Instrumental Measurement Errors, Their Sources and Remedies

Ashutosh Sharma[1], Nitin Kumar[1], Adarsh Kumar[2]

[1] Scholar Mechanical Engineering Global College of Engineering & Technology, Kahnpur Khui Punjab
[2]HOD Mechanical Engineering Department Global College of Engineering & Technology, Kahnpur Khui Punjab,

Abstract

Errors of measurement arise because our observations are affected by many sources of variability, but our conceptual frameworks necessarily ignore much of this variability. Sources of variability that are not included in our models and descriptions of phenomena are treated as error. A good theory of error supports the development of precise measurements, clearly defined constructs and sound public policy. Narrowly defined constructs that do not generalize much beyond the observed performances do not involve many sources of error, but constructs that generalize observed scores over a broad range of conditions of observation necessarily involve many potential sources of error. We can have narrow constructs with small errors or more broadly defined constructs with larger errors. Some errors that are negligible for individuals can have a substantial impact on estimates of group performance, and therefore, can have serious consequences.

Keywords: measurement, errors, control

1. INTRODUCTION

Paradoxically, errors of measurement do not exist, but they are essential. As illustrated below, there is nothing about a single test score or a pair of scores that implies the presence of errors of measurement. However, if two scores are taken to be measures of the same variable for the same person, we expect them to be equal, and if they are not equal, our data are inconsistent with our conceptual framework. We can resolve this dilemma by assuming that one or both of the measurements contain errors. Errors of measurement play a vital role in quantitative analyses, by making it possible to model data without immediately running into inconsistencies. The true score theory is a good simple model for measurement, but it may not always be an accurate reflection of reality. In particular, it assumes that any observation is composed of the true value plus some random error value. But is that reasonable? What if all error is not random? Isn't it possible that some errors are systematic, that they hold across most or all of the members of a group? One way to deal with this notion is to revise the simple true score model by dividing the error component into two subcomponents, **random error** and **systematic error[1]**. Here, we'll look at the differences between these two types of errors and try to diagnose their effects on our research.

1.1 Types of Error

No measurement can be made with perfect accuracy, but it is important to find out what the accuracy actually is and how different errors have entered into the measurement. A study of errors is a first step in finding ways to reduce them. Such a study also allows us to determine the accuracy of the final test result. Errors may come from different sources and are usually classified under three main headings:

(a) **Gross Errors** Largely human errors, among them misreading of instruments, incorrect adjustment and improper application of instruments, and computational mistakes.
(b) **Systematic Errors** Shortcomings of the instruments, such as defective or worn parts, and effects of the environment on the equipment or the user.
(c) **Random Errors** Those due to causes that cannot be directly established because of random variations in the parameter or the system of measurement.

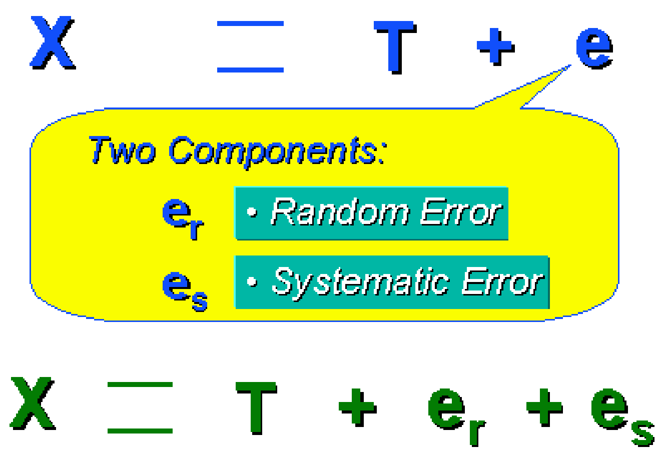

Figure 1 Measurement Errors

1.2 Random Error

Random error is caused by any factors that randomly affect measurement of the variable across the sample. For instance, each person's mood can inflate or deflate their performance on any occasion[2]. In a particular testing, some children may be feeling in a good mood and others may be depressed. If mood affects their performance on the measure, it may artificially inflate the observed scores for some children and artificially deflate them for others. The important thing about random error is that it does not have any consistent effects across the entire sample. Instead, it pushes observed scores up or down randomly. This means that if we could see all of the random errors in a distribution they would have to sum to 0 -- there would be as many negative errors as positive ones. The important property of random error is that it adds variability

ISSN: 2229-6913(Print), ISSN: 2320-0332(Online), UGC Approved Journal (S.No.63019)

to the data but does not affect average performance for the group. Because of this, random error is sometimes considered **noise.**

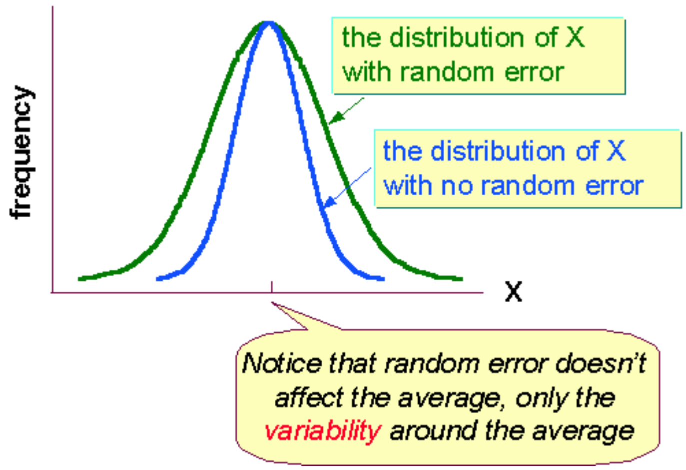

Figure 2 Random Errors

1.2.1 Sources of random error

The random or stochastic error in a measurement is the error that is random from one measurement to the next. Stochastic errors tend to be normally distributed when the stochastic error is the sum of many independent random errors because of the central limit theorem. Stochastic errors added to a regression equation account for the variation in *Y* that cannot be explained by the included *X*s.

1.2.2 Surveys

The term "observational error" is also sometimes used to refer to response errors and some other types of non-sampling error. In survey-type situations, these errors can be mistakes in the collection of data, including both the incorrect recording of a response and the correct recording of a respondent's inaccurate response.

1.3 Systematic Error

Systematic error is caused by any factors that systematically affect measurement of the variable across the sample. For instance, if there is loud traffic going by just outside of a classroom where students are taking a test, this noise is liable to affect all of the children's scores -- in this case, systematically lowering them. Unlike random error, systematic errors tend to be consistently either positive or negative -- because of this; systematic error is sometimes considered to be *bias* in measurement.

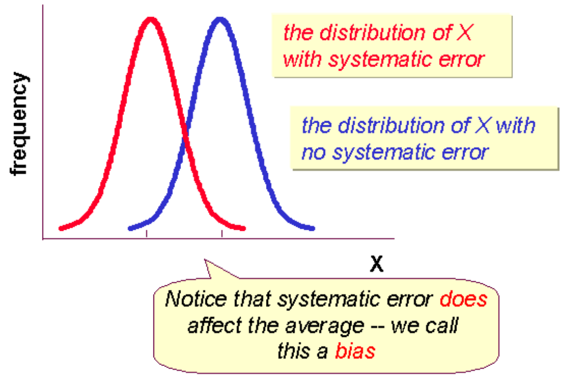

Figure 3 Systematic Errors

1.1.1 1.3.1Sources of systematic error

1.1.2 1.3.1.1 Imperfect calibration

Sources of systematic error may be imperfect calibration of measurement instruments (zero error), changes in the environment which interfere with the measurement process and sometimes imperfect methods of observation can be either zero error or percentage error. If you consider an experimenter taking a reading of the time period of a pendulum swinging past a fiducially marker: If their stop-watch or timer starts with 1 second on the clock then all of their results will be off by 1 second (zero error). If the experimenter repeats this experiment twenty times (starting at 1 second each time), then there will be a percentage error in the calculated average of their results; the final result will be slightly larger than the true period.

Distance measured by radar will be systematically overestimated if the slight slowing down of the waves in air is not accounted for. Incorrect zeroing of an instrument leading to a zero error is an example of systematic error in instrumentation.

Systematic errors may also be present in the result of an estimate based upon a mathematical model or physical law [4]. For instance, the estimated oscillation frequency of a pendulum will be systematically in error if slight movement of the support is not accounted for.

1.1.3 1.3.1.2 Quantity

Systematic errors can be either constant, or related (e.g. proportional or a percentage) to the actual value of the measured quantity, or even to the value of a different quantity (the reading of a ruler can be affected by environmental temperature). When it is constant, it is simply due to incorrect zeroing of the instrument. When it is not constant, it can change its sign. For instance, if a thermometer is affected by a proportional systematic error equal to 2% of the actual temperature, and the actual temperature is 200°, 0°, or −100°, the measured temperature will be 204° (systematic error = +4°), 0° (null systematic error) or −102° (systematic error = −2°), respectively. Thus, the temperature will be overestimated when it will be above zero, and underestimated when it will be below zero.

1.1.4 1.3.1.3 Drift

Systematic errors which change during an experiment (drift) are easier to detect. Measurements indicate trends with time rather than varying randomly about a mean. Drift is evident if a measurement of a constant quantity is repeated several times and the measurements drift one way during the experiment. If the next measurement is higher than the previous measurement as may occur if an instrument becomes warmer during the experiment then the measured quantity is variable and it is possible to detect a drift by checking the zero reading during the experiment as well as at the start of the experiment (indeed, the zero reading is a measurement of a constant quantity). If the zero reading is consistently above or below zero, a systematic error is present. If this cannot be eliminated, potentially by resetting the instrument immediately before the experiment then it needs to be allowed by subtracting its (possibly time-varying) value from the readings, and by taking it into account while assessing the accuracy of the measurement. If no pattern in a series of repeated measurements is evident, the presence of fixed systematic errors can only be found if the measurements are checked, either by measuring a known quantity or by comparing the readings with readings made using a different apparatus, known to be more accurate. For example, if you think of the timing of a pendulum using an accurate stopwatch several times you are given readings randomly distributed about the mean. A systematic error is present if the stopwatch is checked against the 'speaking clock' of the telephone system and found to be running slow or fast. Clearly, the pendulum timings need to be corrected according to how fast or slow the stopwatch was found to be running. Measuring instruments such as ammeters and voltmeters need to be checked periodically against known standards. Systematic errors can also be detected by measuring already known quantities. For example, a spectrometer fitted with a diffraction grating may be checked by using it to measure the wavelength of the D-lines of the sodium electromagnetic spectrum which are at 600 nm and 589.6 nm. The measurements may be used to determine the number of lines per millimeter of the diffraction grating, which can then be used to measure the wavelength of any other spectral line. Constant systematic errors are very difficult to deal with as their effects are only observable if they can be removed. Such errors cannot be removed by repeating measurements or averaging large numbers of results. A common method to remove systematic error is through calibration of the measurement instrument.

2 COMMON ERRORS IN RESEARCH PROCESS

2.1 Population Specification

This type of error occurs when the researcher selects an inappropriate population or universe from which to obtain data. Example: Packaged goods manufacturers often conduct surveys of housewives, because they are easier to contact, and it is assumed they decide what is to be purchased and also do the actual purchasing. In this situation there often is population specification error. The husband may purchase a significant share of the packaged goods, and have significant direct and indirect influence over what is bought. For this reason, excluding husbands from samples may yield results targeted to the wrong audience.

2.2 Sampling

Sampling error occurs when a probability sampling method is used to select a sample, but the resulting sample is not representative of the population concern. Unfortunately, some element of sampling error is unavoidable. This is accounted for in confidence intervals, assuming a probability sampling method is used. Example: Suppose that we collected a random sample of 500 people from the general U.S. adult population to gauge their entertainment preferences. Then, upon analysis, found it to be composed of 70% females. This sample would not be representative of the general adult population and would influence the data. The entertainment preferences of females would hold more weight, preventing accurate extrapolation to the US general adult population. Sampling error is affected by the homogeneity of the population being studied and sampled from and by the size of the sample.

2.3 Selection

Selection error is the sampling error for a sample selected by a non probability method. Example: Interviewers conducting a mall intercept study have a natural tendency to select those respondents who are the most accessible and agreeable whenever there is latitude to do so. Such samples often comprise friends and associates who bear some degree of resemblance in characteristics to those of the desired population.

2.4 Non-responsive

Non response error can exist when an obtained sample differs from the original selected sample.
Example: In telephone surveys, some respondents are inaccessible because they are not at home for the initial call or call-backs. Others have moved or are away from home for the period of the survey. Not-at-home respondents are typically younger with no small children, and have a much higher proportion of working wives than households with someone at home. People who have moved or are away for the survey period have a higher geographic mobility than the average of the population. Thus, most surveys can anticipate errors from non-contact of respondents. Online surveys seek to avoid this error through e-mail distribution, thus eliminating not-at-home respondents.

2.5 Measurement

Measurement error is generated by the measurement process itself, and represents the difference between the information generated and the information wanted by the researcher. Example: A retail store would like to assess customer feedback from at-the-counter purchases. The survey is developed but fails to target those who purchase in the store. Instead, results are skewed by customers who bought items online.

3 MEASUREMENT ERRORS AND BIAS

Epidemiological studies measure characteristics of populations. The parameter of interest may be a disease rate, the prevalence of an exposure, or more often some measure of the association between an exposure and disease. Because studies are carried out on people and have all the attendant practical and ethical constraints, they are almost invariably subject to bias.

3.1 Selection bias

Selection bias occurs when the subjects studied are not representative of the target population about which conclusions are to be drawn. Suppose that an investigator wishes to estimate the prevalence of heavy alcohol consumption (more than 21 units a week) in adult residents of a city. He might try to do this by selecting a random sample from all the adults registered with local general practitioners, and sending them a postal questionnaire about their drinking habits. With this design, one source of error would be the exclusion from the study sample of those residents not registered with a doctor. These excluded subjects might have different patterns of drinking from those included in the study. Also, not all of the subjects selected for study will necessarily complete and return questionnaires, and non-responders may have different drinking habits from those who take the trouble to reply. Both of these deficiencies are potential sources of selection bias. The possibility of selection bias should always be considered when defining a study sample. Furthermore, when responses are incomplete, the scope for bias must be assessed. The problems of incomplete response to surveys are considered further in.

3.2 Information bias

The other major class of bias arises from errors in measuring exposure or disease. In a study to estimate the relative risk of congenital malformations associated with maternal exposure to organic solvents such as white spirit, mothers of malformed babies were questioned about their contact with such substances during pregnancy, and their answers were compared with those from control mothers with normal babies. With this design there was a danger that "case" mothers, who were highly motivated to find out why their babies had been born with an abnormality, might recall past exposure more completely than controls. If so, a bias would result with a tendency to exaggerate risk estimates. Another study looked at risk of hip osteoarthritis according to physical activity at work, cases being identified from records of admission to hospital for hip replacement. Here there was a possibility of bias because subjects with physically demanding jobs might be more handicapped by a given level of arthritis and therefore seek treatment more readily. Bias cannot usually be totally eliminated from epidemiological studies. The aim, therefore, must be to keep it to a minimum, to identify those biases that cannot be avoided, to assess their potential impact, and to take this into account when interpreting results. The motto of the epidemiologist could well be "dirty hands but a clean mind"

4 REDUCING MEASUREMENT ERRORS

So, how can we reduce measurement errors, random or systematic? One thing you can do is to pilot test your instruments, getting feedback from your respondents regarding how easy or hard the measure was and information about how the testing environment affected their performance. Second, if you are gathering measures using people to collect the data (as interviewers or observers) you should make sure you train them thoroughly so that they aren't inadvertently introducing error. Third, when you collect the data for your study you should double-check the data thoroughly. All data entry for computer analysis should be "double-punched" and verified. This means that you enter the data twice, the second time having your data entry machine check that you are typing the exact same data you did the first time. Fourth, you can use statistical procedures to adjust for measurement error. These range from rather simple formulas you can apply directly to your data to very complex modeling procedures for modeling the error and its effects. Finally, one of the best things you can do to deal with measurement errors, especially systematic errors, is to use multiple measures of the same construct. Especially if the different measures don't share the same systematic errors, you will be able to **triangulate** across the multiple measures and get a more accurate sense of what's going on.

5 TOP TEN WAYS TO REDUCE SURVEY MEASUREMENT ERROR

1. Follow basic administrative guidelines
2. Clarify the “central players” in the region and nationally and be certain to consider ways to work with them and reduce the chance of “spoilers”
3. Conduct pre testing for all questionnaires
4. Hire relatively large numbers of interviewers, whom should be tested in the course of training, while setting high goals and providing rewards for success
5. Interviewers should be assigned using interpenetrating sampling techniques
6. Consider all potential errors of non observation, including sampling, coverage and non response
7. Include questions that allow ex post identification of different types of error of measurement
8. Carefully evaluate whether there are systematic non response patterns that might affect interpretation of findings
9. Design clear guidelines for filling missing data, preferably using interviewer teams done shortly after each day of data collection

10. Attempt to compare results with those that might be obtainable from routine statistics or other different data sources

REFERENCES

1 Errors of Measurement in Statistics, W. G. Cochran, Technometrics, Vol. 10, No. 4 (Nov., 1968), pp. 637– 666:http://www.jstor.org/stable/1267450

2 Bland, J. Martin, and Douglas G. Altman. "Statistics notes: measurement error." Bmj 313.7059 (1996): 744.

3 John R. Taylor, An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, 2d Edition, University Science Books, 1997

4 Biemer, P. P., and Fecso, R. S. (1995), "Evaluating and Controlling Measurement Error in Business Surveys," in Business Survey Methods, eds. B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Colledge and P. S. Kott, New York: Wiley-Interscience, pp. 257-281