On the Nonadditivity of Correlation Coefficients

Part 2: Fisher Transformations

Abstract – This is Part 2 of a tutorial series on the nonadditivity of correlation coefficients. This time we discuss Fisher *r*-to-*Z* and *Z*-to-*r* transformations and the risks of arbitrarily implementing these.

Keywords: nonadditivity, correlations, coefficients, fisher transformations, bivariate normality, pearson, spearman
Published: 01-07-2011; Updated: 09-18-2017
© E. Garcia, PhD; admin@minerazzi.com

Note: This article is part of a legacy series that the author published circa 2011 at http://www.miislita.com, now a search engine site. It is republished in pdf format here at http://www.minerazzi.com , with its content edited and updated.

Introduction

In Part 1 of this series we discussed why it is not possible to compute arithmetic averages from Pearson and Spearman correlation coefficients (Garcia, 2017a). Over the years, several workarounds in the form of weighted averages have been proposed, being Fisher Transformations one of them (Garcia, 2012).

The purpose of this article is to revisit said transformations and demonstrate the risk of arbitrarily implementing them. The following conventions are used:

- k is the number of samples.
- n is the sample size or number of paired observations (x, y) in a sample.
- x is an independent variable and $\bar{x}$ its mean.
- y is a dependent variable and $\bar{y}$ its mean.
- r stands for a correlation coefficient which depending on context can be r_p or r_s.
- r_p is Pearson's product-moment correlation coefficient r_p.
- r_s is Spearman's rank-order correlation coefficient.
- Z is a transformed score computed as $\frac{1}{2} ln\left(\frac{1+r}{1-r}\right)$.
- ATANH and arctanh stand for an inverse hyperbolic tangent function or transformation.
- TANH and tanh stand for a hyperbolic tangent function or transformation.

Discussion

In 1915, Fisher introduced the function that goes by his name, the Fisher *r*-to-*Z* Transformation,

$$Z = \frac{1}{2} ln \left(\frac{1+r}{1-r} \right) \quad (1)$$

whose inverse, the Fisher *Z*-to-*r* Transformation, is

$$r = \frac{e^{2Z} - 1}{e^{2Z} + 1} = \frac{e^{Z} - e^{-Z}}{e^{Z} + e^{-Z}} \quad (2)$$

He discovered more or less fortuitously that (1) makes the variability of *r* values that are close to ±1.00 comparable to that of *r* values in the mid-range, effectively acting as a variance stabilizing transformation.

At first, Fisher was not so sure about its efficacy, and at the end of his paper he wrote that the transformation "is not a little attractive, but so far as I have examined it, it does not simplify the analysis, and approaches relative constancy at the expense of the constancy proportionate to the variable…" (Fisher, 1915).

Later, he was more optimistic and proved for data from a bivariate distribution that the sampling distribution of *Z* was approximately normal, too (Fisher, 1921). The exact distribution of *Z* for data from a bivariate Type A Edgeworth distribution was confirmed years later by Gayen (1951). Hotelling (1953) calculated the Taylor series expressions for the moments of *Z* and several related statistics. Hawkins (1989) discovered the asymptotic distribution of *Z* for data from a distribution with bounded fourth moments.

Fisher Transformations are needed for computing weighted means of *r*, standard errors, confidence intervals, and for significance testing. They are also needed for meta-analyses based on Hedges-Olkin's fixed effect model (Hedges & Olkin, 1985; Field, 2001).

Essentially, (1) is an inverse hyperbolic tangent function (arctanh) and (2) a hyperbolic tangent function (tanh). Nowadays these transformations are available in most programming languages and software. The scientific calculator bundled with Windows supports them, and Microsoft Excel includes these as the ATANH and TANH functions. A tool for computing multiple Fisher Transformations is available online (Garcia, 2017b).

Limitations of Fisher Transformations

The *r*-to-*Z* Fisher Transformation (1) and its inverse (2) should not be applied arbitrarily, but only when the *x* and *y* random variables from which *r* is computed are approximately normally distributed; i.e., when the distributions of *x* and *y* are bell-shaped.

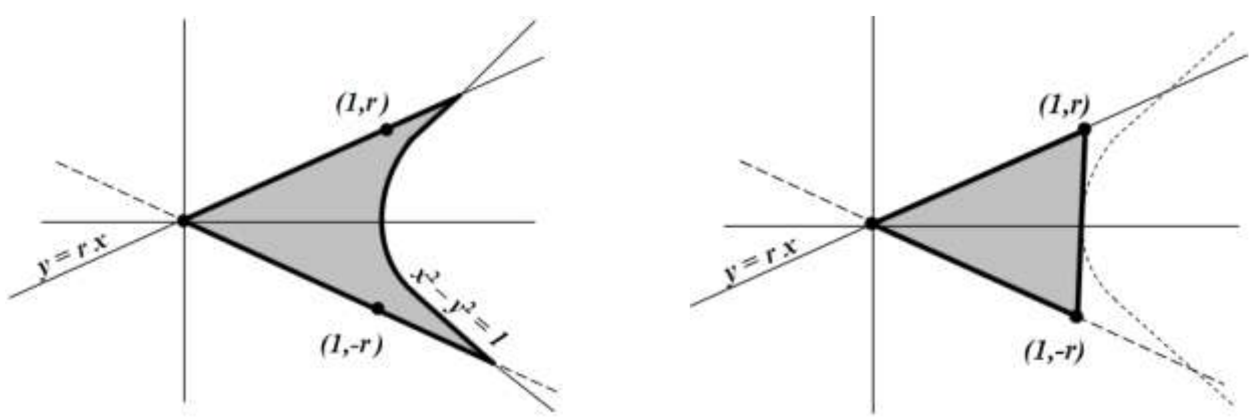
Zimmerman, Zumbo, and Williams (2003) have shown that arbitrarily applying the transformations, especially from distributions that violate bivariate normality can lead to spurious results, even for large sample sizes.

They found that "…significance tests of hypotheses about validity and reliability coefficients or differences between them require an assumption of bivariate normality despite large sample sizes. Researchers certainly should be aware of this assumption before using the *r* to *Z* transformation in data analysis."

The implication here is that an intrinsic geometric requirement must be present in the distribution of both variables for the transformations to be valid. Let us address this point.

Why is Bivariate Normality so Important?

Bond and Richardson (2004) have published the first geometrical interpretations of the Fisher Transformations. These are depicted in Figure 1 with permission from these authors.



(a) (b)

Figure 1. *Z* (a) and *r* (b) as areas in scatterplots. Source: Bond & Richardson, 2004.

Bond and Richardson (2004) explained: "Correlational statistics can be represented as areas in Euclidean space. Suppose that we have data on two variables (*X* and *Y*) which we have standardized in the usual manner via $x_i = \frac{X_i - \bar{X}}{s_X}$ and $y_i = \frac{Y_i - \bar{Y}}{s_Y}$. Then the least-squares regression line for predicting *y* from *x* is, of course, $y = rx$. Let us depict this regression line in a two-dimensional *xy* scatterplot, along with its reflection in the *x*-axis--the line $y = -rx$. Let us also insert into our *xy* scatterplot the unit hyperbola $H^1 = \{(x, y) \mid x^2 - y^2 = 1, x > 0\}$. The quantity Z_r can be regarded as the area enclosed by this hyperbola, the regression line, and its reflection."

"To justify this representation, let us begin by changing to polar coordinates $(x, y) = (u \cos \phi, u \sin \phi)$. Then the equation of the hyperbola becomes $x^2 - y^2 = u^2 (\cos^2 \phi - u \sin^2 \phi) = u^2 \cos (2\phi) = 1$, so that the area indicated in Figure 1 can be expressed as the integral"

$$\frac{1}{2}\int_{-arctan(r)}^{arctan(r)} u(\phi)^2 d\phi = \frac{1}{2}\int_{-arctan(r)}^{arctan(r)} \frac{1}{cos(2\phi)} d\phi = \frac{1}{2} ln\left(\frac{r+1}{r-1}\right) = arctanh(r) = Z_r$$

Clearly the requirement of bivariate normality comes from the transformations themselves. Violation of bivariate normality means that the distribution of at least one of the variables is not bell-shaped, essentially distorting the effective areas to be computed from the integration as shown in Figure 1.

As a result, if we ignore the bivariate normality requirement we may end reporting misleading *r* values along with spurious standard errors, confidence intervals, and significance tests. Increasing the size of a sample does not remove the error introduced. As Zimmerman, et. al (2003) correctly asserted: "…significance tests of non-zero values of correlation based on the *r* to *Z* transformation are grossly distorted for distributions that violate bivariate normality. Also, significance tests of non-zero values of r_S based on the *r* to *Z* transformation are distorted even for normal distributions."

"Another implication of the present findings is that, in practice, the *r* to *Z* transformation can be expected to be sensitive to violation of bivariate normality. This fact is relevant to hypotheses testing, finding confidence intervals, and averaging correlation coefficients. In these applications, neither large samples nor conversion to ranks affords protection. For example, significance tests of hypotheses about validity and reliability coefficients or differences between them require an assumption of bivariate normality despite large sample sizes."

“Researchers certainly should be aware of this assumption before using the *r* to *Z* transformation in data analysis. If it is not tenable, estimates of non-zero values of correlation coefficients can be extremely biased, and significance tests can be invalid. These consequences appear to be more severe than ones typically associated with non-normality in *t* and *F* tests of differences in location.”

Implications to Meta-Analysis

The limitations ascribed to Fisher Transformations (1) and (2) can impact a meta-analysis based on Hedges-Olkin’s fixed effect model (Hedges & Olkin, 1985; Field, 2001; Garcia, 2012).

In said model, *r* values from different studies are taken for effect sizes and transformed into *Z* scores with (1). After that, a weighted *Fisher Z* score of the form $\overline{Z} = \sum_j^k (n_j - 3) Z_j \Big/ \sum_j^k (n_j - 3)$ is computed and, if needed, transformed back into an average correlation score with (2).

For those calculations to be valid, all samples from all the studies should not violate the bivariate normality requirement. If this is not tenable, or said information is not available, a researcher might be tempted to make bold assumptions about normality for all of the studies involved or simply ignore the issue. Either way, the results can be challenged or in the worse case scenario the researcher is opening the door to a rebuttal. The presence of positive and negative correlations can complicate the interpretation of the results (Field, 2003).

Conclusion

We have briefly discussed Fisher Transformations and the risks arbitrarily implementing these when bivariate normality is not present in data sets. We have shown that this requirement comes from the transformations themselves.

In Part 3 of this series we will show that the mathematical relationship between regression, correlation, and association, as well as the biased nature of Pearson’s r_p and Spearman’s r_s provides additional arguments against arithmetically averaging correlations or arbitrarily applying (1) and (2).

References

Bond, C. F. and Richardson, K. (2004). Psychometrika, Vol 69, 2, 291-303. Retrieved from http://faculty.tcu.edu/richardson/seeing/seeingf060403.pdf

Field, A. P. (2001). Meta-analysis of correlation coefficients: a Monte Carlo comparison of fixed- and random-effects methods. Psychological Methods 6-2:161–180 Retrieved from http://psych.colorado.edu/~willcutt/pdfs/Field_2001.pdf

Field, A. P. (2003). Can meta-analysis be trusted? The Psychologist, 16, 642-645. Retrieved from http://sro.sussex.ac.uk/714/1/CAN_META-ANALYSIS_BE_TRUSTED.pdf

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika, 10, 507-521. Retrieved from http://digital.library.adelaide.edu.au/dspace/bitstream/2440/15166/1/4.pdf

Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. Metron 1, 3-31. Retrieved from https://digital.library.adelaide.edu.au/dspace/bitstream/2440/15169/1/14.pdf

Fisher, R. A. (1924). The distribution of the partial correlation coefficient. Metron 3, 329-332. Retrieved from https://digital.library.adelaide.edu.au/dspace/bitstream/2440/15182/1/35.pdf

Garcia, E. (2012). The Self-Weighting Model. Communications in Statistics - Theory and Methods Volume 41, 2012 - Issue 8. Taylor & Francis Online. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/03610926.2011.654037

Garcia, E. (2017a). On the Nonadditivity of Correlation Coefficients Part 1: Pearson's r and Spearman's r_s. Retrieved from http://www.minerazzi.com/tutorials/nonadditivity-correlations-part-1.pdf

Garcia, E. (2017b). Fisher Transformations. Retrieved from http://www.minerazzi.com/tools/fisher/transformations.php

Gayen, A. K. (1951). The Frequency Distribution of the Product-Moment Correlation Coefficient in Random Samples of Any Size Drawn from Non-Normal Universes. Biometrika. Biometrika Trust. 38 (1/2): 219–247. JSTOR 2332329. doi:10.1093/biomet/38.1-2.219. Retrieved from https://academic.oup.com/biomet/article-abstract/38/1-2/219/245776/THE-FREQUENCY-DISTRIBUTION-OF-THE-PRODUCT-MOMENT?redirectedFrom=fulltext

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. 231-235. Academic Press: Orlando.

Hawkins, D. L. (1989). Using U statistics to derive the asymptotic distribution of Fisher's Z statistic. The American Statistician. American Statistical Association. 43 (4): 235–237. JSTOR 2685369. doi:10.2307/2685369. Retrieved from https://www.jstor.org/stable/2685369?seq=1#page_scan_tab_contents

Hotelling, H (1953). New light on the correlation coefficient and its transforms. Journal of the Royal Statistical Society, Series B. Blackwell Publishing. 15 (2): 193–225. JSTOR 2983768. Retrieved from https://www.jstor.org/stable/2983768?seq=1#page_scan_tab_contents

Zimmerman, D. W., Zumbo, B. D. & Williams, R. H. (2003). Bias in estimation and hypothesis testing of correlation. Psicológica 24:133–158. Retrieved from http://www.uv.es/revispsi/articulos1.03/9.ZUMBO.pdf See also http://www.redalyc.org/articulo.oa?id=16924109