Families in Society: The Journal of Contemporary Social Services
1–8

DOI: 10.1177/1044389420913415
journals.sagepub.com/home/fis
SAGE

Research at Work: Sampling, Central Tendency, and Causation

Jeremiah W. Jaggers[1] and Alysse M. Loomis[2]

Abstract

A basic understanding of statistics is an important piece of social work practice, informing the ability of social workers to interpret and evaluate existing research and to conduct evaluations of programs in their agencies. This article gives an overview of basic statistical concepts of variables, central tendency, and correlation versus causation to give social workers a preliminary foundation to critically evaluate and utilize basic research. The basic statistics in the study "Relationship Between the Five Facets of Mindfulness on Mood and Substance Use Relapse" by Temme and Wang (2018) are unpacked as a way to bring to life the statistical concepts discussed throughout the article.

Keywords

statistics, practice evaluation, research, causation

Manuscript received: February 17, 2020; Accepted: February 20, 2020

Disposition editor: Sondra J. Fogel

Introduction

Evaluation is a critical component of evidence-based practice. Social work practitioners undergo extensive training in preparation for professional practice, including training on research methodology. Part of conducting successful evaluation is ongoing education and training, which includes a fundamental understanding of statistics. What follows is an overview of important statistical concepts, including a review of *variables, central tendency*, and *correlation versus causation*. We use a published manuscript to provide a sample that can be referred to for examples of each concept.

Sample Article

The article selected as a sample is "Relationship Between the Five Facets of Mindfulness on Mood and Substance Use Relapse" by Temme and Wang (2018). In the article, the authors describe their research approach and findings. Temme and Wang (2018) describe the relationship between five specific components of mindfulness-based practice and (a) mood and (b) substance use. The article provides a description of each component along with sample items that were asked of study participants. The article concludes with a section describing the study's implications for social work practice. It is suggested that you review the article before proceeding. The full citation may be found in the References section of this article.

[1]PhD, MSW, assistant professor, The University of Utah, Salt Lake City, USA
[2]PhD, LCSW, assistant professor, The University of Utah, Salt Lake City, USA

Corresponding Author:
Jeremiah W. Jaggers, assistant professor, College of Social Work, The University of Utah, 395 S. 1500 E., Salt Lake City, UT 84112, USA.
Email: jeremiah.jaggers@utah.edu

General Overview

Variables

Oftentimes, the first section reported in the results of a research article is that of *descriptive statistics*, which as the name implies is a section in which the researcher describes the variables that were measured within their sample. To understand what is being presented, it can be helpful to understand the different *types* of variables that are typically discussed. In Temme and Wang, the variables measured were the five facets of mindfulness (observing, describing, acting with awareness, nonjudging, and nonreactivity), warning signs of relapse, and mood.

Categorical variables are those where distinct categories are being measured but do not represent any type of numerical measurement. Temme and Wang (2018) measure several categorical variables, including sex (male and female) and race/ethnicity (African American, Hispanic, Mixed, and White) in which they are indicating which *category* individuals in their study identified that they fit into.

There are two types of categorical variables. The examples above represent *nominal categorical variables*, which means that there is no meaning in the order of the variables. *Ordinal* variables are similar to nominal categorical variables in that there are specific categories that are being measured; what is different is that with ordinal variables there is an order to the categories (such as *low, medium*, and *high*). In ordinal variables, there is no specific distance between the categories, just as there is no specific measurable distance between *low* and *medium.* For example, Temme and Wang (2018) ask questions about mood using questions based on a Likert-type scale, where participants are asked whether they felt a certain mood (such as friendly) over the past week and are given five choices from 0 to 5, where 0 = *never* and 4 = *extremely*. This is an example of an ordinal variable considering there is not a specific distance between never and extremely; however, it is still clear that never is less than extremely.

Dichotomous variables are those that take only one of two values. Measuring sex as male/female is one example of a dichotomous variable. Dichotomous variables can also be used to illustrate the presence/absence of something being measured. For example, suppose Temme and Wang (2018) had measured *whether or not* an individual had relapsed as opposed to warning signs of relapse, this would be a dichotomous variable where yes indicates that a relapse occurred and no indicates that a relapse did not occur.

Continuous variables are numerical variables that represent an infinite number of options. As an example, when measuring age, one can measure in years, months, days, minutes, and so on and come up with an infinite number of options to represent age. In Temme and Wang (2018), mood, warning signs of relapse, and the facets of mindfulness are reported as continuous variables. The individual questions related to warning signs of relapse from the Advance Warning of Relapse (AWARE) were captured as ordinal variables (using a Likert-type scale) but were then added together to create a sum score, or continuous variable where scores ranged from 28 to 140. This final score represents warning signs of relapse where higher scores indicate higher risk of relapse and the distance between scores is easy to understand (for example, the distance from 28 to 38 is the same as the distance from 38 to 48).

Central Tendency

Mean, median, mode. Measures of central tendency help to understand the data by describing the "middle" of the data in a number of ways. The mode is the most frequent score, which in the case of age would be the most frequently reported age. The median is the score where 50% of scores fall above and 50% fall below. In the example of age, it would be the age at which about 50% of other participants in the study were younger and about 50% of participants were older. The mean, or average, or a variable is the most commonly reported measure of central tendency. In

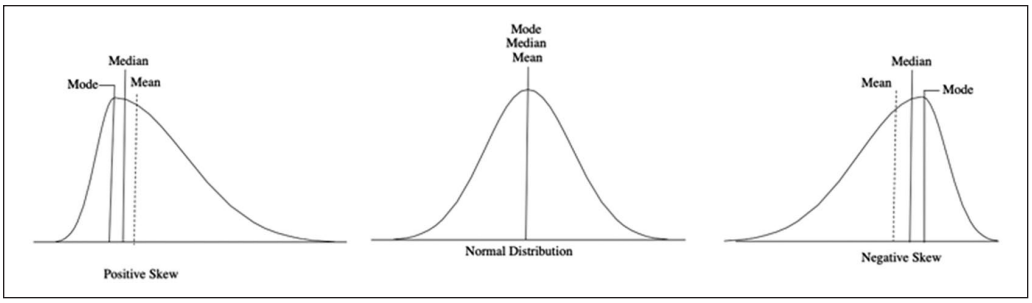

Figure 1. Normal and skewed distribution.

Temme and Wang's (2018) study, the mean age, or average age, of participants is 36.9, meaning that people are on average about 37 years old.

Standard deviation. Standard deviation is a way to describe how much variability there is in data and is often reported alongside the mean. Standard deviation is a measure of how much variance there is from the mean. If the standard deviation is low then the individual scores tend to be close to the mean and if it is large, then the individual scores tend to be further away from the mean (or more spread apart). Standard deviations give information about how spread out numbers are from the mean, about 68% of values are within 1 standard deviation from the mean, 95% are within 2 standard deviations, and 99.7% are within 3 standard deviations from the mean. Looking at Temme and Wang (2018), the mean age of participants is 36.9 and the standard deviation is 10.8. This indicates that 68% of participants are within 10.8 years from the mean age of 36.9 years, or between 26.1 and 47.7. If the standard deviation was much smaller, 2.8 for example, then this would suggest that the majority of participants are very close to the mean age; in this example, 68% of participants would fall between 34.1 and 39.7.

Distribution

Normal. When data have a normal distribution, this means that the mean, median, and mode are all the same (Abu-Bader, 2011). Figure 1a shows an example of data with a normal distribution; as can be seen by the image, normal distributions are often referred to as a "bell curve." As can be seen in the figure, about 50% of scores fall above the mean and about 50% fall below the mean. For many advanced statistical tests, there are assumptions that the data being used is normally distributed, making it important for researchers to share whether the data are normally distributed or skewed so that readers can determine whether the test being used was appropriate.

Skew. As noted above, in normal data there are a similar amount of cases (often people) above and below the average of a given variable. With some variables, however, researchers will find that the data is skewed, or that there are more people represented above (or alternately below) the average. Temme and Wang (2018) state "negative mood and risk for relapse were slightly skewed in a positive direction, showing that there was greater than average scores in this area" (p. 213). This shows that in their data, there were more people who reported negative mood and higher relapse than those on the lower end. If the data were skewed negatively, this would mean that there were *less* people who reported negative mood and higher relapse than those on the higher end.

Correlation Versus Causation

Correlation and causation are ways of demonstrating how the variables that were measured in the study relate to *each other*. Perhaps the

most common error in research is conflating correlation and causation, which is when two variables that are related to each other are implied to have a cause and effect relationship (where one variable is thought to cause the other). While the two concepts are related, correlation is far more common because it is easier to calculate and requires fewer steps.

In looking for how variables are related to each other, variables are commonly discussed in two distinct manners—as being independent or dependent. Independent variables (often referred to using the letter X) are said to influence or change values of a dependent variable (often referred to using the letter Y). Temme and Wang hypothesized that "non-judgement and discrimination would predict mood and awareness of relapse" (p. 211); thus, they were situating two facets of mindfulness (nonjudgment and discrimination) as independent variables that they hypothesized would predict mood and relapse (dependent variables). However, the assumption that X *causes* Y is where confusion often arises when determining whether a causal relationship exists. The following section provides a succinct overview of correlation, causation, and their differences.

Correlation

A statistical correlation shows how two variables change with respect to one another (cf. Sedgwick, 2012) and also provides information about how strongly the two variables relate to each other and in what direction. One of the most common correlations, though definitely not the only one, is the Pearson's product moment correlation, also known as Pearson's correlation (example discussed below). A Pearson's correlation requires the use of two continuous variables. Pearson's correlation is represented using the letter r. Values for r range from −1 to 1. The closer the r value is to 1, the stronger the relationship, regardless of the direction of the relationship. Positive and negative values are used to indicate the direction of the relationship. A positive value (e.g., $r = .3$) suggests that two variables increase together. A negative value (e.g., $r = -.7$) indicates the relationship is inverse—as one increases the other decreases. Comparing a correlation of .3 to −.7, the latter correlation (−.7) is the strongest because it has the largest absolute value. A value of 1 or −1 represents a perfect correlation, where the change in one variable is exactly proportional to change in another variable. Perfect correlations are rare and are typically attributed to a mathematical or methodological error. Similarly, a value of 0 means that the two variables have no correlated relationship.

Temme and Wang (2018) report correlational coefficients from their study on page 214. Their Table 3 shows the Pearson's correlation for each of the aforementioned five factors. The correlation between "Observe" and "Describe" is positive and statistically significant ($r = .54$). These results indicate that when *observing* one's own behavior using mindfulness techniques, individuals are better able to *describe* their own feelings. Positive correlations signal that the two constructs either increase or decrease together; for example, as scores on observing increase, scores on describing also tend to increase (or vice versa). This is known as a direct correlation. Alternatively, negative r values are known as an inverse correlation. Inverse correlations signal that two constructs move in opposite directions of one another. This can be seen in Table 3 where "Observe" and "Act With Awareness" are inversely correlated ($r = -.38$). This is interpreted to mean that an increase in *observing* one's own behaviors using mindfulness techniques results in a decrease in *acting* without thought.

Causation

Causality "assumes that the value of an interdependent variable is the reason for the value of a dependent variable (Allen, 2017)." That is, X is the reason Y happens. Unlike correlation, causation can only be explained by meeting certain rigorous criteria. Temme and Wang (2018) do not focus on causation. Rather, they focused on how specific measures of mindfulness are related to one another. Therefore, this section will not use their article. Instead,

this section will provide a brief overview of Bradford Hill's Criteria for Causality, more commonly known as Hill's Criteria (Hill, 1965). See the appendix for more information. Hill's criteria include nine essential components: (a) strength, (b) consistency, (c) specificity, (d) temporality, (e) analogy, (f) plausibility, (g) coherence, (h) experimental evidence, and (i) biological gradient. Below we provide examples of strength, temporality, plausibility, and experimental evidence as they are the most common and critical factors in determining causation.

1. **Strength.** The strength of a relationship is, as it sounds, how much one variable influences the other. Specifically, strength is concerned with how much influence the independent variable (X) has on the dependent variable (Y). One example is the strength of the relationship between cognitive behavioral therapy (X) and depression (Y). Cognitive behavioral therapy, also known as CBT, is intended to change how people think and behave, especially as it relates to interpersonal difficulties (Martin, 2019). CBT has a strong relationship with depression; when an individual engages in cognitive behavioral treatment, the accompanying change in depression is also strong (Fava et al., 1998).
2. **Temporality.** Using the CBT example above, the CBT treatment should temporally precede improvement in depression. That is, symptoms of depression typically do not improve on their own, though fluctuations are not unheard of. Still, improvements in symptoms of depression should occur after treatment has begun.
3. **Plausibility.** CBT is associated with changes in mental health. Therefore, it is plausible that CBT will change symptoms of depression. On the other hand, trephination, or the placing of a hole through the skull (Petrone et al., 2015), has been used to treat depression for centuries, even though there is no evidence it directly effects symptoms of depression (Dobson, 2000). That is, there is no plausible relationship between trephination and depression.
4. **Experimental Evidence.** Experiments typically include two or more groups. This is done to make a comparison of the effect of a relationship. For example, there is extensive research showing that CBT has a stronger causal effect on improving symptoms of depression when compared with no intervention (placebo) (Hofmann et al., 2012). Simply put, CBT does a better job on improving depression than a placebo alone.

Simply on a definitional basis, correlation is less complex than causation. The same may be said when conducting research that is intended to show causality. Correlations are often reported for each study, when appropriate. On the other hand, causal relationships can only be shown after extensive research, often lasting years and spread across multiple studies. It is therefore critical that any claim of causation be closely examining the language used. For example, terms like "caused" or "predicted" imply a causal relationship. Alternatively, words like "associated" imply a noncausal relationship. The use of these terms do not necessarily reflect the research that has been done as authors often inappropriately invoke causation by using causal terms in studies that do not meet causal criteria. Determining that a causal relationship exists, when in fact it does not can result in significant harm. One contemporary example that made false claims of causation postulated a relationship between receiving the mumps, measles, and rubella (MMR) vaccine and the occurrence of autism (Wakefield et al., 1998)[1] an article that has since been retracted by the journal and admonished by other scholarly outlets (cf. Goodlee et al., 2011). The article violated a number of scientific principles undergirding causal explanations. For example, the

research did not include a control group. In addition, despite attempts at replicating Wakefield's study, scientists could not find a link between MMR and autism. Numerous controlled trials, a more rigorous approach, refuted Wakefield's conclusions. Since its initial publication in the prestigious journal *The Lancet*, there has been a rise in deaths from the measles as parents have refused to vaccinate their children for fear the child will develop autism (Patel et al., 2019). In spite of multiple studies debunking Wakefield's assertions, there continues to be a small but vocal group that support the unfounded claim.

Application to Social Work

With almost indeterminable access to the latest news and research, it is critical that social workers have the ability to determine the trustworthiness of research. Evidence-based social work practice is built upon foundations of research. Numerous social work interventions, from case management to cognitive behavioral therapy, are undergirded by evidence derived from research. That evidence is quantified using statistics. While we are not proposing that social workers also become trained statisticians, it is important that social work practitioners have the ability to interpret basic statistics.

Indeed, the promotion of and ability to critically examine evaluations and other research is noted in the *Code of Ethics of the National Association of Social Workers*. Social work professionals practicing in the field are uniquely poised to understand the real-world implications of research study findings. Understanding basic statistics lays a foundation for social work professionals to critically evaluate a study, from understanding how variables are being measured, to determining whether the researchers are making inflated claims about the relationships between two variables as well as if researchers are making claims about causation that should in fact be attributed to correlation. This critical knowledge can help to guide social work practitioners in discerning the degree to which they use a particular research study to inform their findings.

Appendix

Definitional Overview of Hill's Criteria

1. **Strength.** Weak associations are subject to bias, and unaccounted for variance. That is, when there are weak associations it is hard to know just *how* the variables are related. Strong associations, however, have the benefit of demonstrating a likely causal effect without introducing unknown bias.
2. **Consistency.** Repeated observation of the same relationship, but in differing context (e.g., different population, different setting) should yield similar results. While inconsistent findings do not negate causality as a potential explanation for the examined variables, it is more difficult to discern a causal relationship. If context is the primary factor influencing the relationship, a causal relationship cannot be accurately identified.
3. **Specificity.** A specific, observable outcome or a limited number of outcomes should be observed. That is, the relationship between the variables should produce a quantifiable effect. In cases where multiple effects are identified, a strong theoretical and statistical argument must be made for a specific, rather than a general relationship.
4. **Temporality.** The *cause* must precede the *effect* in time. In cases where temporality cannot be accurately determined, or where two variable change together regardless of time, correlation is a more accurate means to describe the relationship. While circular relationships are possible, the effect must still follow the initial cause. The continuing recurrence of the relationship is inconsequential.
5. **Analogy.** Analogous research to support causation is an important consideration. Similar research conducted

using experimental approaches should result in similar findings for, or against causation.

6. **Plausibility.** A demonstrated relationship must be likely before it can be causally defined. Statistical error, methodological fallacies, and undefined bias can all result in positive findings, when actually a causal relationship does not exist.
7. **Coherence.** A coherent relationship is one that aligns with existing history, research, and theory. That is, the evidence is not likely to be a radical departure from what is already known. However, paradigm shifts (cf. Kuhn, 1962) do occur, though they are quite rare. One such shift is the current focus on neurobiology to explain mental illness. While this type of research is still new, the emphasis is not on symptoms or contextual relationships, rather a focus on brain function and biological predisposition for mental illness.
8. **Experimental Evidence.** The use of experimental research design is an essential, and arguably necessary component to determine whether a relationship is causal. Experimental evidence is derived using two groups at a minimum—an *experimental* group and a *control* group. Experimental methods provide a mechanism to compare two similar groups, with only one group receiving the new intervention; the control group typically receives a placebo.

In addition, it is recommended that you see "Research at Work: Understanding Regression Tables" (Kondrat & Jaggers, 2018) for a brief overview of explained variance.

Please refer to Kendler et al. (1999) work, "Causal Relationship Between Stressful Life Events and the Onset of Major Depression" for an example of a causal study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jeremiah W. Jaggers https://orcid.org/0000-0002-5066-0192

Note

1. The authors cite this article as an example of the negative consequences associated with prematurely asserting causality. The Wakefield article lacks significant scientific credibility and should not be considered when conducting research on autism, vaccines, or good science.

References

Abu-Bader, S. H. (2011). *Using statistical methods in social science research*. Lyceum Books.

Allen, M. (2017). *The SAGE encyclopedia of communications research methods*. SAGE.

Dobson, R. (2000). Doctors warn of the dangers of trepanning. *British Medical Journal*, *320*(7235), Article 602.

Fava, G. A., Rafanelli, C., Grandi, S., Conti, S., & Belluardo, P. (1998). Prevention of recurrent depression with cognitive behavioral therapy: Preliminary findings. *Archives of General Psychiatry*, *55*(9), 816–820.

Goodlee, F., Smith, J., & Marcovitch, H. (2011). Wakefield's article linking MMR vaccine and autism was fraudulent. *The British Medical Journal*, *342*, 1–4.

Hill, A. B. (1965). The environment and disease: Association of causation? *Proceedings of the Royal Society of Medicine*, *58*, 295–300.

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy Research*, *36*(5), 427–440.

Kendler, K. S., Markowski, L. M., & Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. *American Journal of Psychiatry*, *156*(6), 837–841.

Kondrat, D. C., & Jaggers, J. W. (2018). Research at work: Understanding regression tables. *Families in Society*, *97*(2), 142–148.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago University Press.

Martin, B. (2019). *In-depth: Cognitive behavioral therapy*. https://psychcentral.com/lib/in-depth-cognitive-behavioral-therapy/

Patel, M., Lee, A. D., Redd, S. B., et al,. (2019). Increase in Measles Cases—United States, January 1–April 26, 2019. *Morbidly and Mortality Weekly Report*, *68*, 402–404.

Petrone, P., Niola, M., Di Lorenzo, P., Paternoster, M., Graziano, V., Quaremba, G., & Buccelli, C. (2015). Early medical skull surgery for treatment of post-traumatic osteomyelitis 5,000 years ago. *PLoS ONE*, *10*(5), Article e0124790.

Sedgwick, P. (2012). Pearson's correlation coefficient. *British Medical Journal*, *345*, 1–2.

Temme, L. J., & Wang, D. (2018). Relationship between the five facets of mindfulness on mood and substance use relapse. *Families in Society*, *99*(3), 209–218.

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., . . . Smith, J. A. (1998). RETRACTED: Ileal-lymphoid nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, *351*(9103), 637–641.