



F30 Cross-Disciplinary Communication: Understanding Commonly Used Terms in Statistics

Glinda S. Cooper, PhD, Innocence Project, 40 Worth Street, Ste 701, New York, NY 10013; Sarah Chu, MS, Innocence Project, 40 Worth Street, Ste 701, New York, NY 10013; and Kareem Belt, MS, Innocence Project, 40 Worth Street, Ste 701, New York, NY*

The goal of this presentation is to explain the meaning of commonly used statistical terms relating to error, specifically, measurement error; in addition, the ways in which these definitions may differ from those for these same terms used in different contexts or by different professions will be highlighted. Key concepts and the relations between these concepts that will be examined include: (1) random error; (2) systematic (non-random) error or bias; (3) reliability; (4) validity; (5) accuracy; and, (6) precision.

This presentation will impact the forensic science community by increasing the criminal justice system stakeholders' understanding of important statistical concepts and terminology and by improving communication among people with different backgrounds and training. As the work of forensic science expands beyond forensic practitioners, judges, and attorneys to include statisticians and researchers with expertise in adjacent fields, it can be expected that conflicts of terminology will arise more often. Thus, this presentation seeks to facilitate understanding and to better integrate these different communities as they focus on the shared goal of improving forensic science.

There is growing recognition that terminology between forensic science disciplines needs to be standardized as disciplines often develop specialized terminology for their field or specific meanings for words with more general usage. The way terms are understood across groups must also be improved. Criminal justice system stakeholders may associate meanings with terms that people outside their discipline or expertise may interpret in very different ways. These differences can lead to miscommunication, particularly with statistical terminology relating to measurement error, as the specific meanings of some commonly used terms are often misinterpreted. Adding to the difficulty in communication is that different terms can be used for the same concept, and that concepts can be both distinct and interrelated.

One important concept is the difference between random error and systematic error. Both of these types of error must be considered for any kind of measurement, including categorical classification outcomes (i.e., "included" or "excluded") and continuous variables. Random error is "directionless," that is, equally likely to result in over- and under-estimations of the true value. Random error influences the precision and reliability of a measure; reliability refers to the repeatability or consistency of a measure, or how much variation one would see in a measure or test that was conducted multiple times.

Systematic or non-random error refers to a consistent inaccuracy that leads to a "biased" result. Bias in this context refers to an error that occurs in a particular direction (i.e., an error that leads to an over-estimation, or an error that leads to an under-estimation). Note that "bias" as used in this statistical context does not convey the connotation that this term can imply in more general usage (i.e., prejudiced or unfair). Systematic error will influence the accuracy and validity of a measure; validity refers to how well a measure reflects the "true" value of a test or measure used in a specific population and conditions.

In general usage, the terms reliability and validity are sometimes used interchangeably, and it is sometimes said that reliability ensures validity; however, in statistical usage these terms refer to different concepts, and it is important to note that a test that is reliable may or may not be valid. A common analogy used to describe these



Jurisprudence - 2017

terms uses darts that are thrown onto a dartboard. If all of the darts cluster closely around the center of the dartboard, the “test” (the person throwing the darts) is both precise (reliable), with little random error, and accurate (valid), with little systematic error. But if the darts cluster in the lower right-hand corner of the dartboard, the “test” would be reliable (precise) but inaccurate (biased or invalid); there is little variation in the results, but the results are off the mark.

Reliability, Validity, Bias