Pearson's Goodness of Fit Statistic as a Score Test Statistic

Gordon K. Smyth

Abstract

For any generalized linear model, the Pearson goodness of fit statistic is the score test statistic for testing the current model against the saturated model. The relationship between the Pearson statistic and the residual deviance is therefore the relationship between the score test and the likelihood ratio test statistic, and this clarifies the role of the Pearson statistic in generalized linear models. The result is extended to cases in which there are multiple reponse observations for the same combination of explanatory variables.

Keywords: Pearson statistic; score test; chi-square statistic; generalized linear model; exponential family nonlinear model; saturated model

1 Introduction

Goodness of fit tests go back at least to Pearson's (1900) article establishing the asymptotic chi-square distribution for a goodness of fit statistic for the multinomial distribution. Pearson's chi-square statistic includes the test for independence in two-way contingency tables. It has been extended in generalized linear model theory to a test for the adequacy of the current fitted model. Given a generalized linear model with responses y_i, weights w_i, fitted means $\hat{\mu}_i$, variance function $v(\mu)$ and dispersion $\phi = 1$, the Pearson goodness of fit statistic is

$$X^2 = \sum \frac{w_i(y_i - \hat{\mu}_i)^2}{v(\hat{\mu}_i)}$$

[14]. If the fitted model is correct and the observations y_i are approximately normal, then X^2 is approximately distributed as χ^2 on the residual degrees of freedom for the model.

The Pearson goodness of fit statistic X^2 is one of two goodness of fit tests in routine use in generalized linear models, the other being the residual deviance. The residual deviance is the log-likelihood ratio statistic for testing the fitted model against the saturated model in which there is a regression coefficient for every observation. The Pearson statistic is a quadratic form alternative to the residual deviance, and is often preferred over the residual deviance because of its moment estimator character. The