DOCUMENT RESUME

ED 050 151 TM 000 540

AUTHOR	Stanley, Julian C.; Livingston, Samuel A.
TITLE	Correcting Four Similar Correlational Measures for Attenuation Due to Errors of Measurement in the Dependent Variable: Eta, Epsilon, Omega, and Intraclass r.
INSTITUTION	Johns Hopkins Univ., Baltimore, Md. Center for the Study of Social Organization of Schools.
PUB DATE	Feb 71
NOTE	10p.; Part of a symposium entitled "Some Attenuating Effects of Errors of Measurement," presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Analysis of Variance, *Cluster Grouping, *Correlation, Data Analysis, *Reliability, Sampling, *Standard Error of Measurement, Statistical Analysis, Surveys, True Scores
IDENTIFIERS	*Attenuation (Correction)

ABSTRACT

Besides the ubiquitous Pearson product-moment r, there are a number of other measures of relationship that are attenuated by errors of measurement and for which the relationship between true measures can be estimated. Among these are the correlation ratio (eta squared), Kelley's unbiased correlation ratio (epsilon squared), Hays' omega squared, and the intraclass coefficient of correlation, expressed as a ratio of variance components by Ronald Fisher. This paper shows how to correct each of these for attenuation. Such corrections permit estimates of relationships between true scores to be made when estimates of relevant reliability coefficients are available. (Author)

Correcting Four Similar Correlational Measures for Attenuation Due to Errors of Measurement in the Dependent Variable: Eta, Epsilon, Omega, and Intraclass r*

Julian C. Stanley and Samuel A. Livingston
The Johns Hopkins University

As Miss Finucci indicated, for two-thirds of a century measurement specialists have been concerned with correcting the Pearson product-moment coefficient of correlation for diminution due to errors of measurement. How errors of measurement affect other simple measures of co-relationship does not seem to have been studied as thoroughly. At the risk of repeating results that may already be in the literature somewhere, though unknown to me and probably to you, Mr. Livingston and I will provide the rationale and results for several correlational statistics. These are eta, epsilon, omega, and intraclass r .

Let us begin with the venerable correlation ratio, eta squared. In analysis of variance language for a one-way classification it may be expressed as the ratio of the sum of squares between groups to the total sum of squares.

Your handout shows the abstract layout. Eta has perhaps been most useful recently for studying the clustering of groups in survey sampling, where group membership constitutes a nominal classification. There may be any number of groups from 2 upward. How great are differences among groups on a

*The third paper of a five-paper symposium entitled "Some Attenuating Effects of Errors of Measurement," American Educational Research Association convention, New York City, 5 February 1971.

dependent variable, compared with differences within groups? Errors of classification may occur, as for example when an attorney is misclassified as a physician, but these are not errors of measurement in the classical sense. In this paper we deal only with the attenuating influence of errors of measurement in the dependent variable. We do not consider correlation ratios for those situations where one may compute two different etas, one for Y on X and the other for X on Y.

The method and results are shown in the handout. We start with eta squared and give it rather detailed treatment. Eta squared is defined in formula (1). There the SS's are sums of squares from an analysis of variance.

The correction for attenuating errors of measurement in Y is given in formula (6). The first term there is eta squared divided by the reliability coefficient of the Y's, which is the same correction one makes in r^2. However, there is a non-negative subtraction term that becomes 0 only for infinite N or for perfect reliability of the Y's. This reduction term seems required because the chance value of eta squared is not zero, as it is for r^2, but instead is capital G minus one, divided by capital N minus one. Thus, for finite sample size the range of possible non-chance etas is less than the range of r's.

Formula (7) shows Kelley's so-called unbiased correlation ratio, epsilon squared, as a function of eta squared. Its attenuation due to errors of measurement in the Y's is caused by attenuation in eta squared.

Hays' ratio, which he called omega squared, is shown in formula (8). It differs from epsilon squared only by having the mean-squared-within in its denominator. This of course causes it to yield lower values than epsilon squared. The attenuating effects of errors of measurement tend to affect its numerator little, but they can greatly increase its denominator.

Eta, epsilon, and omega are closely related, as Glass and Hakstian showed splendidly in their 1969 AERJ article that won for them the Palmer O. Johnson Memorial Award they received last night. The fourth correlation ratio which we consider, intraclass r , is similar in appearance. Compare formulas (8) and (10). Omega squared and intraclass r have the same numerator and the same mean squares in their denominators, but the coefficients of the mean-squares-within is considerably different in the two denominators. Intraclass r has a rather different rationale from eta, epsilon, and omega. As explained in the handout, it is based on a ratio of variance components in the situation where levels of the ANOVA factor have been sampled randomly from a target population of such levels.

Intraclass r is especially useful when, for example, one wishes to study the correlation of the heights of brothers. "Family" then constitutes the ANOVA factor. A sample of capital G families is drawn randomly from a large population of families to which one wishes to generalize. Then the heights of the brothers in the sample are obtained, and an analysis of variance is performed. The variance of the means of all families in the population is estimated. This estimated variance component is then divided by itself plus the estimated within-group variance, as shown in the first part of formula (9).

The numerator of formula (10) is affected little by errors of measurement, whereas the denominator can be increased considerably by them. However, intraclass r would seem to be less attenuated by errors of measurement in the Y's than epsilon squared or omega squared are. One must remember, though, that these three statistics do not have the same value for a given set of data. Intraclass r will be largest, and omega squared smallest.

In this brief paper and handout we have merely scratched the surface of an important and seemingly neglected area. Perhaps it has had little attention from educational researchers and psychologists because analytical survey sampling has been the province mainly of sociologists. Survey researchers have not, until recently, seemed much interested in errors of measurement. There are signs that this may be changing.

Persons who delve into this area will quickly discern problems that we have only implied in our presentation. Among these are unbiasedness versus maximum likelihood estimation, computation of appropriate reliability coefficients, correction for errors of measurement in the X variable, and the influence of technical errors in setting factor levels. For some surveys misclassification may be an important cause of lowered correlation ratios. For other surveys or experiments, imprecise determination of the factor levels, particularly of an ordered variable, may have serious consequences.

Most of the techniques and procedures needed to further the study of correlation ratios already exist. There are unsolved problems enough for at least one worthwhile doctoral dissertation.

Handout for paper by Julian C. Stanley and Samuel A. Livingston entitled "Correcting Four Similar Correlational Measures for Attenuation Due to Errors of Measurement in the Dependent Variable," presented on 5 February 1971 as one of the five papers of a symposium entitled "Some Attenuating Effects of Errors of Measurement," American Educational Research Association convention, New York City.

Y_{ig} notation, the ith observation ($i = 1, 2, \ldots, n_g$) in the gth group ($g = 1, 2, \ldots, G$). $\sum_{g=1}^{G} n_g = N$. Groups constitute the independent variable (X).

	1	2	...	G	
	Y_{11}	Y_{12}	...	Y_{1G}	
	Y_{21}	Y_{22}	...	Y_{2G}	
	.	.		.	
	.	.		.	
	.	.		.	
	$Y_{n_1 1}$	$Y_{n_2 2}$	...	$Y_{n_G G}$	
Means	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	...	$\bar{Y}_{.G}$	$\bar{Y}_{..}$

Pearson's (1903, 1905, 1911, 1923) Correlation Ratio

Define the correlation ratio as

$$\eta_{YX}^2 = \frac{SS_{between}}{SS_{total}} = 1 - \frac{SS_{within}}{SS_{total}} . \qquad (1)$$

Take its expectation over an infinite number of random samples of N observations each:

$$E\left[\eta_{YX}^2\right] = E\left[1 - \frac{SS_w}{SS_{tot}}\right] = 1 - E\left[\frac{SS_w}{SS_{tot}}\right] \doteq 1 - \frac{E\left[SS_w\right]}{E\left[SS_{tot}\right]} = 1 - \frac{E\left[(N - G)MS_w\right]}{E\left[(N - 1)MS_{tot}\right]} =$$

$$1 - \frac{(N - G)E\left[MS_w\right]}{(N - 1)E\left[MS_{tot}\right]} = 1 - \frac{(N - G)\sigma_w^2}{(N - 1)\sigma_{tot}^2} . \quad (2)$$

Analytically, the approximation seems fairly good if N is moderately large.

For error-free Y measures one has

$$\eta_{T_Y X}^2 = 1 - \frac{(N - G)\sigma_{T_w}^2}{(N - 1)\sigma_{T_{tot}}^2} , \quad (3)$$

where T designates true scores on the dependent (i.e., Y) variable. Formula (3) may be rewritten as

$$\eta_{T_Y X}^2 = 1 - \frac{(N - G)(\sigma_w^2 - \sigma_e^2)}{(N - 1)(\sigma_{tot}^2 - \sigma_e^2)} , \quad (4)$$

because of the familiar model of classical test-score theory that produces the relationship $\sigma_Y^2 = \sigma_T^2 + \sigma_e^2$; σ_e^2 is the variance of measurement errors.

A computing form of (4) is

$$\hat{\eta}_{T_Y X}^2 = 1 - \frac{(N - G)(MS_w - \hat{\sigma}_e^2)}{(N - 1)(MS_{tot} - \hat{\sigma}_e^2)} , \quad (5)$$

where $\hat{\sigma}_e^2 = MS_{tot}(1 - \hat{\rho}_{T_Y Y}^2) = MS_w(1 - \hat{\rho}_{T_y y}^2)$. $\rho_{T_Y Y}^2$ is the overall reliability coefficient of the Y's ; $\hat{\rho}^2_{T_y y}$ is the within-groups reliability coefficient of the Y's . Via classical test-score theory it is assumed that $\sigma^2_{e_t} = \sigma^2_{e_w}$. See Stanley (1971).

Formula (5) may also be written as

$$\hat{\eta}^2_{T_Y X} = \frac{\eta^2_{YX}}{\hat{\rho}^2_{T_Y Y}} - \left(\frac{G-1}{N-1}\right)\left(\frac{1-\hat{\rho}^2_{T_Y Y}}{\hat{\rho}^2_{T_Y Y}}\right) . \tag{6}$$

This makes clear that, when reliability of the Y's is perfect (i.e., $\rho^2_{T_Y Y} = 1$) , $\eta^2_{T_Y X} = \eta^2_{YX}$.

Kelley's (1935) Epsilon Squared

$\hat{\eta}^2_{T_X X} > \hat{\eta}^2_{YX}$ when $\hat{\eta}^2_{YX} > \frac{G-1}{N-1}$, because $\frac{G-1}{N-1}$ is the chance value of η^2_{YX} . This is the "bias" that Kelley tried to remove via his epsilon squared (see Glass and Hakstian, 1969):

$$\epsilon^2_{YX} = 1 - \frac{MS_w}{MS_{tot}} = 1 - \left(\frac{N-1}{N-G}\right)\left(1 - \eta^2_{YX}\right) . \tag{7}$$

When $MS_w = MS_b$, epsilon squared becomes 0 .

Correcting ϵ^2 for attenuation merely involves correcting the η^2 in its formula via formula (5) or (6), above, or directly by subtracting the variance error of measurement from numerator and denominator of (7).

Hay's (1963) Omega Squared

Hays' ω^2 differs from ε^2 in the denominator only (see Glass and Hakstian, 1969, formulas 13 and 14), so it is also "unbiased."

$$\hat{\omega}^2 = \frac{SS_b - (G - 1)MS_w}{SS_{tot} + MS_w} = \frac{(N - 1)(MS_{tot} - MS_w)}{(N - 1)MS_{tot} + MS_w} . \qquad (8)$$

If one merely subtracts $\hat{\sigma}_e^2$ from MS_{tot} and also from MS_w in the numerator, it is unaffected. (Of course, $MS_{tot} - MS_w$ is somewhat affected by errors of measurement, though not systematically.) The denominator, however, will be reduced by $N\hat{\sigma}_e^2$, and hence the corrected value may greatly exceed ω^2.

The Intraclass Coefficient of Correlation

Intraclass ρ (Harris, 1913; Fisher, 1925) is defined in random-effects-model ANOVA components of variance as follows:

$$\rho_{intraclass} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_w^2} ,$$

where

$$E[MS_b] = \sigma_w^2 + \left(\frac{N - \frac{\sum_{g=1}^{G} n_g^2}{N}}{G - 1} \right) \sigma_\alpha^2 , \text{ and } E[MS_w] = \sigma_w^2 .$$

Call the coefficient in parentheses c .

$Y_{ig} = \mu + \alpha_g + w_{ig}$. $\alpha \sim$ n.d. $(0 , \sigma_\alpha^2)$; $w \sim$ n.d. $(0 , \sigma_w^2)$ for every g .

$\alpha_g = \mu_g - \mu$; $w_{ig} = Y_{ig} - \mu_g$.

$$\hat{\rho}_{intraclass} = \frac{\frac{MS_b - MS_w}{c}}{\frac{MS_b - MS_w}{c} + MS_w} = \frac{MS_b - MS_w}{MS_b + (c - 1)MS_w}$$

$$= \frac{(N - 1)(MS_{tot} - MS_w)}{(N - 1)MS_{tot} - [(N - 1) - c(G - 1)]MS_w} . \quad (9)$$

If $n_1 = n_2 = . . . = n_G = n = N/G$, then formula (9) reduces to

$$\hat{\rho}_{intraclass} = \frac{(N - 1)(MS_{tot} - MS_w)}{(N - 1)MS_{tot} - (n - 1)MS_w} . \quad (10)$$

Thus, though its numerator is the same as that for eta squared and epsilon squared, its denominator is smaller. This occurs because of the way that $\rho_{intraclass}$ is defined via components of variance for a random-effects ANOVA. The other three statistics are for the fixed-effects situation only. Because errors of measurement apply to repeated testing of the same examinees, however, rather than to sampling fluctuations, the correction for attenuation procedure is similar. A random sample of examinees would receive a random sample of factor levels. Repeated testing (perhaps conceptual rather than actual) would be with the same examinees but with different factor levels.

References

Fisher, Ronald A. Statistical methods for research workers. (1st ed.) Edinburgh: Oliver and Boyd, 1925.

Glass, Gene V, and Hakstian, A. Ralph. Measures of association in comparative experiments: their development and interpretation. American Educational Research Journal, 1969, 6, 403-414.

Harris, J. A. On the calculation of intra-class and inter-class coefficients of correlation from class moments when the number of possible combinations is large. Biometrika, 1913, 9, 446-472.

Hays, William L. Statistics for psychologists. New York: Holt, Rinehart and Winston, 1963.

Kelley, Truman L. An unbiased correlation ratio measure. Proceedings of the National Academy of Sciences, 1935, 21, 554-559.

Pearson, Karl. Mathematical contributions to the theory of evolution.--On homotyposis in homologous but differentiated organs. Proceedings of the Royal Society, 1903, 71, 288-313. See the footnote on pages 303, 304.

Pearson, Karl. Mathematical contributions to the theory of evolution.--XIV. On the general theory of skew correlation and non-linear regression. Drapers' Company Research Memoirs, Biometric Series, II, 1905.

Pearson, Karl. On a correction to be made to the correlation ratio η . Biometrika, 1911-1912, 8, 254-256.

Pearson, Karl. On the correction necessary for the correlation ratio η . Biometrika, 1922-23, 14, 412-417.

Stanley, Julian C. Reliability. Ch. 13 in Robert L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

"Student" (W. S. Gosset). The correction to be made to the correlation ratio for grouping. Biometrika, 1913, 316-320.