A Tutorial on Standard Errors

E. Garcia
admin@miislita.com

Published on: June 25, 2010; Last Modified: June 25, 2010 © 2010 E. Garcia

Keywords: tutorial, standard errors, correlation coefficients, Fisher Transformation

Abstract: This is a tutorial on standard errors. Every statistic from a sample distribution has a standard error that is specific to that statistic. Using the incorrect definition for a standard error invalidates any research study.

Introduction

The standard deviation of the sampling distribution of a statistic is known as the *standard error* of that statistic (1). A standard error, denoted by *stderr* or SE, measures the accuracy of a statistic. According to Arsham (2):

> "To express the accuracy of the estimates of population characteristics, one must also compute the standard errors of the estimates. These are measures of accuracy that determine the possible errors arising from the fact that the estimates are based on random samples from the entire population, and not on a complete population census. "

In general, SE values are computed in different ways depending on the statistic used. A comprehensive list of SE definitions is provided elsewhere (2, 3). This tutorial covers a limited number of these.

The Standard Error of the Mean

This is probably the best known standard error. The standard error of the mean is the standard deviation of the sample mean estimate $\bar{x}$ of a population mean (1); i.e.

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}} \qquad \text{(Eq 1)}$$

where s is the sample standard deviation (i.e., the sample-based estimate of the standard deviation of the population), and n is the size of the sample or number of observations. To compute s and $\bar{x}$ the observations must be additive.

The Standard Error of the Median

For large samples with normal distributions (3), the SE of the median can be approximated as

$$SE_m = 1.25\frac{s}{\sqrt{n}} = 1.25\, SE_{\bar{x}} \qquad \text{(Eq 2)}$$

In most cases, $SE_{\bar{x}}$ is preferred since $SE_{\bar{x}} < SE_m$ (3 - 5). This formula can yield wrong results for extremely non-normal distributions. Although SE_m is computed when the data is normally distributed, bootstrapping SE_m avoids this requirement. Bootstrap is a useful treatment for computing confidence intervals and standard errors of difficult statistics like the median (6 - 9).

The Standard Error of the Standard Deviation

The SE of the standard deviation (SE_s) is about 71% that of the mean (3),

$$SE_s = 0.71 \frac{s}{\sqrt{n}} \quad \text{(Eq 3)}$$

The distribution of the standard deviation is positively skewed for small n but is approximately normal if the sample size is 25 or greater. Procedures for calculating the area under the normal curve work for the sampling distribution of the standard deviation as long as the sample size is at least 25 and the distribution is approximately normal (3).

The Standard Error of a Correlation Coefficient

The SE of a correlation coefficient r is computed by normalizing the fraction of the unexplained variations with respect to $n - 2$ degrees of freedom (10); i.e.

$$SE_r = \frac{\sqrt{1-r^2}}{\sqrt{n-2}} \quad \text{(Eq 4)}$$

where r^2 is the Coefficient of Determination which expresses the fraction of the explained variations (variations in y as the result of variations in x). For example, if r^2 is 0.90, then the independent variable y is said to explain 90% of the variance in the dependent variable x, but does not explain $1 - r^2$ or 10% of the variance in the dependent variable.

The Standard Error of a Fisher Transformation

The Fisher Transformation converts a correlation coefficient into a *z score* sometimes also known as a *normal score* (2, 10 – 13). The profile of this transformation is shown in Figure 1 and is taken from reference 11.

This transformation is used to compare any two correlation coefficients. It is computed for each r value as follows.

$$z = \frac{1}{2}[ln(1 + r) - ln(1 - r)] \quad \text{(Eq 5)}$$

where z scores approach a normal distribution. The standard error associated to a z value is

$$SE_z = \sqrt{\frac{1}{n-3}} = \frac{1}{\sqrt{n-3}} \quad \text{(Eq 6)}$$

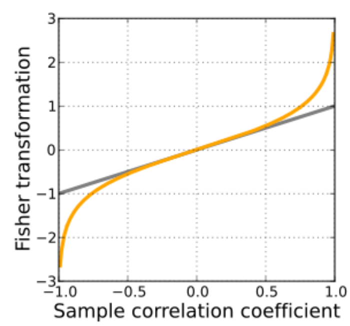

Figure 1. Fisher z Transformation. Source: Reference 11.

Pooled Standard Errors

To test whether any two correlation coefficients, r_1 and r_2, are significantly different, these are transformed into Fisher *z-scores*. Their difference, computed as $z_1 - z_2$, is tested using a pooled standard error (2, 10 - 13) which is defined as follows:

$$SE_z = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \quad \text{(Eq 7)}$$

Beware of Incorrect Standard Error Calculations

A statistical analysis makes sense only if there is a sound theory behind it. Using incorrect *SE* definitions invalidates arguments stated or implied in any statistical study.

For instance, recently search engine marketers from SEOmoz.org published reports using standard errors of correlation coefficients (14, 15). It was later acknowledged that the *SE* of a correlation coefficient was literally computed as $\frac{s}{\sqrt{n}}$ where s was a standard deviation computed out of several correlation coefficients (16). Conclusions were then drawn from these calculations.

The problem with this approach is that there is no such thing as a 'standard deviation of a correlation coefficient', at least not as computed by SEOmoz. Why?

To compute such a standard deviation it is necessary to compute first a mean correlation coefficient. But there is a problem: **correlation coefficients are not 'additive'**.

As stated by StatSoft, creators of *Statistica* (17):

> "Are correlation coefficients "additive?" No, they are not. For example, an average of correlation coefficients in a number of samples does not represent an "average correlation" in all those samples. Because the value of the correlation coefficient is not a linear function of the magnitude of the relation between the variables, correlation coefficients cannot simply be averaged. In cases when you need to average correlations, they first have to be converted into additive measures. For example, before averaging, you can square them to obtain *coefficients of determination* which are additive (as explained before in this section), or convert them into so-called *Fisher z* values, which are also additive."

In other words, it is not possible to add r values and then average these in order to compute a so-called mean and standard deviation of correlation coefficients. In general, a correlation coefficient is not a linear function of the magnitude of the relation between variables, but a function of the covariance between two variables already normalized by their standard deviations. The covariance itself is defined in terms of the expectation (mean) values of the variables; i.e.

$$r = \frac{covar(x,y)}{s_x * s_y} \quad \text{(Eq 8)}$$

$$r = \frac{E(x*y) - E(x)*E(y)}{s_x * s_y} \quad \text{(Eq 9)}$$

where $E(x*y)$, $E(x)$, and $E(y)$ are expectation values.

In the case of a Spearman Correlation Coefficient, the mere notion of constructing such a linear function by averaging Spearman values is highly questionable because what is considered are ranks, not the magnitude of the relation between the variables. It is true that is possible to compute a *Spearman* value as a *Pearson* value for the rank data, but still the computed correlation coefficients are not additive.

Applications

Once a standard error is properly computed for a particular statistic: what we do with it?

First, the statistic in question divided by its standard error gives a way of testing whether the statistic is significantly different from zero.

A second application of the standard error is the production of confidence intervals (18).

A third application consists in testing whether any two statistics of the same kind are significantly different. In an upcoming tutorial we discuss some of these applications.

SE computations can get messy, depending on the statistic under consideration. As noted by Simon (18):

> "There are a lot of subtleties in the use of the standard error, especially in more complex problems. Sometimes, for example, the standard error applies not to the statistic itself, but to the logarithm of that statistic. For example, a logistic regression model will compute an odds ratio for your data, but the standard error refers not to the odds ratio, but to the log odds ratio. In this situation, you need to compute the confidence interval on the log scale and then transform the results back to the original scale of measurement."

Conclusion

As mentioned before, a statistical analysis makes sense only if there is a sound theory behind it. Using incorrect *SE* definitions invalidates arguments stated in any statistical study. Building a theory based on spurious statistical measures or ideas made out of thin air amounts to what is commonly deemed as "quack" Science (19, 20). In an upcoming tutorial we describe several applications for standard errors. In particular, we explain how *SE* values are used in correlation coefficient testing and in the proper way of comparing *r* values.

References

1. Kurtz, M. *Handbook of Applied Mathematics for Engineers and Scientists*, Chapter 10, p10.4. McGraw-Hill, New York (1991).
2. Arsham, H. *Statistical Thinking for Managerial Decisions* (1994).
   http://www.mirrorservice.org/sites/home.ubalt.edu/ntsbarsh/Business-stat/opre504.htm
3. Galvan, A. *Distribution Equations*
   http://cnx.org/content/m10245/latest/
4. Stockburger, D. W. *Introductory Statistics: Concepts, Models, and Applications* (1996)
   http://www.psychstat.missouristate.edu/introbook/sbk19m.htm
5. Lane, D. M., *HyperStat Online Contents: Sampling Distribution of Median*
   http://davidmlane.com/hyperstat/index.html
6. Maindonald, J., Brown, J. *Data Analysis and Graphics Using R: An Example-Based Approach*, Chapter 4. Cambridge University Press (2003).
   http://wwwmaths.anu.edu.au/~johnm/r-book/daagur3.html
7. Rolke, W. A. *ESMA 6661 Theory of Statistics: The Bootstrap*
   http://math.uprm.edu/~wrolke/esma6661/boot.htm
8. Rolke, W. A. *ESMA 6661 Theory of Statistics: Some Applications of the Bootstrap*
   http://math.uprm.edu/~wrolke/esma6661/boot1.htm
9. Caffo, B. *Lecture 12* (2007).
   http://www.docstoc.com/docs/43117393/lecture12
10. Price, D. BioStatistics Topic 23 *Correlation Analyses*
    http://www2.hawaii.edu/~donaldp/cbes685/doc/BioStatistics%20Topic%2023.pdf
11. Fisher Transformation
    http://en.wikipedia.org/wiki/Fisher_transformation
12. Garson, D. Correlation (1998).
    http://faculty.chass.ncsu.edu/garson/PA765/correl.htm
13. Yaffee, R. A., *Common Correlation and Reliability Analysis with SPSS for Windows* (2003).
    http://www.nyu.edu/its/statistics/Docs/correlate.html
14. SEOmoz: *The Science of Ranking Correlations* (2010).
    http://www.seomoz.org/blog/the-science-of-ranking-correlations
15. SEOmoz: *Google vs. Bing Correlation Analysis of Ranking Elements* (2010).
    http://www.seomoz.org/blog/google-vs-bing-correlation-analysis-of-ranking-elements
16. Sphinn: *SEO is Mostly a Quack Science* (2010).
    http://sphinn.com/story/151848/#77530
17. Statsoft Textbook: *Basic Statistics*
    http://www.statsoft.com/textbook/basic-statistics/#Correlationso
18. Simon, S. *What is a Standard Error?* (2002).
    http://www.childrens-mercy.org/stats/definitions/stderr.htm
19. Dziuba, T. *SEO Is Mostly Quack Science* (2010).
    http://teddziuba.com/2010/06/seo-is-mostly-quack-science.html
20. Garcia, E. *Beware of SEO Statistical Studies* (2010).
    http://irthoughts.wordpress.com/2010/04/23/beware-of-seo-statistical-studies/