

## Measurement Error

*Definitions.* For two variables, X and Y, the following hold:

Parameter	Explanation
$E(X) = \frac{\sum X_i}{N} = \mu_x$	Expectation, or Mean, of X
$V(X) = E[(X - \mu_x)^2] = \sigma_x^2$	Variance of X
$SD(X) = \sqrt{V(X)} = \sigma_x$	Standard Deviation of X
$COV(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = \sigma_{xy}$	Covariance of X and Y
$CORR(X, Y) = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = r_{xy} = r_{yx}$	Correlation of X and Y
$\beta_{yx} = \frac{\sigma_{xy}}{\sigma_x^2}$	Slope coefficient for the Bivariate regression of Y on X (Y dependent)

*Question:* Suppose X suffers from random measurement error - that is, the values of X that we observe differ randomly from the true values that we are interested in. For example, we might be interested in income. Since people do not remember their income exactly, reported income will sometimes be higher and sometimes be lower than true income. In such a case, how does random measurement error affect the various statistical measures we are typically interested in? That is, how does unreliability affect our statistical measures and conclusions?

*Revised Question:* Let us put the question more formally. Let  $X = X_t + \varepsilon$ , where  $\varepsilon$  is a random error term (i.e. has mean 0 and variance  $s_\varepsilon^2$ ). That is,  $X_t$  is the “true” value of the variable, and X is the flawed measure of the variable that is observed. We want to see how the statistics for the observed variable, X, differ from the statistics for the true variable,  $X_t$ . When thinking about this question, keep in mind that, because  $\varepsilon$  is a random error term, it is independent from all other variables (except itself), e.g.  $COV(X, \varepsilon) = COV(Y, \varepsilon) = 0$ .

*Definition of Reliability:* The reliability of a variable is defined as:

$$REL(X) = \frac{\sigma_{x_t}^2}{\sigma_x^2} = r_{x_t X}^2$$

The first equality says reliability is true variance divided by total variance. The second equality says the reliability of a variable is the squared correlation between the true value of the variable and the observed value that suffers from random measurement error. If there is no random measurement error, reliability = 1.

*Some additional rules for expectations.* Before answering the question, the following additional rules are helpful. Let A, B, C, and D be random variables. Then,

- (1)  $E(A + B) = E(A) + E(B)$
- (2) If A and B are independent,  $V(A + B) = V(A) + V(B)$
- (3)  $COV(A + B, C + D) = COV(A, C) + COV(A, D) + COV(B, C) + COV(B, D)$

## Effects of Unreliability

A. For the mean:

$$E(X) = E(X_t + \varepsilon) = E(X_t) + E(\varepsilon) = E(X_t) \quad [\text{Expectations rule 1}]$$

NOTE: Remember, since errors are random,  $\varepsilon$  has mean 0.

*Implication:* Random measurement error does not bias the expected value of a variable - that is,  $E(X) = E(X_t)$

B. For the variance:

$$V(X) = V(X_t + \varepsilon) = V(X_t) + V(\varepsilon) \quad [\text{Expectations rule 2}]$$

NOTE: Remember,  $\text{COV}(X_t, \varepsilon) = 0$  because  $\varepsilon$  is a random disturbance.

*Implication:* Random measurement error does result in biased variances. The variance of the observed variable will be greater than the true variance.

C. For the covariance (we'll let  $Y_t$  stand for the perfectly measured Y variable):

$$\begin{aligned} \text{COV}(X, Y_t) &= \text{COV}(X_t + \varepsilon, Y_t) = \text{COV}(X_t, Y_t) + \text{COV}(\varepsilon, Y_t) \\ &= \text{COV}(X_t, Y_t) \quad [\text{Expectations rule 3}] \end{aligned}$$

NOTE: Remember,  $\text{COV}(\varepsilon, Y_t) = 0$  because  $\varepsilon$  is a random disturbance.

*Implication:* Covariances are not biased by random measurement error.

D. For the correlation:

$$r_{xy_t} = \frac{\sigma_{XY_t}}{\sigma_X \sigma_{Y_t}}, \quad r_{x_t y_t} = \frac{\sigma_{X_t Y_t}}{\sigma_{X_t} \sigma_{Y_t}} = \frac{\sigma_{XY_t}}{\sigma_{X_t} \sigma_{Y_t}}$$

Thus, when X and  $Y_t$  covary positively,  $\text{CORR}(X, Y_t) \leq \text{CORR}(X_t, Y_t)$

*Implication:* Random measurement error produces a downward bias in the bivariate correlation. This is often referred to as attenuation.

E. For  $\beta_{Y_t X}$ : ( $Y_t$  is perfectly measured, X has random measurement error)

$$\beta_{Y_t X} = \frac{\sigma_{XY_t}}{\sigma_X^2}, \quad \beta_{Y_t X_t} = \frac{\sigma_{X_t Y_t}}{\sigma_{X_t}^2} \quad \text{Thus, when X and } Y_t \text{ covary positively, } \beta_{Y_t X} \leq \beta_{Y_t X_t}$$

*Implication:* Random measurement error in the Independent variable produces a downward bias in the bivariate regression slope coefficient.

F. For  $\beta_{Y X_t}$ : (Now Y is measured with random error, while  $X_t$  is measured perfectly)

$$\beta_{Y X_t} = \frac{\sigma_{X_t Y}}{\sigma_{X_t}^2}, \quad \beta_{Y_t X_t} = \frac{\sigma_{X_t Y_t}}{\sigma_{X_t}^2}. \quad \text{Thus, } \beta_{Y X_t} = \beta_{Y_t X_t}$$

*Implication:* Random measurement error in the Dependent variable does not bias the slope coefficient. HOWEVER, it does lead to larger standard errors. Recall that the formula for the standard error of b is

$$s_b = \sqrt{\frac{1 - R^2}{(N - K - 1)}} * \frac{s_Y}{s_X}$$

When you have random measurement error in Y,  $R^2$  goes down because of the previously noted downward bias. This increases the numerator. Also, the variance of Y goes up, which further increases the standard error.

### *Additional implications*

- When you have more than one independent variable, random measurement error can cause coefficients to be biased either upward or downward. As you add more variables to the model, all you can really be sure of is that, if the variables suffer from random measurement error (and most do) the results will probably be at least a little wrong!
- Reliability is a function of both the total variance and the error variance. True variance is a population characteristic, error variance is a characteristic of the measuring instrument. As Wiley and Wiley note, an item rarely has a “unique” reliability. Any social process that tends to increase consensus results in a reduction in the true variability in attitudes without necessarily affecting the characteristics of the measuring instrument. An increase in reliability can always be achieved by conducting measurements on a more heterogeneous population.
- Comparisons of any sort can be distorted by differential reliability of variables. For example, if comparing effects of two variables, one variable may appear to have a stronger effect simply because it is better measured. If comparing, say, husbands and wives, the spouse who gives more accurate information may appear more influential. For a more detailed discussion of how measurement error can affect group comparisons, see Thomson, Elizabeth and Richard Williams (1982) “Beyond wives family sociology: a method for analyzing couple data” *Journal of marriage and the family* Vol 44 999:1008
- The fact that reliabilities differ between groups does not necessarily mean that one group is more “accurate.” It may just mean that there is less true variance in one group than there is in another.
- Examining error variances is therefore often more informative than just looking at reliabilities.

*Dealing with measurement error.* For the most part, this is a subject for a research methods class or a more advanced statistics class. I’ll toss out a few ideas for now:

- Collect better quality data in the first place. Make questions as clear as possible.
- Measure multiple indicators of concepts. When more than one question measures a concept, it is possible to estimate reliability and to take corrective action. For a more detailed discussion on measuring reliability, see Reliability and Validity Assessment, by Edward G. Carmines and Richard A. Zeller. 1979. Paper # 17 in the Sage Series on Quantitative Applications in the Social Sciences. Beverly Hills, CA: Sage.
- Create scales from multiple indicators of a concept. The scales will generally be more reliable than any single item would be. In SPSS you might use the FACTOR or RELIABILITY commands; in Stata relevant commands include `factor` and `alpha`.
- Use advanced techniques, such as LISREL, which let you incorporate multiple indicators of a concept in your model. Ideally, LISREL “purges” the items of measurement error hence producing unbiased estimates of structural parameters.