The Self-Weighting Model Tutorial: Part 2

Abstract – This is the second of a two-parts tutorial on the Self-Weighting Model (SWM), a new weighting model for statistical analysis. SWM allows within/between-set comparisons, producing estimates with a discriminatory power not found through current weighting strategies. The model is applicable to a wide range of statistical problems for which conditional weighted means are required.

Keywords: standard errors, sampling distributions, correlation coefficients, fisher transformations
Published: 03-10-2015; Updated: 10-08-2016
 admin@minerazzi.com

Introduction

This is the second of a two-parts tutorial on the Self-Weighting Model (SWM), a framework for computing weighted means.

The model was first published in *Communications in Statistics - Theory and Methods* (Garcia, 2012a) and presented at SES, New York (Garcia, 2012b). The content of this tutorial is based on these two references.

In part 1 (Garcia, 2015c), we stated the problem of computing valid averages from non-additive statistics like correlation coefficients. Previous weighting strategies for doing this were reviewed and the SWM was introduced. We now present a systematic derivation of the model.

Essentially, we identify local and global weights and the variability information stored in these weights. We also compare the results from the model with other weighting strategies and meta-analysis models (Dell, 2016; Hunter & Schmidt, 2000; Hedges & Olkin,1985).

Derivation

Let $S = \{Y_1, Y_2, \ldots Y_k\}$ be a set of k statistics, where Y_j consists of m number of constituent statistics

$$Y_j = f\,(X_{1_j}, X_{2_j}, \cdots X_{m_j}) \tag{1}$$

Assume that w_j is a local weight defined in terms of the constituent statistics of Y_j. The significance of this weight will soon be evident. Multiplying (1) by w_j

$$w_j Y_j = w_j \; f\,(X_{1_j}, X_{2_j}, \cdots X_{m_j}) \tag{2}$$

Raising this result to a power and taking summations, from j to k,

$$\sum_j^k (w_j)^p (Y_j)^p \;=\; \sum_j^k (w_j)^p [f(X_{1_j}, X_{2_j}, \cdots X_{m_j})]^p \tag{3}$$

Assume that $g = 1 \big/ \sum_j^k (w_j)^p$ (or its inverse, $g' = \sum_j^k (w_j)^p$) defines a global weight over a given set, S. Since any two sets, S_1 and S_2, have their own global weights, incorporating these weights into the model should allow comparisons between the sets.

So multiplying (3) by $g = 1 \big/ \sum_j^k (w_j)^p$ and taking the p^{th} root leads to a statistic that we call the *self-weighted power mean, M(p, w)*

$$M(p,w) \;=\; \left[\frac{\sum_j^k (w_j)^p (Y_j)^p}{\sum_j^k (w_j)^p} \right]^{1/p} \;=\; \left\{ \frac{\sum_j^k (w_j)^p [\, f(X_{1_j}, X_{2_j}, \cdots X_{m_j})]^p}{\sum_j^k (w_j)^p} \right\}^{1/p} \tag{4}$$

where

$$c_j \;=\; \frac{(w_j)^p (Y_j)^p}{\sum_j^k (w_j)^p (Y_j)^p} \tag{5}$$

is the contribution of each weighted Y to the $M(p, w)$ of a given set, S.

Thus, within-set comparisons are also possible with (5). Several examples are provided in the next section.

Illustrative Examples

As mentioned in part 2 of this tutorial, there are at least $2^m - 1$ ways of defining a local weight, w. For Pearson's r, m = 3 because its constituent terms are cov_{xy}, s_x, and s_y. In this case there are at least 7 ways of defining a local weight, w

- $w = s_y$
- $w = s_x$
- $w = s_x s_y$
- $w = 1/cov_{xy}$
- $w = s_y/cov_{xy}$
- $w = s_x/cov_{xy}$
- $w = (s_x s_y)/cov_{xy} = 1/r$

By contrast for a coefficient of variation, $cv_x = s_x / \bar{x}$ and m = 2 so there are least 3 different w's. Tables 1 and 2 summarize all these weights and the family of $M(p, w)$ expressions that can be derived from these.

When p = 1, a family of self-weighted means is obtained. When p = 2, the result is a family of self-weighted means equivalent to root mean square (*rms*) ratios. When p = 3, a family of asymmetric estimates (in theory, applicable to datasets with mixed signs) is obtained, and so on.

It can also be demonstrated that if k number of $w_j Y_j$ values form a vector, $\boldsymbol{wY}$, with p-norm equal to $\left[\sum_j^k \left|w_j Y_j\right|^p\right]^{1/p}$ and k number of w_j values form another vector, $\boldsymbol{w}$, with p-norm equal to $\left[\sum_j^k \left|w_j\right|^p\right]^{1/p}$, then when p = 2 an $M(p, w)$ is equivalent to an L_2 *norm ratio*.

Similarly when p = 1 and all $w_j Y_j$ and w_j are real positive quantities, $M(p, w)$ is equivalent to an L_1 *norm ratio*. See (2 – 4).

TABLE 1. The *M(p, w)* family of Pearson's correlation, where $Y = r = cov_{xy}/(s_x s_y)$ **.**

w	*M(p, w)*
s_y	$M(p,w) = \left[\frac{\sum_j^k (s_{y_j})^p (r_j)^p}{\sum_j^k (s_{y_j})^p}\right]^{1/p} = \left[\frac{\sum_j^k (cov_{xy_j}/s_{x_j})^p}{\sum_j^k (s_{y_j})^p}\right]^{1/p}$
s_x	$M(p,w) = \left[\frac{\sum_j^k (s_{x_j})^p (r_j)^p}{\sum_j^k (s_{x_j})^p}\right]^{1/p} = \left[\frac{\sum_j^k (cov_{xy_j}/s_{y_j})^p}{\sum_j^k (s_{x_j})^p}\right]^{1/p}$
$s_x s_y$	$M(p,w) = \left[\frac{\sum_j^k (s_{x_j} s_{y_j})^p (r_j)^p}{\sum_j^k (s_{x_j} s_{y_j})^p}\right]^{1/p} = \left[\frac{\sum_j^k (cov_{xy_j})^p}{\sum_j^k (s_{x_j} s_{y_j})^p}\right]^{1/p}$
$1/cov_{xy}$	$M(p,w) = \left[\frac{\sum_j^k (1/cov_{xy_j})^p (r_j)^p}{\sum_j^k (1/cov_{xy_j})^p}\right]^{1/p} = \left[\frac{\sum_j^k (1/s_{x_j})^p (1/s_{y_j})^p}{\sum_j^k (1/cov_{xy_j})^p}\right]^{1/p}$
s_y/cov_{xy}	$M(p,w) = \left[\frac{\sum_j^k (s_{y_j}/cov_{xy_j})^p (r_j)^p}{\sum_j^k (s_{y_j}/cov_{xy_j})^p}\right]^{1/p} = \left[\frac{\sum_j^k (1/s_{x_j})^p}{\sum_j^k (s_{y_j}/cov_{xy_j})^p}\right]^{1/p}$
s_x/cov_{xy}	$M(p,w) = \left[\frac{\sum_j^k (s_{x_j}/cov_{xy_j})^p (r_j)^p}{\sum_j^k (s_{x_j}/cov_{xy_j})^p}\right]^{1/p} = \left[\frac{\sum_j^k (1/s_{y_j})^p}{\sum_j^k (s_{x_j}/cov_{xy_j})^p}\right]^{1/p}$
$1/r$	$M(p,w) = \left[\frac{\sum_j^k (1/r_j)^p (r_j)^p}{\sum_j^k (1/r_j)^p}\right]^{1/p} = \left[\frac{k}{\sum_j^k (1/r_j)^p}\right]^{1/p}$

TABLE 2. The *M(p, w)* family of a coefficient of variation, where $Y = cv_x = s_x / \bar{x}$ **.**

w	$M(p, w)$
$\bar{x}$	$M(p,w) = \left[\dfrac{\sum_j^k (\bar{x}_j)^p (cv_{x\,j})^p}{\sum_j^k (\bar{x}_j)^p}\right]^{1/p} = \left[\dfrac{\sum_j^k (s_{x\,j})^p}{\sum_j^k (\bar{x}_j)^p}\right]^{1/p}$
$1/s_x$	$M(p,w) = \left[\dfrac{\sum_j^k (1/s_{x\,j})^p (cv_{x\,j})^p}{\sum_j^k (1/s_{x\,j})^p}\right]^{1/p} = \left[\dfrac{\sum_j^k (1/\bar{x}_j)^p}{\sum_j^k (1/s_{x\,j})^p}\right]^{1/p}$
$1/cv_x$	$M(p,w) = \left[\dfrac{\sum_j^k (1/cv_{x\,j})^p (cv_{x\,j})^p}{\sum_j^k (1/cv_{x\,j})^p}\right]^{1/p} = \left[\dfrac{k}{\sum_j^k (1/cv_{x\,j})^p}\right]^{1/p}$

Unfeasible Solutions

Depending on the nature of the statistics involved, some self-weighted power means might not be statistically feasible.

For example, in Tables 1 and 2 the first two *M(p, w)* expressions are valid solutions when $p = 2$, but not when $p = 1$. The reason is that when $p = 1$ these expressions involve additions of standard deviations which are not additive quantities. Thus, before modeling with SWM, one must pay attention to the nature of the statistical constituents to be used as the building blocks of candidate *M(p, w)* expressions.

Applications

From Table 1, setting $p = 2$ and squaring the result yields an expression equivalent to the weighted average coefficient of determination derived by Faller (Faller, 1981, 1982; Glahn, 1982); i.e.

$$\bar{r}^2 = \frac{\sum_j^k (s_{y\,j})^2 (r_j)^2}{\sum_j^k (s_{y\,j})^2} \qquad (6)$$

To understand how SWM compares with the weighting strategies mentioned earlier, consider the two sets, S_1 and S_2, given in Table 3 and adapted from Faller's paper.

TABLE 3. SWM vs. meta-analysis results for two sets of correlation coefficients at the same sample size level.

Data						SWM $M(p, w) = \bar{r}, p = 2, w = s_y$			Hunter-Schmidt $\bar{r} = \sum_j^k n_j r_j \big/ \sum_j^k n_j$	Hedges-Olkin $\bar{Z} = \sum_j^k (n_j - 3) Z_j \big/ \sum_j^k (n_j - 3)$
	j	${s_y}_j^2$	r_j^2	r_j	Z_j	c_j	$\bar{r}^2$	$\bar{r}$		
S_1	1	1.00	0.90	0.95	1.82	0.99	0.83	0.91	0.64	$\bar{Z} = 1.07$; $\bar{r} = 0.79$
	2	0.10	0.10	0.32	0.33	0.01				
S_2	1	1.00	0.10	0.32	0.33	0.53	0.17	0.42	0.64	$\bar{Z} = 1.07$; $\bar{r} = 0.79$
	2	0.10	0.90	0.95	1.82	0.47				

In Table 3 we assumed that all four samples are of same size. Notice that Hunter-Schmidt's model reduces to computing an arithmetic mean correlation of 0.64 for both sets. By contrast, Hedges-Olkin's fixed effect model returns an average mean *Z* score of 1.07, which when is *Z-to-r* transformed returns a correlation of 0.79 for both sets. That is, both meta-analysis models fail to discriminate between the two sets.

Computing an average correlation of the form $\bar{r} = \sqrt{\bar{R}}$, where $\bar{R} = (1/k)\sum_j^k R_j$ and $R_j = r_j^2$, as suggested by StatSoft, currently owned by Dell, would not help either as this approach returns $\bar{R} = 0.50$ and $\bar{r} = 0.71$ for both sets (Dell, 2016). The reason as to why all these weighting strategies fail to discriminate between the sets can be ascribed to the fact that these do not incorporate variability information present in the original data sets.

By contrast, SWM incorporates the missing piece of information through the computed local and global weights. Thus in Table 3, SWM returns a self-weighted mean correlation of 0.91 for S_1 and of 0.42 for S_2. Therefore, between-set comparisons are possible.

Within-set comparisons are also possible. In S_1, $c_1 = 0.99$ and $c_2 = 0.01$, meaning that the first sample influences more the self-weighted mean correlation of the set than the second sample. In S_2, the difference in contributions from individual samples to the self-weighted mean is now smaller: $c_1 = 0.53$ vs. $c_2 = 0.47$. These results agree with those of Faller (1981); i.e., we have shown that Faller's transformation is a particular solution of the SWM framework.

Conclusion

The Self-Weighting Model and a new measure, the self-weighted power mean, *M(p, w)*, have been presented and compared with current meta-analysis models.

We have shown that Fisher, Hunter-Schmidt, and Hedges-Olkin transformations can fail to discriminate correlations cases. When a constant sample size is used, Hunter-Schmidt's model returns an arithmetic average. Studies based on such averages can be challenged on the grounds that correlations are not additive and, therefore, that the computed means are invalid statistics.

Valid weighted means are essential to many fields that involve inference, predictions, and risk assessment. For instance, if $s_x = s_y$, $r = \beta = cov_{xy} /(s_y^2)$. In portfolio management, β is referred to as *financial elasticity* or *correlated relative volatility*. This is the formula for the beta of an asset within a portfolio where *x* measures the rate of return of the asset, *y* the rate of return of the portfolio, and cov_{xy} is the covariance between the rates of return (Wikipedia, 2015). Thus, we can use SWM to compute valid weighted $\bar{\beta}$ in a straightforward manner.

Our model is not limited to correlation coefficients or coefficients of variations. SWM can be incorporated into current meta-analysis strategies and software as a new discriminatory layer for statistical modeling.

If $p = 1$ and all w_jY_j and w_j are real positive quantities, *M(p, w)* is equivalent to an L_1 *norm ratio*. On the other hand when $p = 2$, *M(p, w)* is equivalent to an *rms ratio* and L_2 *norm ratio*. Establishing the meaning of *M(p, w)* for higher *p* values requires further studies.

Moreover, from the last row of Table 1 and Table 2 when $p = 1$, $w_j = 1/Y_j$, and $\sum_j^k (w_j)^p (Y_j)^p = k$ the *M(p, w)* statistic reduces to the *harmonic mean*, a statistic that arises frequently in engineering and science. These findings suggest that SWM could be used for a broad range of engineering, science, information retrieval, and data mining problems.

References

Dell (2016). What are Basic Statistic? Retrieved from http://www.statsoft.com/textbook/basic-statistics/#Correlationso

Faller, A. J. (1981). An average correlation coefficient. Journal of Applied Meteorology 20:203–205. Retrieved from http://journals.ametsoc.org/doi/pdf/10.1175/1520-0450(1981)020%3C0203:AACC%3E2.0.CO;2

Faller, A. J. (1982). Reply. Journal of Applied Meteorology 21:1203–1205. Retrieved from http://journals.ametsoc.org/doi/pdf/10.1175/1520-0450(1982)021%3C1203%3AR%3E2.0.CO%3B2

Garcia, E. (2012a). The Self-Weighting Model, Communications in Statistics - Theory and Methods, 41:8,1421-1427. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/03610926.2011.654037

Garcia, E. (2012b). True Lies: Search Marketing Data Errors Uncovered. Retrieved from http://web.archive.org/web/20121101095532/http:/sesconference.com/archive/2012/newyork/agenda-day1.php

Garcia, E. (2015c): The Self-Weighting Model Tutorial: Part 1 Retrieved from http://www.minerazzi.com/tutorials/self-weighting-model-tutorial-part-1.pdf

Glahn, H. R. (1982). Comments on An average correlation coefficient. Journal of Applied Meteorology 21:1202–1203. Retrieved from http://journals.ametsoc.org/doi/pdf/10.1175/1520-0450(1982)021%3C1202%3ACOACC%3E2.0.CO%3B2

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

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: implications for cumulative knowledge in Psychology. International Journal of Selection and Assessment 8:275–292. Retrieved from http://www.biz.uiowa.edu/faculty/fschmidt/meta-analysis/hunter_schmidt_2000_rev.pdf

Wikipedia (2015). *Beta (Finance)*. Retrieved from http://en.wikipedia.org/wiki/Beta_(finance)