DISCUSSION

Norman Breslow and Dirk Moore

University of Washington

In our first reading of this stimulating paper, we were struck by the apparent close correspondence between the proposed random effects model for contingency table analysis and the theory of generalized linear models and quasi-likelihood developed by Nelder and Wedderburn (1972) and Wedderburn (1974) and investigated further in the recent monograph by McCullagh and Nelder (1983). For some simple parametric models, the one-parameter family of densities (5.6) appears identical to the corresponding quasi-likelihood functions with $\nu^{-1} = (n\theta)^{-1}$ playing the role of the dispersion parameter that multiplies the variance function. Further, parallels between the two developments are seen in the fact that residual deviance or chi-square goodness-of-fit measures, divided by their degrees of freedom, are used to estimate ν (cf. equations (4.8) and (5.20)). Furthermore, $\hat{\nu}$ acts as a scaling factor for the asymptotic distribution of the vector of sufficient statistics for the β parameters of primary interest (5.23), as it does for the parameter estimates under quasi-likelihood. It is quite interesting that the relatively simple asymptotic results, already known to hold unconditionally from quasi-likelihood theory, apply even to the complicated conditional distributions considered in the paper. Since we suspect that such matters will receive more thorough discussion from others with greater knowledge of quasi-likelihood techniques, however, we turn our attention to a possible alternative measure of the degree to which a given table conforms to the hypothesis of independence.

As is well known (Bishop, Fienberg and Holland, 1975), the hypothesis of independence of row and column classifications in a contingency table may be expressed as a log-linear model for the expected cell frequencies $E(m_{ij})$. The likelihood calculations are simplest when the m_{ij} have independent Poisson distributions, and we keep to this in order to make the discussion as transparent as possible. Conditioning on the grand total leads to the hypothesis of independence for the multinomial distribution considered by Diaconis and Efron.

In order to accommodate the idea of a sample size that increases while the number of cells IJ remain fixed, we suppose that the Poisson means have the form $E(m_{ij}) = N\lambda_{ij}$. We further suppose that the λ_{ij} are sampled independently from distributions with means $\zeta_{ij} = \exp(\mu_0 + \alpha_i + \beta_j)$, the hypothesis of independence, and variances that represent the degree of departure from this hypothesis. The usual quasi-likelihood generalization of the Poisson model results from the assumption $\mathrm{Var}(\lambda_{ij}) = \sigma^2\zeta_{ij}$. It follows that $E(m_{ij}) = N\zeta_{ij} = \mu_{ij}$ and $\mathrm{Var}(m_{ij}) = (1 + N\sigma^2)\mu_{ij}$. This generalized linear model with dispersion parameter $\theta^{-1} = (1 + N\sigma^2)$ corresponds closely to the situation considered by Diaconis and Efron. Imposition of the parameter constraint $\sum_{i,j} \zeta_{ij} = 1$ via an appropriate choice of μ_0 ensures that N is estimated by the grand total n.

An alternative random effects model that seems to us more natural in the context of log-linear theory expresses the random effects on the same scale as