

# Asymptotic normality of goodness-of-fit statistics for sparse Poisson data

Ursula U. Müller and Gerhard Osius

Department of Mathematics and Computer Sciences

University of Bremen

P.O. Box 33 04 40

D-28334 Bremen, Germany

November 9, 2000

## **Abstract**

Goodness-of-fit tests for discrete data and models with parameters to be estimated are usually based on Pearson's  $\chi^2$  or the Likelihood Ratio Statistic. Both are included in the family of Power-Divergence Statistics  $SD_\lambda$  which are asymptotically  $\chi^2$  distributed for the usual sampling schemes. We derive a limiting standard normal distribution for a standardization  $T_\lambda$  of  $SD_\lambda$  under Poisson sampling by considering an approach with an increasing number of cells. In contrast to the  $\chi^2$  asymptotics we do not require an increase of all expected values and thus meet the situation when data are sparse. Our limit result is useful even if a bootstrap test is used, because it implies that the statistic  $T_\lambda$  should be bootstrapped and not the sum  $SD_\lambda$ . The peculiarity of our approach is that the models under test only specify associations. Hence we have to deal with an infinite number of nuisance parameters. We illustrate our approach with an application.

Key words: contingency tables, goodness-of-fit, odds ratios, Poisson data, Power-Divergence Statistics, sparse data.