Communications in Applied Geometry.
ISSN 2249-4286 Volume 13, Number 2 (2023) pp. 185-186

www.ripublication.com/cag.htm

A Statistical Margin of Error from a Geometric Perspective

David A. Swanson
University of California Riverside, Riverside CA 92521
Center for Studies in Demography and Ecology
University of Washington
Seattle, Washington USA 98195I

Abstract

A statistical margin of error for the mean of variable x taken from a random sample of size n from a population of interest of very large, possibly infinite, size can be presented as an extension of the Pythagorean distance between two vectors, where one vector, V_m, represents the mean of the random sample, and the other, $V_{1,}$ is found by summing the n values of x and letting these n summed values form the elements of $V_{1,}$

Suppose we have a random variable, x, where the values of x taken from a random sample of size n from a population of size N, where N is possibly infinite, but at the least very large and not necessarily known. Let the sum of the n values of random variable x equal Σn_i. These same n values form the n elements of vector V_1, There are exactly n - 1 additional distinguishable vectors, V_2, V_3, ... , V_n, that can be formed by permuting the n elements of V_1 such that $V_1 + V_2 + V_3 + ,..., + V_n = (\Sigma ni_1, , \Sigma ni_2, \Sigma ni_3, , \Sigma ni_n) = V_s$. Multiplying V_s by the scalar (l/n) gives V_m, the point in n – space that is $(l/n)^{th}$ the distance from the origin to Vs. V_m can be interpreted as the mean of the random sample of size n.

Now consider an infinite number of random samples of size n from a population of size N, where N is possibly infinite, but at the least very large and not necessarily known, for which this same process was repeated. Let the mean of all of these sample means be V_n, which is equal to the mean of the entire population. Because we have an infinite number of samples, V_n, is unknown to us but of interest. We can estimate the distance from V_m to V_n with a selected degree of probability and an extension of this distance can be viewed as the Margin of Error, a statistic expressing the level of confidence in the estimate of the distance between V_m to $V_{n.}$.

How can we do this? First, Let *D* equal the Pythagorean distance between V_1 and V_m.

We then multiply *D* by the scalar $(1/n)^{½}$, which allows us to present the standard deviation of our random sample as $s = (1/n)^{½}(D)$. It follows that the standard error associated with our sample mean, V_m, is $se = D$. With a given level of certainty, knowledge of *D* allows us to assess how far V_m is from V_n. We can do by constructing a "Margin of Error," (MOE). If $n \leq 120$, then MOE = $t_c D$, where t_c can be taken from a "student's t table" as found in Spiegel [1, p.344]. If, for example, n = 25 and we want to be 95% certain of the distance between V_m and V_n, then we select $t_c = 2.06$.

If n >120, then MOE = $z_c D$, where z_c is taken from a "z table" as found in Spiegel [1, p. 343] Thus, if n > 120 and we want to be 95% certain of the distance between V_m and V_n, then we select $z_c = 1.96$.

Thus, from the perspective of geometry (and linear algebra), an MOE is an extension of D, where the extension is determined by n and the level of confidence we want in terms of the distance between V_m and V_n. Note that if D = zero, random variable x is a constant value, k, which implies $V_m = V_n$.

Reference

[1] Spiegel, M. (1961). *Theory and Problems of Statistics*. Schaum's Outline Series. McGraw-Hill. New York.