

How Large Should a Statistical Sample Be?

Violeta C. Menil and Ruili Ye, Hostos CC (CUNY)



Violeta Menil is an associate professor at the Mathematics Department of Hostos CC of the City University of New York. Her research career started with a computer simulation (Monte Carlo study) evaluating two multidimensional scaling algorithms for fitting the weighted Euclidean distance model (Ph.D. dissertation). She has authored 12 publications and was awarded 2 research grants on developmental mathematics. Her interests focus on multidimensional scaling, univariate and multivariate statistics.



Ruili Ye is an assistant professor at Hostos CC. Her Ph.D. thesis was a first-order logic formalization of the sense and reference notions introduced by the German philosopher Gottlob Frege. Her interests include topics in mathematical and philosophical logic as well as a keen interest in Bertrand Russell.

This study serves as a teaching aid for teachers of introductory statistics. The aim of this study was limited to determining various sample sizes (n) when estimating population proportion, p . Tables on sample sizes were generated using a C++ program, which depends on population size (N), degree of precision or error level (E), and confidence level (Z). Nineteen different population sizes, five degrees of precision, and three levels of confidence were utilized. The study found out that the larger the population size, (N), the higher the degree of precision, (E) and the higher the probability/confidence level, (Z), the larger the sample size must be.

Two values for the sample estimate () of the population proportion (p) were used in this study. Practical applications of randomly pulling appropriate number of samples from huge data sets were also discussed.

