By Jensen’s Inequality: , so regardless of the correlation, we have , with equality only when the number of matched samples is very large (I think). The observed test statistic therefore has more weight in the tails and for a given critical value , , with equality only in the limit (with m).

By using the 20th quantile estimator, we set . Equivalently, we set . In other words, we force our inflation factor to be conservative, allowing an “inflated” estimate only 20% of the time.

Alternatively, we can try to correct for this inflation factor directly. Let . Under the null hypothesis, note that . Let so that , where is the Fisher transformed correlation and . Let refer to and respectively.

Since , we can write:

We now create a scaled test statistic which “corrects” the modified t-test statistic to have the same variance as

We can solve for in the denominator:

Note that, since is normal, is log-normal, and

So, we have that

Plugging this into our expression for ,

Finally, plugging in our observed value for , we have

Under the null hypothesis, and , so we expect to be calibrated better than which has much inflated variance relative to . In fact, since is a scalar multiple of a distribution, we would say that has a non-standardized t distribution, with the scaling parameter estimated as:

This relationship is confirmed in the inf\_factor.R script. We should compare inferential results from the scaled t-distribution with alternative estimators for across different sample sizes.