Confidence Intervals for Percentages and Counts

Introduction

This document describes the algorithms for computing confidence intervals for percentages and counts for bar charts. The data are assumed to be from a simple random sample, and each confidence interval is a separate or individual interval, based on a binomial proportion of the total count. The computed binomial intervals are equal-tailed Jeffreys prior intervals (see Brown, Cai, & DasGupta, 2001, 2002, 2003). Note that they are generally not symmetric around the observed proportion. Therefore, the plotted interval bounds are generally not symmetric around the observed percentage or count.

Notations

The following notation is used throughout this chapter unless otherwise noted:

X_i	Distinct values of the category axis variable
w_i	Rounded sum of weights for cases with value X_i
$W = \sum_i w_i$	Total sum of weights over values of X
p_i	Population proportion of cases at X_i
α	Specified error level for $100(1-\alpha)$% confidence intervals

IDF.BETA(p,shape1,shape2) in COMPUTE gives the p^{th} quantile of the beta distribution or incomplete beta function with shape parameters shape1 and shape2. For a precise mathematical definition, see page 2 of "Appendix 12: Cumulative Distribution, Percentile Functions, and Random Numbers."

Confidence Intervals for Counts (Wp_i)

Lower bound for $W p_i = W$ [IDF.BETA(α/2,w_i +.5,W-w_i +.5)].

Upper bound for $W p_i = W$ [IDF.BETA(1-α/2,w_i +.5,W-w_i +.5)].

Confidence Intervals for Percentages ($100p_i$)

Lower bound for $100\ p_i = 100$ [IDF.BETA(α/2,w_i +.5,W-w_i +.5)].

Upper bound for $100\ p_i = 100$ [IDF.BETA(1-α/2,w_i +.5,W-w_i +.5)].

References

Brown, L. D., Cai, T., & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, **16**(2): 101-133.

Brown, L. D., Cai, T., & DasGupta, A. (2002). Confdence intervals for a binomial Proportion and asymptotic expansions. The Annals of *Statistics*, **30**(4): 160-201.

Brown, L. D., Cai, T., & DasGupta, A. (2003). Interval estimation in exponential families. *Statistica Sinica*, 13: 19-49.