sigma_{ij}}{\sigma_i \sigma_j}}{k(k-1)} = \frac{\sum\limits_{j \neq i}\sum\limits_{i} \frac{\sigma_{ij}}{\sigma_i \sigma_i}}{k(k-1)} = \frac{\sum\limits_{j \neq i}\sum\limits_{i} \frac{\sigma_{ij}}{\sigma_i^2}}{k(k-1)} = \frac{\sum\limits_{j \neq i}\sum\limits_{i} \sigma_{ij}}{k(k-1)\sigma_i^2} \tag{2}$$

In Equation (2), part two is equal to part three due to equal item variances, and part four is equal to part five because that item variance is a constant across all items.

Finally, inserting Equation (2) into the typical Spearman-Brown formula, we have

$$
\begin{aligned}
\text{SB} = \frac{k\bar{r}}{1+(k-1)\bar{r}} &= \frac{k\left[\frac{\sum\sum\sigma_{ij}}{k(k-1)\sigma_i^2}\right]}{1+(k-1)\left[\frac{\sum\sum\sigma_{ij}}{k(k-1)\sigma_i^2}\right]} \\
&= \frac{\left[\frac{\sum\sum\sigma_{ij}}{(k-1)\sigma_i^2}\right]}{1+\left[\frac{\sum\sum\sigma_{ij}}{k\sigma_i^2}\right]} \\
&= \frac{\left[\frac{\sum\sum\sigma_{ij}}{(k-1)\sigma_i^2}\right]}{\left[\frac{k\sigma_i^2+\sum\sum\sigma_{ij}}{k\sigma_i^2}\right]} \\
&= \frac{k\sigma_i^2\sum\sum\sigma_{ij}}{(k-1)\sigma_i^2(k\sigma_i^2+\sum\sum\sigma_{ij})} \\
&= \frac{k}{k-1}\left[\frac{\sum\sum\sigma_{ij}}{k\sigma_i^2+\sum\sum\sigma_{ij}}\right] \\
&= \frac{k}{k-1}\left[\frac{\sigma_C^2-\sum\sigma_i^2}{\sigma_C^2}\right] = \frac{k}{k-1}\left[1-\frac{\sum\sigma_i^2}{\sigma_C^2}\right] = \alpha \qquad (3)
\end{aligned}
$$

Because Equation (3) is the formula for coefficient alpha, both formulas are equivalent under the equal-variances assumption.

The following code demonstrates that the Spearman-Brown formula and coefficient alpha are identical under equal item variances and using average inner item correlations.

```
> set.seed(789234)
> n <- 100       # people
> k <- 10        # items
> t.var1 <- 0.5  # true-score variance (people)
> t.var2 <- 0.2  # true-score variance (item)
> e.var  <- 1.0  # error variance
> # Average true-score for each person/item:
> t.ppl  <- rnorm(n = n, mean = 0, sd = sqrt(t.var1))
> t.it   <- rnorm(n = k, mean = 0, sd = sqrt(t.var2))
> # Average true-score for each person on each item:
> t.mat  <- t( sapply( t.ppl,
+                      FUN = function(x) x + t.it ) )
> # The observed scores of each person on each item:
> x.mat  <- scale( t.mat + rnorm(n*k, 0, sqrt(e.var)) )
> # NOTE: ITEMS ALL HAVE SAME VARIANCE DUE TO scale #
>
> ##################
> # Spearman-Brown #
> ##################
>   cor.mat <- cor(x.mat)
>   cor.vec <- cor.mat[lower.tri(cor.mat, diag = FALSE)]
>   avg.cor <- mean(cor.vec)
```

```
>   SB      <- k*(avg.cor)/(1 + (k - 1)*avg.cor)
> #####################
> # Coefficient Alpha #
> #####################
>   item.var <- sum( apply(x.mat, MARGIN = 2, FUN = var) )
>   test.var <- var( rowSums(x.mat) )
>   alpha    <- ( k/(k - 1) )*(1 - item.var/test.var)
> # Are they the same?
>   SB
[1] 0.8610683
>   alpha
[1] 0.8610683
>
> # Yup!
```

Note the Spearman-Brown formula and coefficient alpha resulted in the same number only because the items all had the same variance (due to the `scale` function). Without scaling the items to have the same variance, the formulas result in slightly different numbers.

```
> set.seed(9123)
> n <- 100      # people
> k <- 10       # items
> t.var1 <- 0.5  # true-score variance (people)
> t.var2 <- 0.2  # true-score variance (item)
> e.var  <- 1.0  # error variance
> # Average true-score for each person/item:
> t.ppl  <- rnorm(n = n, mean = 0, sd = sqrt(t.var1))
> t.it   <- rnorm(n = k, mean = 0, sd = sqrt(t.var2))
> # Average true-score for each person on each item:
> t.mat  <- t( sapply( t.ppl,
+                      FUN = function(x) x + t.it ) )
> # The observed scores of each person on each item:
> x.mat  <- t.mat + rnorm(n*k, 0, sqrt(e.var))
> # NOTE: DO NOT ALL HAVE SAME VARIANCE DUE TO scale #
>
> ##################
> # Spearman-Brown #
> ##################
>   cor.mat <- cor(x.mat)
>   cor.vec <- cor.mat[lower.tri(cor.mat, diag = FALSE)]
>   avg.cor <- mean(cor.vec)
>   SB      <- k*(avg.cor)/(1 + (k - 1)*avg.cor)
> #####################
> # Coefficient Alpha #
> #####################
>   item.var <- sum( apply(x.mat, MARGIN = 2, FUN = var) )
```

```
>   test.var <- var( rowSums(x.mat) )
>   alpha    <- ( k/(k - 1) )*(1 - item.var/test.var)
> # Are they the same?
>   SB
[1] 0.8151104
>   alpha
[1] 0.8146178
>
> # No!  Weird!
```

References

[1] Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, *3*, 296–322.

[2] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

[3] Lattin, J., Carroll, J. D., & Green, P. E. (2003). *Analyzing multivariate data*. Pacific Grove, CA: Brooks/Cole.

[4] Spearman, Charles, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, *3*, 271–295.