One-Tailed or Two-Tailed P Values in PLS-SEM?

Ned Kock, Division of International Business and Technology Studies, Texas A&M International University, Laredo, TX, USA

ABSTRACT

Should P values associated with path coefficients, as well as with other coefficients such as weights and loadings, be one-tailed or two-tailed? This question is answered in the context of structural equation modeling employing the Partial Least Squares Method (PLS-SEM), based on an illustrative model of the effect of e-collaboration technology use on job performance. A one-tailed test is recommended if the coefficient is assumed to have a sign (positive or negative), which should be reflected in the hypothesis that refers to the corresponding association. If no assumptions are made about coefficient sign, a two-tailed test is recommended. These recommendations apply to many other statistical methods that employ P values, including path analyses in general, with or without latent variables, plus univariate and multivariate regression analyses.

INTRODUCTION

A common question often arises in the context of discussions about structural equation modeling (SEM) employing the partial least squares (PLS) method, referred to here as PLS-SEM (Kock, 2013b, 2014; Kock & Lynn, 2012), among researchers in the field of e-collaboration (Kock, 2005; Kock & Nosek, 2005) as well as many other fields. Should P values associated with path coefficients be one-tailed or two-tailed?

This is an important question because normally one-tailed tests yield lower P values than two-tailed tests. In fact, this is always the case when symmetrical distributions of path coefficients are assumed, such as Student’s t-distributions. Therefore, the decision as to whether to use one-tailed or two-tailed tests can influence whether one or more hypotheses are accepted or rejected. This decision also influences the statistical power of a PLS-SEM analysis, with the power being higher with tests employing one-tailed P values.

We try to provide an answer to this question, which requires brief ancillary discussions of related topics – e.g., PLS-SEM’s treatment of measurement error. While our discussion addresses path coefficients, it also applies to other coefficients such as weights and loadings. Even though the focus is on PLS-SEM, much of what is said here applies to many other statistical analysis techniques. Among these are path analyses in general, without or without latent variables, as well as univariate and multivariate regression analyses.

DOI: 10.4018/ijec.2015040101
ILLUSTRATIVE MODEL

The discussion presented in this study is based on the illustrative model shown in Figure 1. This model contains two latent variables, e-collaboration technology use ($L$) and job performance ($J$). Each latent variable is measured indirectly through three indicators.

Let us assume that $J$, $L$, $x_{Li}$, and $x_{ji}$ ($i = 1 \ldots 3$) are scaled to have a mean of zero and a standard deviation of one (i.e., these variables are standardized). Our illustrative model can then be described by equations (1), (2), and (3).

$$x_{Li} = \lambda_{Li} L + \theta_{Li}, \ i = 1 \ldots 3.$$  \hspace{1cm} (1)

$$x_{ji} = \lambda_{ji} J + \theta_{ji}, \ i = 1 \ldots 3.$$  \hspace{1cm} (2)

$$J = \beta L + \varepsilon.$$  \hspace{1cm} (3)

The path coefficient $\beta$ and loadings $\lambda_{Li}$ and $\lambda_{ji}$ ($i = 1 \ldots 3$) are assumed to describe the model at the population level, as true values. The population is made of teams of individuals who use an integrated e-collaboration technology including e-mail and voice conferencing to different degrees. That is, the unit of analysis is the team, not the individual.

The e-collaboration technology facilitates the work of the teams. Different values of job performance by the teams, where performance is evaluated by managers, are associated with different degrees of use of the e-collaboration technology.

PLS-SEM AND MEASUREMENT ERROR

PLS-SEM algorithms estimate latent variable scores as exact linear combinations of their indicators (i.e., as “composites”). As such, they do not properly account for measurement error. This can be illustrated through (4) and (5); where latent variable scores are calculated properly accounting for, and not properly accounting for, the measurement error $\varepsilon$. Both equations denote the number of indicators as $n$.

$$F = \sum_{i=1}^{n} \omega_{Fi} x_{Fi} + \varepsilon.$$  \hspace{1cm} (4)

$$\hat{F} = \sum_{i=1}^{n} \omega_{Fi} \hat{x}_{Fi}.$$  \hspace{1cm} (5)

A full discussion of the effects of PLS-SEM not properly accounting for measurement error is outside the scope of this study. Nevertheless, one effect that will be noticed in the next section is that the path coefficient is attenuated, due to the correlation attenuation property (Nunnally & Bernstein, 1994) expressed in (6).

$$r(\hat{F}_i, \hat{F}_j) = r(F_i, F_j) \sqrt{\alpha_i \alpha_j}.$$  \hspace{1cm} (6)

In this correlation attenuation equation, $\alpha_i$ and $\alpha_j$ denote the true reliabilities of the true latent variables $F_i$ and $F_j$, which are estimated via PLS-SEM as $\hat{F}_i$ and $\hat{F}_j$. These true reliabilities can be estimated through the Cronbach’s alpha coefficients for the latent variables.

Figure 1. Illustrative model
Collaborative Distance: A Framework for Distance Factors Affecting the Performance of Distributed Collaboration
www.igi-global.com/article/collaborative-distance-framework-distance-factors/42100?camid=4v1a

Spatial Planning and Regional Growth: A Benchmarking Study for North-South Aegean and Crete
www.igi-global.com/chapter/spatial-planning-and-regional-growth/206043?camid=4v1a