Rostra
a cura di
By
Benito V. Frosini

A UNIFIED FRAMEWORK FOR CONDITIONAL BINOMIAL MODELS: THE BETA–BINOMIAL, ALTHAM'S AND THE MARKOV–LIKE SUSCEPTIBILITY MODELS

Maria Lea Correa Leite

Epidemiologic Unit, C.N.R Milano.

Mario Braga

Epidemiologic Unit, Ospedale Maggiore Milano.

Ettore Marubini

Istituto di Statistica Medica e Biometria, Università degli Studi di Milano.

The paper considers three probability models: the beta–binomial, Altham's and the Markov–like (zero–inflated Poisson) processes, which have found wide application in modelling correlated binary outcomes. All models are conditional binomial models differing only for the flexibility in describing the prior distribution of the binomial probabilities. The paper presents the three models and shows how to derive them analytically. For the special case where the number of subunits per unit is fixed and equal to two, the equality of the three models is shown.

Key word: Correlated binomial outcomes, Beta–Binomial distribution, Altham's distribution, Markov–like susceptibility models

1. INTRODUCTION

The biomedical literature presents many examples of researches, such as family studies, teratology and measurements on eyes or teeth, where the study design has a hierarchical structures. The responses on the sub–units (members of the family, eyes or teeth of an individual) are said to be nested within the units (families, subjects). Under these circumstances, it is plausible to assume a relative homogeneity of the sub–units within each unit or in other terms we may say that the responses tend to be correlated. In order to deal with dependent dichotomous outcomes, an appealing approach may be obtained by resorting to the intraclass correlation coefficient, which is a measure of the dependence between subunits within each unit, and by inserting it as one of the parameters of the unconditional distribution function. In the biomedical field as well as in that of industrial production, several models have been suggested: the beta–binomial model, which has been extensively used in toxicology (Williams, 1975; 1982; 1988), Altham's distribution (Altham, 1976), and the zero–inflated Poisson or Markov–like processes (Campbell et al., 1991; Lambert, 1992; Zucker et al., 1992), which assume the population under study to be composed of subjects with no events and subjects with one or more occurrences. This is like assuming the existence of two processes: one modelling the probability of becoming susceptible, the other defining the probability of developing the event in a susceptible individual. All models considered in the paper are conditional binomial models. They differ in that for the beta–binomial the binomial parameter is modelled via a continuous distribution (beta density function), whereas for the Altham and the Markov–like susceptibility models the prior distribution of the binomial parameter is discrete: the former model has a 3–point and the latter a 2–point distribution.

2. METHODS

Let y_{ijk} denote the dichotomous outcome from the k–th subunit (observational) of the j–th unit (experimental) in the i–th group, $i=1,2,...,g, j = 1,2,...,m_i$ and $k=1,2,...,n_{ij}$.

$$y_{ijk} = \begin{cases} 1 & \text{positive outcome} \\ 0 & \text{otherwise} \end{cases}$$

The random variable y_{ij+} (where + indicates the sum over k) is assumed to be binomially distributed with parameters z_{ij} and n_{ij}, that is:

$$f(y_{ij+} = r \mid z_{ij}, n_{ij}) = \binom{n_{ij}}{r} z_{ij}^r (1 - z_{ij})^{(n_{ij} - r)}$$

where $r=0,1,...,n_{ij}$.

In order to allow for the heterogeneity of different units, we can assume that the binomial parameter, z_{ij}, is, in its turn, a random variable whose distribution can be modelled by:

case I: A probability density function beta with parameters α_i and β_i:

$$f(z_{ij} \mid \alpha_i, \beta_i) = \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i) \cdot \Gamma(\beta_i)} z_{ij}^{\alpha_i - 1} (1 - z_{ij})^{\beta_i - 1} \qquad 0 \le z_{ij} \le 1 \qquad \alpha_i > 0, \quad \beta_i > 0$$

The beta distribution is extremely flexible and can assume different shapes according to the parameters α_i and β_i. For a detailed discussion of the property of this distribution the reader is referred to Johnson and Kotz (1970).

case II: A discrete probability model, where the r.v. z_{ij} follows a 3–point distribution:

$$f(z_{ij} = 0) = \rho_i (1 - p_i)$$

$$f(z_{ij} = p_i) = (1 - \rho_i)$$

$$f(z_{ij} = 1) = \rho_i p_i$$

where $p_i = pr\,(y_{ijk} = 1)$ is the probability of a positive outcome in the i–th group, and ρ_i is the intraunit correlation coefficient. The latter accounts for the possible non–independence among the outcomes of the subunits belonging to the same unit.

case III: A discrete probability model where the r.v. z_{ij} follows a 2–point distribution:

$$pr(z_{ij} = 0) = 1 - \frac{p_i}{\tilde{p}_i}$$

$$pr(z_{ij} = \tilde{p}_i) = \frac{p_i}{\tilde{p}_i}$$

where $\tilde{p}_i = pr(y_{ijk} = 1 \mid y_{ijh} = 1) > p_i$ for $k \ne h$

The joint probability function of y_{ij+} and z_{ij} is given by the product of the conditional distribution of y_{ij+}, given z_{ij} and the prior ditribution of z_{ij}, that is:

$$f(y_{ij+} = r \mid z_{ij}) \cdot f(z_{ij})$$

For case I the joint probability is:

$$\binom{n_{ij}}{r} \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i) \cdot \Gamma(\beta_i)} z_{ij}^{(\alpha_i + r - 1)} (1 - z_{ij})^{(\beta_i + n_{ij} - r - 1)}$$

For cases II and III, the joint distribution can be easily obtained by resorting to a contingency table:

case II:

$f(y_{ij+} \mid z_{ij})$

y_{ij+} / z_{ij}	0	1	·	r	·	n_{ij}	$f(z_{ij})$
0	1	0	·	0	0	0	$\rho_i(1-p_i)$
p_i	$(1-p)^{n_j}$	$n_{ij}p_o(1-p_i)^{h_{ij}}$	·	$\binom{n_{ij}}{r} p_i^r(1-p_i)^{(n_{ij}-r)}$	·	$p_i^{n_{ij}}$	$(1-\rho_i)$
1	0	0	·	0	·	1	$\rho_i p_i$

case III:

$f(y_{ij+} \mid z_{ij})$

y_{ij+} / z_{ij}	0	1	·	r	·	n_{ij}	$f(z_{ij})$
0	1	0	·	0	0	0	$1-\frac{p_i}{\tilde{p}_i}$
$\tilde{p}_i$	$(1-\tilde{p}_i)^{n_{ij}}$	$n_{ij}\tilde{p}_i(1-\tilde{p}_i)^{n_{ij}-1}$	·	$\binom{n_{ij}}{r} \tilde{p}_i^r(1-\tilde{p}_i)^{(n_{ij}-r)}$	·	$\tilde{p}_i^{n_{ij}}$	$\frac{p_i}{\tilde{p}_i}$

The unconditional (marginal) distribution of y_{ij+} is obtained integrating out z_{ij}, that is:

$$
\begin{aligned}
f(y_{ij+}=r) &= \int_0^1 f(y_{ij+}=r \mid z_{ij})\, df(z_{ij}) \\
&= \binom{n_{ij}}{r} \frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\cdot\Gamma(\beta_i)} \int_0^1 z_{ij}^{(\alpha_i+r-1)}(1-z_{ij})^{(\beta_i+n_{ij}-r-1)} dz_{ij} \\
&= \binom{n_{ij}}{r} \frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\cdot\Gamma(\beta_i)} \cdot \frac{\Gamma(\alpha_i+r)\Gamma(\beta_i+n_{ij}-r)}{\Gamma(\alpha_i+\beta_i+n_{ij})}
\end{aligned}
$$

The resulting distribution is the beta–binomial one.

A better insight of the beta–binomial distribution is attained by reparametrizing to:

$$p_i = \frac{\alpha_i}{(\alpha_i + \beta_i)} \quad \text{and} \quad \phi_i = (\alpha_i + \beta_i)^{-1}$$

The first parameter is the expectation of the i–th beta distribution and, given p_i, the second parameter, ϕ_i, determines the distribution shape.

Furthermore, Crowder (1979) pointed out that:

$$\phi_i(\phi_i + 1)^{-1} = (\alpha_i + \beta_i + 1)^{-1} = \rho_i$$

In the case of discrete distributions, the sum replaces the integral operator. Consequently, we have respectively for case II and case III the following results:

$$f(y_{ij+} = r) = \sum_{z_{ij}=0}^{1} f(y_{ij+} \mid z_{ij}) \cdot f(z_{ij})$$

case II

$$f(y_{ij+} = 0) = \rho_i(1 - p_i) + (1 - \rho_i)(1 - p_i)^{n_{ij}}$$

$$pr(y_{ij+} = r) = \binom{n_{ij}}{r}(1 - \rho_i)p_i^r(1 - p_i)^{n_{ij} - r} \qquad (1 \le r \le n_j - 1)$$

$$pr(y_{ij+} = n_{ij}) = \rho_i p_i + (1 - \rho_i)p_i^{n_{ij}}$$

which is the Altham distribution (1976), and

case III

$$f(y_{ij+} = 0) = 1 - \frac{p_i}{\tilde{p}_i} + (1 - \tilde{p}_i)^{n_{ij}} \cdot \frac{p_i}{\tilde{p}_i}$$

$$f(y_{ij+} = r) = \binom{n_{ij}}{r}\tilde{p}_i^r(1 - \tilde{p}_i)^{(n_{ij} - r)} \cdot \frac{p_i}{\tilde{p}_i}$$

$$= \binom{n_{ij}}{r} p_i \tilde{p}_i^{(r-1)}(1 - \tilde{p}_i)^{(n_{ij} - r)} \qquad (1 \le r \le n_j)$$

which is the Markov–like susceptibility model discussed by Zucker and Wittes (1991). By an easy reparametrization of the prior distribution of the 2–point model, that is:

$$f\left(z_{ij} = 0\right) = \tau_i$$

one obtains

$$f\left(y_{ij+} = 0\right) = \tau_i + \left(1-\tau_i\right)\left(1-\tilde{p}_i\right)^{n_{ij}}$$

$$f\left(y_{ij+} = r\right) = \left(1-\tau_i\right)\binom{n_{ij}}{r}\tilde{p}_i^{\,r}\left(1-\tilde{p}\right)^{\left(n_{ij}-r\right)}$$

which, for analogy with Lambert's definition (1992), will be called the zero–inflated binomial model. This author, in fact, models $f(y_{ij+} = r \mid z_{ij})$ via a Poisson model and calls the resulting unconditional distribution the zero–inflated Poisson model. The correlation between subunits within each unit, can be expressed by:

$$\rho_i = \frac{\tilde{p} - p_i}{\left(1-p_i\right)}$$

These three models have been widely applied in different areas of biological research where repeated observations were conducted on the same subject, leading to non–independent outcomes. If we consider, in particular, medical branches such as ophthalmology, audiology and nephrology, where the number of subunits within unit is fixed and equal to two, it is noteworthy that the three models are equivalent. For this purpose the following equalities are helpful:

		CASE I		CASE II–III
$p(y_{ijk} = 1)$	=	$\alpha_i/(\alpha_i + \beta_i)$	=	p_i
$p(y_{ijk} = 0)$	=	$\beta_i/(\alpha_i + \beta_i)$	=	$(1 - p_i)$
$p(y_{ijk} = 1/y_{ijh} = 1)$	=	$(\alpha_i + 1)/(\alpha_i - \beta_i + 1)$	=	$p_i(1 - \rho_i)+\rho_i = \tilde{p}_i$
$p(y_{ijk} = 0/y_{ijh} = 1)$	=	$\beta_i/(\alpha_i + \beta_i + 1)$	=	$(1 - p_i)(1 - \rho_i)$
$p(y_{ijk} = 1/y_{ijh} = 0)$	=	$\alpha_i/(\alpha_i + \beta_i + 1)$	=	$p_i(1 - \rho_i)$
$p(y_{ijk} = 0/y_{ijh} = 0)$	=	$(\beta_i + 1)/(\alpha_i - \beta_i + 1)$	=	$1 - p_i(1 - \rho_i)$
ρ_i	=	$1/(\alpha_i + \beta_i + 1)$	=	$(\tilde{p}_i - p_i)/(1 - p_i)$

It is easy to see that:

$$f(y_{ij+}=0) = \frac{\Gamma(\alpha_i+\beta_i)\Gamma(\alpha_i)\Gamma(\beta_i+2)}{\Gamma(\alpha_i)\Gamma(\alpha_i+\beta_i+2)}$$

$$= \frac{\beta_i(\beta_i+1)}{(\alpha_i+\beta_i)(\alpha_i+\beta_i+1)} =$$

$$=(1-p_i)\left[1-p_i(1-\rho_i)\right]=(1-p_i)\left[1-p_i+\rho_i\cdot p_i+\rho_i-\rho_i\right]$$

$$=(1-p_i)\rho_i+(1-p_i)^2(1-\rho_i) \text{ Altham's model}$$

$$=1-2p_i\frac{\tilde{p}_i}{\tilde{p}_i}+p_i\frac{\tilde{p}_i^2}{\tilde{p}_i}=1-\frac{p_i}{\tilde{p}_i}\left(1-1+2\tilde{p}_i-\tilde{p}_i^2\right)$$

$$=1-\frac{p_i}{\tilde{p}_i}\left[1-(1-\tilde{p}_i)^2\right] \text{ Markov-like model}$$

$$f(y_{ij+}=1) = \frac{2\Gamma(\alpha_i+\beta_i)\Gamma(\alpha_i+1)\Gamma(\beta_i+1)}{\Gamma(\alpha_i)\Gamma(\beta_i)\Gamma(\alpha_i+\beta_i+1)}$$

$$= \frac{2\alpha_i\beta_i}{(\alpha_i+\beta_i)(\alpha_i+\beta_i+1)} =$$

$$=p_i(1-p_i)(1-\rho_i)+(1-p_i)p_i(1-\rho_i)=2p_i(1-p_i)(1-\rho_i) \text{ Altham's model}$$

$$=2p_i(1-\tilde{p}_i) \text{ Markov-like model}$$

$$f(y_{ij+}=2) = \frac{2\Gamma(\alpha_i+\beta_i)\Gamma(\alpha_i+2)\Gamma(\beta_i+1)}{\Gamma(\alpha_i)\Gamma(\beta_i)\Gamma(\alpha_i+\beta_i+2)} =$$

$$= \frac{\alpha_i(\alpha_i+1)}{(\alpha_i+\beta_i)(\alpha_i+\beta_i+1)} =$$

$$=p_i\left[p_i(1-\rho_i)+\rho_i\right]=p_i^2(1-\rho_i)+p_i\rho_i \text{ Altham's model}$$

$$=p_i\tilde{p}_i \text{ Markov-like model}$$

3. DISCUSSION

The three models are based on three different correlation structures, which, in their turn, may represent three different mechanisms of dependency with decreasing levels of complexity among the subunits. In the beta–binomial model, the probability of an event in a given subunit appears to be conditioned by the various possible patterns of response observed in the remaining subunits. In contrast, in Altham's model this probability is conditioned by two patterns of response only: all or none of the remaining subunits showing an event. As regards the Markov–like susceptibility model, the probability of an event in a given subunit is conditioned merely by at least one event in any of the remaining subunits. In the presence of two subunits only, the correlation structure is unique and the three models coincide, as was shown algebraically in the second part of the previous section.

RIASSUNTO

Questo articolo considera tre modelli probabilistici: il modello Beta–Binomiale, il modello di Altham ed il modello "Markov–like". Questi modelli hanno trovato una diffusa applicazione nel caso di dati binari correlati. Tutti e tre i modelli appartengono alla classe dei modelli condizionati e si differenziano solamente per la flessibilità con cui vengono composte le probabilità binomiali. Nel presente articolo, i tre modelli vengono descritti e viene derivata la loro forma analitica. Nel caso particolare che il numero di sotto–unità per ciascuna unità sia uguale a due, si dimostra facilmente che i tre modelli sono equivalenti.

REFERENCES

Altham, P. (1976), "Discrete variable analysis for individuals grouped into families," *Biometrika,* 63, 263–269.

Campbell, M.J., Machin, D., D'Arcagues, C. (1991), "Coping with extra Poisson variability in the analysis of factors influencing vaginal ring expulsion," *Statistics in Medicine,* 10, 241–254.

Crowder, MJ. (1979) "Inference about the intraclass correlation coefficient in beta–binomial ANOVA for proportions," *Journal of the Royal Statistical Society* (B), 41, 230–234.

Johnson, NL. Kotz, S. (1970) "Distribution in statistics," *Continuous univariate distribution – 2.* Boston, Houghton and Mifflin 63(112–212).

Lambert, D. (1992) "Zero–inflated Poisson regression, with an application to defects in manufacturing," *Technometrics,* 34, 1–14.

Williams, DA. (1975) "The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity," *Biometrics,* 31, 949–952.

Williams, DA. (1982) "Extra–binomial variation in logistic linear models," *Applied Statistics,* 31, 144–148.

Williams, DA. (1988) "Extra–binomial variation in toxicology," Proceedings of the *XIV International Biometric Conference.* Namur, July 18–23.

Zucker, D., Wittes, J. (1991) *"Testing the effect of treatment in experiments with correlated binary outcomes,"* Personal communication in text.