

Tools For Presenting Examples of Sampling Distributions and Hypothesis Testing

Mark C. Fulcomer¹, S. David Kriska², Marcia M. Sass³

¹ Rutgers School of Public Health, 48 Trainor Circle, Bordentown, NJ 08505

² Walden University, 3758 Surrey Hill Place, Upper Arlington, OH 43220

³ Rutgers School of Public Health, 48 Trainor Circle, Bordentown, NJ 08505

Abstract

This poster describes three Excel spreadsheets for presenting real-time illustrations of some basic concepts of sampling distributions and hypothesis testing in the classroom. Incorporating “recursive” formulae to simplify calculations (Fulcomer et al., 2007), the first and second spreadsheets (for the binomial and Poisson distributions, respectively) can aid in explaining important properties of sampling distributions for discrete random variables as well as in comparing two competing sets of values for the distributions’ parameters. The third spreadsheet utilizes the normal distribution to display power calculations for simple hypothesis testing for a wide variety of instructor-specified scenarios. These can include comparisons of distributions with 641 different means by five levels of sample sizes and four levels of Type I errors (significance), each with both one- vs. two-tailed rejection regions. All three of the spreadsheets have built-in graphs to accompany the tabled calculations. Because of the ease and speed with which examples can be modified, the spreadsheets can also be used to introduce topics related to sample size determination.

Key Words: sampling distributions; hypothesis testing; power calculations; significance levels; one- vs. two-tailed tests; sample size determination

1. Introduction

Our interest in facilitating improved classroom presentations of methods and examples for introductory and intermediate courses goes back over a decade (Fulcomer et al., 2002) and includes treatments of the use of historical health and bibliographic information (Fulcomer et al, 2004) and student projects (Sass et al., 2005), in health, health-related, and other diverse disciplines (Fulcomer et al., 2005). Spreadsheets for two major random variables (the binomial and Poisson) were introduced next (Fulcomer et al., 2007), followed by more general illustrations of the construction of sampling distributions for discrete random variables (Fulcomer et al., 2008).

2. Descriptions of the Three Spreadsheets

The first and second spreadsheets can be helpful in conveying some major properties of “point estimators” (e.g., unbiasedness, consistency, relative efficiency, and sufficiency) as well as introducing the topics of hypothesis testing and sample size determination to satisfy desired levels for Type I and Type II errors. Because the spreadsheets effectively eliminate the need for students to refer to traditional textbook tables for these two discrete random variables, they enable presentations and exercises to address somewhat more “realistic” and “interesting” examples.

The third spreadsheet introduced here covers the classic, simplest case of a continuous normally distributed random variable with “**known**” parameters with the user specifying values for the population mean μ_x and standard deviation σ_x , respectively. Beyond the case of $N = 1$, users also provide **four additional sample size (N) levels** and **four levels of Type I (Alpha or α) errors**. For each combination of sample size (N) and Type I error α , the **standard error** is calculated as σ_x/\sqrt{n} . Also found for the 20 combinations of the sample size and Type I error levels are the **critical values (CVs) for 2-tailed and 1-tailed rejection regions** (i.e., where cumulative probabilities are $\alpha/2$ and $1-\alpha/2$ and $1-\alpha$, respectively). Finally and in each of the combinations of the sample size and Type I error levels, the spreadsheet finds the necessary cumulative probabilities to calculate the **Type II (Beta or β) error** to determine **$1-\beta$, the statistical power** of 2-tailed and 1-tailed test over 641 values for the mean in 2-digit standardized values in the range from -3.20 to 3.20, corresponding to the tables that appear in many standard textbooks. Like its two discrete counterparts, the cumulative values when $N = 1$ can be used to compare results for both the original and the standardized values to those found in many textbooks, once again helping students to better understand the critically important normal distribution.

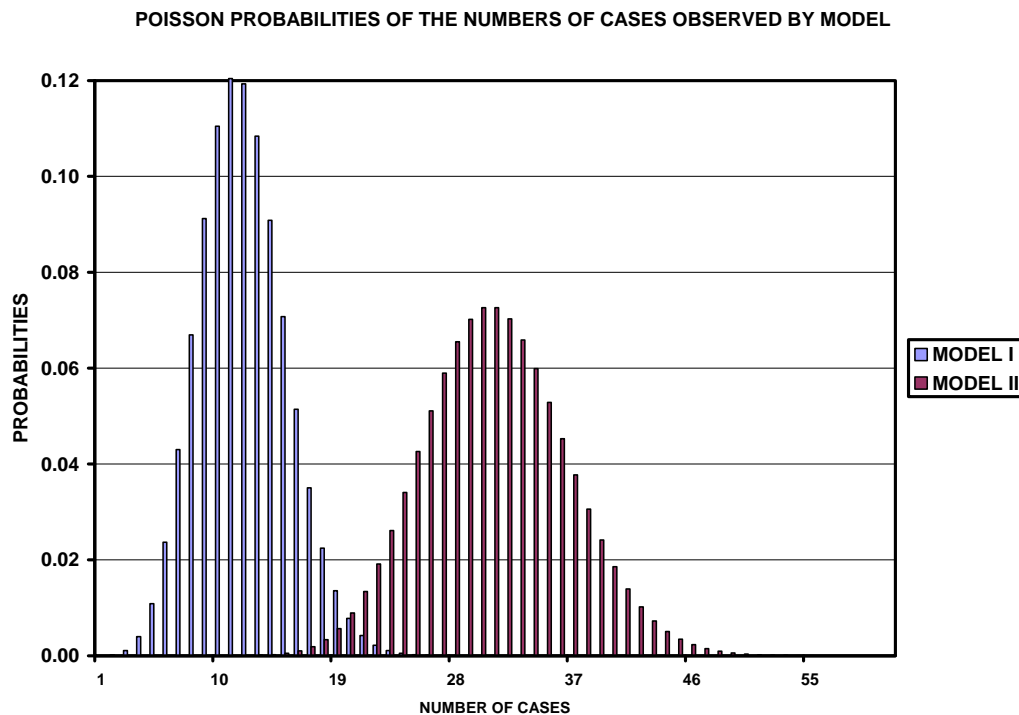
2. Using the Spreadsheets

There are color-coded instructions found at the end or bottom of each of the three spreadsheets. Values for specifying any input values needed are quite easy to enter and are indicated by blue shading, while calculated values use yellow shading; some shadings of red and green are also used to provide visual separations and emphases.

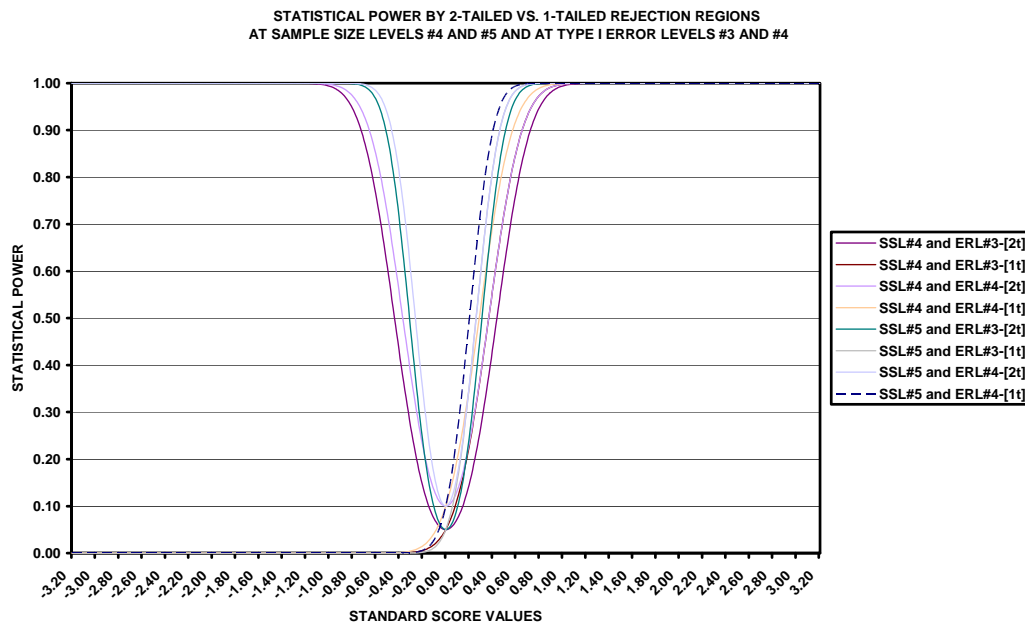
Once needed values have been entered, the spreadsheets return tables of results very quickly, even in the case of the normal distribution which produces 181 pages of output. Each of the spreadsheets provide results under a null hypothesized value H_0 and an alternative H_A . These tables of values to be used by students to compare results to textbook examples as well as for other assignments.

The three spreadsheets also automatically provide color graphs of the results. For the binomial and Poisson distributions there are graphs for both the number of events as well as their proportion, which except for the scales of measurement are identical and provide students with insights into “invariance” properties of some statistics. Displayed in the figure below are the number of cases (up to 60 are displayed here) when $N = 2000$ and probabilities of observing a case are .00545 and .015 under H_0 and H_A , respectively (i.e., the corresponding “intensity” parameter values are 10.90 and 30.00, respectively).

Figure 1



For the normal distribution spreadsheet, there are many more graphs: (1) statistical power by Type I level and 2-tailed vs. 1-tailed test for each of the 5 sample size levels; (2) statistical power with 2-tailed rejections by sample size levels for each of the 4 Type I error levels; and (3) statistical power by 2-tailed and 1-tailed rejection regions at the 4th and 5th sample size levels and at the 3rd and 4th Type I error levels. With the example shown in the figure that follows, the third type of graph can be especially helpful when the 4th and 5th sample size levels are set for modestly large and large sample sizes with slightly higher Type I errors for the 3rd and 4th levels, say at .05 and .10 or at .10 and .20, for illustrations of “realistic” power calculations for clinical trials.

Figure 2

Clearly, these spreadsheets provide considerable flexibility in presenting some major concepts related to sampling distributions and hypothesis testing, including: (1) the construction of sampling distributions; (2) independence and independent random sampling; (3) definitions and computations of parameters and statistics; (4) means, variances, standard deviations, and standard errors; (5) expected values; (6) convergence to continuity; (7) unbiasedness; (8) consistency; (9) normal approximations and the central limit theorem; (10) hypothesis testing and sample size determination for null and alternative hypotheses; (11) critical values for 1-tailed and 2-tailed rejection regions; (12) confidence intervals; (13) Type I errors and false positives; (14) Type II errors and false negatives; (15) statistical power and sensitivity; and (16) increasing statistical power by a) by 1-tailed vs. 2-tailed rejection regions, b) by increasing in Type I error levels, c) by increasing sample size (N), and d) by increasing effect sizes.

Perhaps more importantly, these spreadsheets provide students with tools that can be effectively utilized in reading and mastering some of the seminal works of statistical power and effect sizes, including Hays (1963), Cohen (1969), and Cohen and Cohen (1983).

References

- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fulcomer, M. C., Headley, A. J., Sass, M. M., Kriska, S. D., Chung, K., Clayton, J. L., LaRue, M., and Holton, M. W. (2002). *Incorporating large and historical bibliographies into the teaching of applied statistics and epidemiology*. Proceedings of the American Statistical Association, 1187-1189.

- Fulcomer, M. C., Sass, M. M., Holton, M. W., Clayton, J. L., Kriska, S. D., and Hirschman, C. L. (2004). *Some classroom roles for historical health information*. Proceedings of the American Statistical Association, 4722-4724.
- Fulcomer, M. C., Lyke, J. A., Colón, M., Sass, M. M., and Kriska, S., D. (2005) *Trans-disciplinary approaches to the teaching of “applied” statistics: ‘Similar Strokes for Different Folks’*. Proceedings of the American Statistical Association, 3741-3743.
- Fulcomer, M. C., Kriska, S. D., Jauregui, M., Sass, M. M., and Kennedy, J. (2007). *Applying some computational simplifications for binomial and Poisson random variables to the teaching of statistical concepts*. Proceedings of the American Statistical Association, 3545-3552.
- Fulcomer, M. C., Kriska, S. D., Sass, M. M., and Jauregui, M. (2008). *Classroom illustrations of the construction of discrete sampling distributions*. Proceedings of the American Statistical Association, 3229-3234.
- Hays, W. L. (1963). *Statistics for Psychologists*. New York: Holt, Rinehart and Winston.
- Sass, M. M., Fulcomer, M. C., Kriska, S. D., Jones, A. E., Little, B. T., Holton, M. W., Polit, C., Knoblauch, K., DiBruno, D., and Kennedy, J. M. (2005). *Examples of student projects using historical health statistics*. Proceedings of the American Statistical Association, 3754-3755.