Supplement to “Statistical methods research done as science rather than mathematics”

James S. Hodges
Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota USA 55414
email: hodge003@umn.edu
May 18, 2019

The predictor of \( \hat{\rho} = -1 \), reproduced from the main paper’s equation (10), is

\[
\left( \frac{Ns - N - 1}{(1 + r/s)(1 + r/q)} \right) \left[ 1 - \left( \frac{Ns - 2}{Ns - N - 1} \right) \frac{1 - \frac{N-1}{N(s-2)} \rho}{1 + \frac{2(N-1)(1+r/s)(1+r/q)}{N(s-2)(1+r/s)(1+r/q)-1}} \right].
\]

(1)

1 The predictor of \( \hat{\rho} = -1 \) is positive

In the expression for the predictor in equation (1) above,

\[
\frac{1 - \frac{N-1}{N(s-2)} \rho}{1 + \frac{2(N-1)(1+r/s)(1+r/q)}{N(s-2)(1+r/s)(1+r/q)-1}} = \frac{N(s-2) - (N-1) \rho}{N(s-2) + 2(N-1)(1+r/s)(1+r/q)} \cdot \frac{1}{1 + \frac{2(N-1)(1+r/s)(1+r/q)}{N(s-2)(1+r/s)(1+r/q)-1}}
\]

\[
< \frac{N(s-2) + (N-1)}{N(s-2) + 2(N-1)} = \frac{Ns - N - 1}{Ns - 2}
\]

for all finite positive \( N \), \( s \), and \( r \) and \( \rho \in (-1, 1) \). It follows immediately that the predictor is positive for these \( N \), \( s \), \( r \) and \( \rho \).

2 The predictor of \( \hat{\rho} = +1 \)

The predictor for \( \hat{\rho} = +1 \) is derived by the same sequence of steps used to derive the predictor for \( \hat{\rho} = -1 \), which gives

\[
- \left( \frac{Ns - N - 1}{(1 + r/s)(1 + r/q)} \right) \left[ 1 - \left( \frac{Ns - 2}{Ns - N - 1} \right) \frac{1 + \frac{N-1}{N(s-2)} \rho}{1 + \frac{2(N-1)(1+r/s)(1+r/q)}{N(s-2)(1+r/s)(1+r/q)-1}} \right].
\]

(3)

Comparison to equation (1) above shows that equation (3) above differs because of three sign changes: a minus sign at the far left of equation (3) and changes in the signs of the summands including \( \rho \) in the numerator and denominator of the complicated fraction at the far right of equation (3).

In the main paper’s tables, to make this predictor more similar to the predictor for \( \hat{\rho} = -1 \), the minus sign at the far left of equation (3) has been omitted.
3 Three bullets early in the main paper’s Section 3

The three bullets are:

- Given \( N, s, \) and \( \rho \), as \( r \) increases — i.e., as the error variance \( \sigma^2_e \) increases relative to the random-effect variance \( \sigma^2_r \) — the predictor goes to zero.
- Given \( \rho \) and \( r \), as either \( N \) or \( s \) increases, the predictor goes to infinity.
- Given \( N, s, \) and \( r \), as \( \rho \) goes to \(-1\), the predictor goes to zero.

3.1 Increasing \( r \)

In equation (1) above, as \( r \) increases,

\[
\frac{(1 + r/s)(1 + r/q) + \rho}{(1 + r/s)(1 + r/q) - 1} \to 1, \tag{4}
\]

so the expression in square brackets goes to a finite positive number for given legal values of \( N, s, \) and \( \rho \). However, as \( r \) increases, the expression in round brackets at the far left of equation (1) goes to 0, so the predictor goes to zero.

3.2 Increasing \( N \)

In equation (1) above, inside the square brackets are three fractions involving \( N \), and as \( N \) increases, each such fraction goes to a finite, positive number:

\[
\frac{Ns - 2}{Ns - N - 1} \to \frac{s}{s - 1}; \quad \frac{N - 1}{N(s - 2)} \to \frac{1}{s - 2}; \quad \frac{2(N - 1)}{N(s - 2)} \to \frac{2}{s - 2}. \tag{5}
\]

Therefore, the expression in square brackets goes to a finite, positive number. However, as \( N \) increases, the expression in round brackets at the far left of equation (1) increases without bound, so the predictor does as well.

3.3 Increasing \( s \)

Because \( s = 2m + 1 \) and \( q = (2m^2 + 3m + 1)/3m \), as \( s \) increases, so does \( q \). As \( s \) increases,

\[
N(s - 2)[(1 + r/s)(1 + r/q) - 1] = \frac{N(s - 2)}{s}[(s + r)(1 + r/q) - s] \to Nr, \tag{6}
\]

so the expression in square brackets in equation (1) goes to a finite, positive number. However, as \( s \) increases, the expression in round brackets at the far left of equation (1) increases without bound, so the predictor does as well.
3.4 Letting $\rho$ go to $-1$

As $\rho \to -1$, 

$$
\frac{1 - \frac{N-1}{N(s-2)}\rho}{1 + \frac{2(N-1)(1+r/s)(1+r/q)+\rho}{N(s-2)(1+r/s)(1+r/q)-1}} \to \frac{1 + \frac{N-1}{N(s-2)}}{1 + \frac{2(N-1)}{N(s-2)}} = \frac{Ns - N - 1}{Ns - 2}.
$$

(7)

Thus the expression in square brackets in equation (1) goes to zero and so does the predictor.

4 Section 5’s results for $\hat{\rho} = \pm 1$ and $\hat{\rho} = \text{NaN}$

Figures 1 and 2 below are analogous to the main paper’s Figure 2, showing the percent of simulated datasets giving $\hat{\rho} = \pm 1$ and $\hat{\rho} = \text{NaN}$ respectively as a function of $r$ and each of the other factors. Note that the vertical scales in these plots are 0 to 60% and 0 to 40%, while Figure 2’s vertical scale is 0 to 100%. The Monte Carlo standard error for each plotted percent is less than 0.8 percentage points.
Figure 1: Percent datasets giving $\hat{\rho} = \pm 1$ as a function of $\log_{10} r$. 

**Percent +/-1 estimates as a function of $r$, for different $N$**

**Percent +/-1 estimates as a function of $r$, for different $s$**

**Percent +/-1 estimates as a function of $r$, for different rho**
Figure 2: Percent datasets giving $\hat{\rho} = \text{NaN}$ as a function of $\log_{10} r$. 