Bacon With Your Eggs? Applications of a New Bivariate Beta-Binomial Distribution

Peter J. DANAHER and Bruce G. S. HARDIE

We present two everyday applications of a new bivariate beta-binomial distribution. Although the applications are familiar, they share unique characteristics that cannot be handled adequately by existing bivariate discrete distributions. These features are high levels of between- and within-trial correlation for the bivariate random variables. Our model is derived from a broad class of bivariate models based on a versatile, but little-known, family of distributions due to Sarmanov. For the applications presented, we show that our model fits observed data very well. In addition, for the bacon and eggs application, the model can be used to help improve the management of inventory in grocery stores.

KEY WORDS: Beta-binomial; Bivariate discrete distribution; Sarmanov distribution.

1. INTRODUCTION

Even in this diet-conscious age, an enduring favorite breakfast choice is bacon and eggs (Backas 2003). Since bacon is often *eaten* with eggs, it seems reasonable to expect that it is often *purchased* with eggs. Consider a succession of shopping trips made by a sample of household shoppers. Each shopper has a particular propensity to buy bacon and/or eggs. In this setting we might observe a correlation in the purchase incidence of bacon or eggs across successive shopping trips, as well as a correlation between bacon and egg purchases. For example, when observing ten successive shopping trips, one person might buy eggs seven times, while another buys eggs only once. The first person clearly has a higher propensity to purchase eggs, resulting in the conditional probability of an eggs purchase for the present trip, given a purchase on the previous trip being higher than the marginal probability of purchasing eggs on any one trip. Such a sequence of egg purchases gives rise to correlated binary data, which are typically observed in repeated observations on the same person, where the occurrence of a particular event is correlated from one trial to the next. There are several models available for analyzing correlated binary data (see, e.g., Altham 1978 and Kupper and Haseman 1978), but these are only for univariate random variables. For analyzing bacon and eggs purchases together, a bivariate correlated model is required.

In the case of bacon and eggs, there are two binary observations per shopper at each shopping trip (trial), repeated over a number of trials. Hence, bacon and egg purchases can potentially be correlated both within each trial and across repeated trials. The purpose of this article is to develop a simple model that accurately accounts for both the inter- and intra-trial correlations for bivariate Bernoulli trials. We show that the model we derive is equivalent to a previous model based on canonical expansions, but the new model has a much more intuitive derivation and has much broader application. Our model can be used to first establish whether there is a correlation between bacon and egg purchases and second to help grocery stores manage their inventory when sales of different (but potentially related) product categories are correlated. For instance, if there is a correlation between bacon and egg purchases and a price promotion stimulates sales of eggs by 50% or more, what is the expected increase in the sales of bacon? Anticipating such sales changes enables store managers to preorder more bacon, thereby avoiding the opportunity cost of lost sales due to stockouts.

We further illustrate our model with a second example looking at people's reading habits for pairs of magazines. Two magazines might have related content and therefore be jointly read by many people, giving rise to a correlation between each pair of magazines (inter-magazine correlation). In addition, magazines often have high reading loyalty from one issue to the next (Danaher 1992) and this gives rise to high correlation across successive issues of the same magazine, known as intra-magazine correlation (Danaher 1989, 1992). Ignoring either of these correlations results in a poor estimate of the audience for an advertising campaign that places advertisements in both magazines.

2. A BIVARIATE BETA-BINOMIAL DISTRIBUTION

We begin with a univariate distribution for the number of successes in k trials. Let Y_j denote a Bernoulli random variable, where $Y_j = 1$ if the jth trial is a "success," with probability p. Then $X = \sum_{j=1}^{k} Y_j$ has a $\text{bin}(k, p)$ distribution. It is well known that this simple binomial model does not capture overdispersion in the variance of X, which has been postulated to result from correlation among the Y_js over successive trials. Altham (1978) and Kupper and Haseman (1978) allowed for this by developing correlated binomial distributions. An alternative and more common explanation for the correlated binary observations is one of heterogeneity. Here it is conjectured that each person has their own personal probability of success (p), with

Peter J. Danaher is Professor of Marketing at the University of Auckland, Department of Marketing, University of Auckland, Private Bag 92019, Auckland, New Zealand (E-mail p.danaher@auckland.ac.nz). Bruce G. S. Hardie is Associate Professor of Marketing, London Business School, London Business School, Sussex Place, London NW1 4SA, U. K. (E-mail: bhardie@london.edu).

DOI: 10.1198/000313005X70939

p itself being a random variable. Under this scenario it is assumed that $Y_j|p$ are Bernoulli trials, so that $X|p \sim \text{bin}(k, p)$. The *unconditional* correlation among the binary observations is captured by allowing p to vary according to a beta distribution, leading to a beta-binomial distribution for the distribution of X.

We now generalize to the bivariate case by defining $X_i = Y_{i1} + Y_{i2} + \cdots + Y_{ik_i}$, where Y_{ij} is a Bernoulli random variable indicating success with probability p_i and $X_i = 0, 1, \ldots, k_i$, $i = 1, 2$. As mentioned earlier, for either of the marginal distributions, a number of examples have been observed where $\text{corr}(Y_{ij}, Y_{ij'}) \neq 0$. These situations are handled reasonably well by allowing X_i to have a beta-binomial distribution. In addition to possible correlation within trials, it might happen that there is correlation between X_1 and X_2. Examples were given by Johnson, Kotz, and Balakrishnan (1997) and Kocherlakota and Kocherlakota (1992). However, their bivariate models have only binomial marginal distributions and therefore cannot accommodate the within-variable correlation discussed above. In our two examples, it will be seen that correlation both between and within trials is anticipated, so that beta-binomial marginals are required. Hence, previously developed models are inappropriate and require enhancements.

Using an analogous approach to the derivation of the beta-binomial model, we propose that $(X_1, X_2|p_1, p_2)$ be a bivariate binomial distribution where X_1 and X_2 are conditionally independent given (p_1, p_2), where each $X_i|p_i \sim \text{bin}(k_i, p_i)$, $i = 1, 2$. The unconditional correlation between X_1 and X_2 is introduced via a bivariate distribution of (p_1, p_2) which permits correlation between p_1 and p_2 (note that p_1 is normally different from p_2 so these random variables are not exchangeable). Our aim is to develop a bivariate distribution of (p_1, p_2) which has beta marginals, so that the unconditional distribution of (X_1, X_2) has beta-binomial marginals, thus accommodating the within-variable correlation.

We therefore require a joint density for (p_1, p_2) that can accommodate both positively and negatively correlated variates and also has beta marginals. Such a bivariate distribution was proposed by Lee (1996), using the framework introduced by Sarmanov (1966). The general form of the Sarmanov bivariate distribution for (p_1, p_2) with specified marginals $f_1(p_1)$ and $f_2(p_2)$ is given by

$$g(p_1, p_2) = f_1(p_1) f_2(p_2) \left[1 + \omega \phi_1(p_1) \phi_2(p_2)\right],$$

where $\phi_i(p_i)$ is the called the "mixing function," being a bounded nonconstant function such that $\int \phi_i(t) f_i(t) dt = 0$. The parameter ω determines the correlation between p_1 and p_2 and must satisfy the condition $1 + \omega \phi_1(p_1) \phi_2(p_2) > 0$ for all p_1 and p_2 for $g(p_1, p_2)$ to be a joint density function.

For beta marginals, with

$$\begin{aligned}
f(p_i|\alpha_i, \beta_i) &= \frac{1}{B(\alpha_i, \beta_i)} p_i^{\alpha_i - 1} (1 - p_i)^{\beta_i - 1}, \ \alpha_i, \beta_i > 0, \\
E(p_i) &= \frac{\alpha_i}{\alpha_i + \beta_i} \equiv \mu_i, \\
\text{var}(p_i) &= \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i)^2 (\alpha_i + \beta_i + 1)} \\
&= \frac{\mu_i (1 - \mu_i)}{(\alpha_i + \beta_i + 1)} \equiv \sigma_i^2,
\end{aligned}$$

Lee (1996) proposed a mixing function of the form $\phi_i(p_i) = p_i - \mu_i$. The resulting bivariate beta distribution is

$$\begin{aligned}
g(p_1, p_2) = {} & f(p_1|\alpha_1, \beta_1) f(p_2|\alpha_2, \beta_2) \\
& \times \left[1 + \omega (p_1 - \mu_1)(p_2 - \mu_2)\right]. \quad (1)
\end{aligned}$$

When $\omega = 0$, this collapses to the product of two univariate beta densities (i.e., p_1 and p_2 are independent). It is easy to show that the correlation between p_1 and p_2 is given by

$$\text{corr}(p_1, p_2) = \omega \sigma_1 \sigma_2. \quad (2)$$

Hence, under this model, p_1 and p_2 are independent if $\text{corr}(p_1, p_2) = 0$ (assuming $\sigma_1^2, \sigma_2^2 > 0$).

We obtain the unconditional distribution of (X_1, X_2) by mixing the distribution of (X_1, X_2), conditional on (p_1, p_2) with the Sarmanov bivariate beta distribution as given in Equation (1):

$$\begin{aligned}
& P(X_1 = x_1, X_2 = x_2 | k_1, k_2, \alpha_1, \alpha_2, \beta_1, \beta_2, \omega) \\
&\quad = \int_0^1 \int_0^1 P(X_1 = x_1 | k_1, p_1) \\
&\quad \quad \times P(X_2 = x_2 | k_2, p_2) g(p_1, p_2) \, dp_1 dp_2 \\
&\quad = P_{\text{BB}}(X_1 = x_1 | k_1, \alpha_1, \beta_1) P_{\text{BB}}(X_2 = x_2 | k_2, \alpha_2, \beta_2) \\
&\quad \quad \times \left[1 + \omega \frac{(x_1 - k_1 \mu_1)(x_2 - k_2 \mu_2)}{(\alpha_1 + \beta_1 + k_1)(\alpha_2 + \beta_2 + k_2)}\right] \quad (3)
\end{aligned}$$

where $P_{\text{BB}}(X_i = x_i | k_i, \alpha_i, \beta_i)$ is the pmf of a beta-binomial distribution with parameters α_i and β_i, given by

$$P_{\text{BB}}(X_i = x_i | k_i, \alpha_i, \beta_i) = \binom{k_i}{x_i} \frac{B(\alpha_i + x_i, \beta_i + k_i - x_i)}{B(\alpha_i, \beta_i)}.$$

To ensure that the pmf given by Equation (3) is nonnegative, ω must satisfy the condition

$$\begin{aligned}
& \frac{(\alpha_1 + \beta_1 + k_1)(\alpha_2 + \beta_2 + k_2)}{k_1 k_2} \\
& \quad \times \max \left[\frac{-1}{\mu_1 \mu_2}, \frac{-1}{(1 - \mu_1)(1 - \mu_2)}\right] \leq \omega \\
& \quad \leq \frac{(\alpha_1 + \beta_1 + k_1)(\alpha_2 + \beta_2 + k_2)}{k_1 k_2} \\
& \quad \times \min \left[\frac{1}{\mu_1 (1 - \mu_2)}, \frac{1}{(1 - \mu_1) \mu_2}\right].
\end{aligned}$$

This is obtained by setting (x_1, x_2) to be at the extreme points of $(0, 0)$, $(0, k_2)$, $(k_1, 0)$, and (k_1, k_2) in Equation (3). Practically it can imply that the correlation between X_1 and X_2 is narrower than $(-1, 1)$. This is a common problem with nonnormal bivariate distributions, such as the Farlie-Gumbel-Morgenstern distribution (Lee 1996). However, Shubina and Lee (2004) showed that the Sarmanov distribution suffers less from this problem than many other bivariate distributions.

Note that alternative "mixing functions," $\phi_i(p_i)$, will yield different bivariate beta distributions—see, for example, Cole, Lee, Whitmore, and Zaslavsky (1995). This will lead to different constraints on the parameters and alternative expressions for $\text{corr}(p_1, p_2)$.

The covariance between X_1 and X_2 is given by

$$\text{cov}(X_1, X_2) = \omega k_1 k_2 \frac{\mu_1 (1 - \mu_1)}{(\alpha_1 + \beta_1 + 1)} \frac{\mu_2 (1 - \mu_2)}{(\alpha_2 + \beta_2 + 1)}. \quad (4)$$

As this bivariate (correlated) beta-binomial distribution has beta-binomial marginals,

$$E(X_i) = k_i \mu_i, \tag{5}$$

and

$$\begin{aligned} \mathrm{var}(X_i) &= k_i \frac{\mu_i(1-\mu_i)(\alpha_i+\beta_i+k_i)}{(\alpha_i+\beta_i+1)} \\ &= k_i \sigma_i^2 (\alpha_i+\beta_i+k_i). \end{aligned} \tag{6}$$

Given Equations (2), (4), and (6), we have

$$\mathrm{corr}(X_1, X_2) = \mathrm{corr}(p_1, p_2) \sqrt{\frac{k_1 k_2}{(\alpha_1+\beta_1+k_1)(\alpha_2+\beta_2+k_2)}}. \tag{7}$$

From Equation (2) we note that $\mathrm{corr}(p_1, p_2) = 0$ if and only if $\omega = 0$; this implies $\mathrm{corr}(X_1, X_2) = 0$ if and only if $\omega = 0$. Hence X_1 and X_2 are independent if $\mathrm{corr}(X_1, X_2) = 0$. A further consequence of Equation (7) is that $|\mathrm{corr}(X_1, X_2)| < |\mathrm{corr}(p_1, p_2)|$ and hence the correlation between the observed variables, X_1 and X_2, will always be less than the underlying correlation between the unobserved individual-level success probabilities p_1 and p_2. Because $\mathrm{corr}(X_1, X_2) \to \mathrm{corr}(p_1, p_2)$ as $k_1, k_2 \to \infty$, the sample correlation tends to the latent correlation as the number of trials increases.

2.1 Relationship with the Canonical Expansion Model

As noted earlier, another application where within- and between-trial correlations are important is the reading of two magazines (Rust 1986). Danaher (1992) reviewed previous models used for this application. Only one explicitly accounts for the dual sources of correlation, being Danaher's (1991) "canonical expansion" model. This model uses Lancaster's (1969) general method of canonical expansions to derive a bivariate distribution for two magazines that had beta-binomial marginals. Denoting X_i as the number of issues read for magazine i, Danaher's (1991) canonical expansion model is

$$\begin{aligned} &P_{\mathrm{CE}}(X_1 = x_1, X_2 = x_2 | k_1, k_2) \\ &\quad = P_{\mathrm{BB}}(X_1 = x_1|k_1) P_{\mathrm{BB}}(X_2 = x_2|k_2) \\ &\quad\quad \times \left[1 + \mathrm{corr}(X_1, X_2) \frac{[x_1 - E(X_1)]\,[x_2 - E(X_2)]}{\sqrt{\mathrm{var}(X_1)\mathrm{var}(X_2)}}\right]. \end{aligned} \tag{8}$$

For a particular estimate of ω it happens that this canonical expansion model is identical to one derived via the Sarmanov model given in Equation (3). To see this, substitute Equations (5), (6), and (7) into Equation (3) to obtain

$$\begin{aligned} &P_{\mathrm{Sarmanov}}(X_1 = x_1, X_2 = x_2 | k_1, k_2) \\ &= P_{\mathrm{BB}}(X_1 = x_1|k_1) P_{\mathrm{BB}}(X_2 = x_2|k_2) \\ &\times \left[1 + \omega \frac{\mathrm{corr}(X_1, X_2)\sigma_1\sigma_2}{\mathrm{corr}(p_1, p_2)} \frac{[x_1 - E(X_1)]\,[x_2 - E(X_2)]}{\sqrt{\mathrm{var}(X_1)\mathrm{var}(X_2)}}\right]. \end{aligned} \tag{9}$$

Lee (1996) noted that if the parameters of $f_i(p_i)$ are known (or at least can be estimated separately from the respective univariate distributions), then the maximum likelihood estimator for ω is $\mathrm{corr}(p_1, p_2)/\sigma_1\sigma_2$. Substituting this into Equation (9) gives

$$\begin{aligned} &P_{\mathrm{Sarmanov}}(X_1 = x_1, X_2 = x_2 | k_1, k_2) \\ &\quad = P_{\mathrm{BB}}(X_1 = x_1|k_1) P_{\mathrm{BB}}(X_2 = x_2|k_2) \\ &\quad \times \left[1 + \mathrm{corr}(X_1, X_2) \frac{[x_1 - E(X_1)]\,[x_2 - E(X_2)]}{\sqrt{\mathrm{var}(X_1)\mathrm{var}(X_2)}}\right], \end{aligned} \tag{10}$$

which is the same as Equation (8). Hence the canonical expansion model gives rise to the identical Sarmanov model in this case. Although the final models are the same, the method used to derive them is very different. The canonical expansion model is like a Fourier series expansion of a probability distribution with some terms assumed to be zero in order to obtain Equation (8). On the other hand, Equation (10) results from compounding a bivariate distribution for (p_1, p_2) with the product of two conditionally independent binomial distributions. That is, the Sarmanov approach is rather more elegant as a model because it generalizes naturally from the beta-binomial distribution by compounding a bivariate binomial model conditional on individual-level parameters (p_1, p_2) with a Sarmanov bivariate distribution for (p_1, p_2). The correlation between X_1 and X_2 is introduced through (p_1, p_2) in the Sarmanov approach. In the canonical expansion development, the correlation between X_1 and X_2 is an artifact of the Fourier series coefficients and it is rather a coincidence that the first-order coefficient is the correlation. The second-order and higher canonical expansion coefficients do not have such a natural interpretation.

2.2 Parameter Estimation

We have two suggestions for parameter estimation. One is to use maximum likelihood to simultaneously estimate all five parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2, \omega)$ in Equation (3) by maximum likelihood. The log-likelihood function is

$$\begin{aligned} \mathrm{LL} = &\sum_{x_1=0}^{k_1} \sum_{x_2=0}^{k_2} n_{x_1,x_2} \\ &\times \ln\left[P(X_1 = x_1, X_2 = x_2 | k_1, k_2, \alpha_1, \alpha_2, \beta_1, \beta_2, \omega)\right], \end{aligned}$$

where n_{x_1,x_2} is the observed frequency of occurence of (x_1, x_2). This log-likelihood function can be maximized by Newton's method (we used the Solver facility in Excel without any problems).

An alternative estimation method is to take two steps to create pseudo maximum likelihood parameters (Gong and Samaniego 1981). In the first step, parameters α_1 and β_1 are estimated by maximum likelihood (or the method of moments if that is preferred) using just the observed marginal distribution of X_1. Parameters α_2 and β_2 are estimated similarly for the observed marginal distribution for X_2. In the second step a method of moments estimate of ω is obtainable using the fact that

$$\mathrm{corr}(X_1, X_2) = \omega \sigma_1 \sigma_2 \sqrt{\frac{k_1 k_2}{(\alpha_1+\beta_1+k_1)(\alpha_2+\beta_2+k_2)}}. \tag{11}$$

Table 1. Observed Bivariate Distribution of the Number of Times Bacon and Eggs Were Purchased on Four Consecutive Shopping Trips

	Eggs					
Bacon	0	1	2	3	4	Total
0	254	115	42	13	6	430
1	34	29	16	6	1	86
2	8	8	3	3	1	23
3	0	0	4	1	1	6
4	1	1	1	0	0	3
Total	297	153	66	23	9	548

Using the observed correlation between X_1 and X_2 in Equation (11) and the previously obtained estimates of $\alpha_1, \alpha_2, \beta_1, \beta_2$ from the marginal distributions, gives a reasonable estimate of ω. As it turns out, in our two examples below, both estimation methods give very similar results. Because maximum likelihood works well and has desirable properties, such as the ability to produce standard errors, we suggest the use of this method and so we only report the MLE results.

3. EXAMPLES

We now fit the bivariate beta-binomial distribution to two datasets: one for the purchase of bacon and eggs, the other for the reading of two magazines.

3.1 Example 1: Purchasing of Bacon and Eggs

We first consider the purchase of bacon and eggs: is a household's number of bacon purchases correlated with its number of egg purchases? We use data from Information Resources, Inc., a consumer panel based in a large U.S. city [see Bell and Lattin (1998) for further details]. Starting in June 1991, we track purchases in the bacon and fresh eggs product categories for a sample of 548 households over four consecutive store trips. We consider only those grocery shopping trips with a total basket value of at least five dollars. For each household, we count the total number of bacon purchases in their four eligible shopping trips and the total number of egg purchases (usually a package of eggs) for the same trips.

The observed bivariate distribution of the number of times bacon and eggs were purchased on these four consecutive shopping trips is reported in Table 1. We see that $254/548 = 46\%$ of the sample purchased neither bacon nor eggs. By contrast, three households purchased bacon on all four shopping trips, while nine households purchased eggs on every shopping trip.

The maximum likelihood parameter estimates are: $\hat{\alpha}_{\text{bacon}} = .357$, $\hat{\beta}_{\text{bacon}} = 4.455$, $\hat{\alpha}_{\text{eggs}} = .858$, $\hat{\beta}_{\text{eggs}} = 3.981$, and $\hat{\omega} = 25.290$. We undertook a chi-squared goodness-of-fit test to see how well the observed distribution matched the estimated distribution. The result is encouraging, as the χ^2 statistic (with 9 df) is only 15.2, resulting in a p value of .09, which is nonsignificant at the 5% level. Additional goodness-of-fit tests on just the marginal distributions are also not significant, with very small χ^2 values, indicating that any lack of fit is due more to the joint structure than the beta-binomial marginals.

Under the full model the value of the log-likelihood function is -995.2. When ω is constrained to zero, the log-likelihood reduces to -1007.9, showing that ω is significantly nonzero (the likelihood ratio test is $2 \times (1007.9 - 995.2) = 25.4 > 3.84 = \chi_1^2$). Hence, there *is* a statistically significant correlation between the number of bacon and egg purchases. Indeed, the observed correlation between the number of trips in which bacon was purchased and the number of trips in which eggs were purchased is .23. The estimated correlation between p_{bacon} and p_{eggs} is .43, being higher than the observed correlation, as anticipated by Equation (7).

Table 2. Estimated Bivariate Distribution for the Probability of Bacon and Eggs Purchases on a Single Shopping Trip

	Eggs		
Bacon	0	1	Total
0	.7691	.1566	.9258
1	.0536	.0206	.0742
Total	.8227	.1773	1.0

We can use the fitted model to predict the probability of, for example, a bacon purchase given an eggs purchase on a particular shopping trip. This is achieved by setting $k_1 = k_2 = 1$ and using the parameter estimates above. This estimated bivariate distribution is given in Table 2. Using the partition theorem and the probabilities in Table 2 it easy to show that

$$\Pr(\text{buy bacon}) = \frac{.0536}{.8277} + \left[\frac{.0206}{.1773} - \frac{.0536}{.8277}\right] \Pr(\text{buy eggs}).$$

During a sales promotion of eggs, often achieved by lowering the price or offering two-for-one deals, the sales of eggs could double or triple in the promotion week (Bell and Lattin 1998). Given that bacon purchases are associated with egg purchases, a large increase in egg sales is likely to also increase sales of bacon. If the grocery store does not plan for this, bacon stocks can run out, resulting in lost potential revenue. If it happens that egg sales are doubled, so that the probability of an eggs purchase becomes $2 \times .1773$, then the predicted probability of a bacon purchase in that same week rises to .0833 from .0742, which is a 12.2% increase in the probability of a bacon purchase. Hence, a wise retailer can use this information to predict the downstream demand on bacon sales when eggs are promoted and avoid a bacon stockout.

3.2 Example 2: Reading of Two Monthly Magazines

We now consider the joint readership of two magazines: is an individual's probability of reading one magazine correlated with their probability of reading another? We use data from the "National Media Survey" conducted by AGB:McNair New Zealand [see Danaher (1989, 1991) for a more detailed description of these data and the particular questions asked of the survey respondents]. For a sample of 3,000 individuals, we know how many issues of two monthly magazines (*Auto Age* and *Signature*) are read in a six-month period. The observed bivariate distribution of the number of times each magazine was read is reported in Table 3. Some 9.9% of the sample read every issue of *Auto Age* while 2.6% read every issue of *Signature*, with only 22 people reading every issue of both magazines.

For these data the parameter estimates are: $\hat{\alpha}_{\text{AA}} = .012$, $\hat{\beta}_{\text{AA}} = .092$, $\hat{\alpha}_{\text{Sig}} = .008$, $\hat{\beta}_{\text{Sig}} = .191$, and $\hat{\omega} = 2.384$. A good

Table 3. Observed Bivariate Distribution of the Number of Issues of *Auto Age* and *Signature* Read in a Six-Month Period

	Signature							
Auto Age(AA)	0	1	2	3	4	5	6	Total
0	2463	20	17	8	8	3	52	2571
1	35	2	2	0	0	0	2	41
2	44	2	0	0	1	0	1	48
3	14	0	2	1	1	0	0	18
4	16	0	0	0	0	0	2	18
5	7	0	0	0	0	0	0	7
6	262	2	3	4	3	1	22	297
Total	2841	26	24	13	13	4	79	3000

fit to the observed distribution is evidenced by a nonsignificant chi-squared goodness-of-fit test ($\chi^2 = 9.1, p = .06$), at the 5% level (some collapsing across the 1 to 5 exposure levels is required due to sparse data). Under the full model the value of the log-likelihood function is -2552.9, which reduces to -2569.0 when ω is constrained to zero. A likelihood ratio test confirms that the readership of these two magazines is significantly associated. In fact, the estimated correlation between p_{AA} and p_{Sig} is .129, which is just higher than the observed correlation between the readership of the two magazines, being .119.

The importance of allowing for possible correlation between magazine readership is best demonstrated by predicting the audience size of an advertising campaign when advertisements are placed in both magazines for six months. Advertisers are particularly interested in the *reach* of a campaign, which is the proportion exposed to either of the two ads (Danaher 1992). Reach after six months is estimated using Equation (3) with $k_1 = k_2 = 6$ to obtain

$$\text{Reach} = 1 - P_{\mathrm{BB}}(X_1 = 0|k_1 = 6)P_{\mathrm{BB}}(X_2 = 0|k_2 = 6) \times \left[1 + \omega \frac{(0 - 6\mu_1)(0 - 6\mu_2)}{(\alpha_1 + \beta_1 + 6)(\alpha_2 + \beta_2 + 6)}\right].$$

The estimated reach when the between-magazine correlation is ignored ($\omega = 0$) is 18.9%, which reduces to 18.0% when the estimated value of ω is used. Table 3 enables us to calculate the "true" reach as $1 - 2463/3000 = 17.9\%$. Clearly, the full model-based estimate of reach (18.0%) is very close to the "true" observed value than the estimate which ignores the pairwise correlation between magazines. In general, ignoring ω results in an overestimate of reach. Although the difference may seem small for these two magazines, in cases where the correlation is higher, the difference in reach predictions with and without ω can be as much as nine percentage points. Overestimating reach by this amount would result in campaign effectiveness expectations (such as sales) being much higher than would actually be observed.

4. CONCLUSION

We present two examples where there are repeated trials of a binary process, namely, buy/do not buy and read/do not read, where there is correlation both between and within trials. Although the applications are everyday occurrences, no existing bivariate model has been developed that fully captures the underlying dynamics of these buying and reading events. Our model can handle both correlations in an elegant and intuitive way. Not only does the model fit the data well, but it can also be used to make predictions that result in more efficient business decision making.

The heart of our model is a little-known, but versatile, class of distributions initially developed by Sarmanov (1966) and used firstly by Lee (1996). In our applications, the Sarmanov model is used to create a bivariate beta distribution, but, as Lee (1996) pointed out, many other bivariate Sarmonov distributions are possible, such as bivariate gamma and survival distributions. Another recent application of the Sarmanov distribution has been to examine visits to related Web sites like amazon.com and barnesandnoble.com (Park and Fader 2004). We hope this article stimulates further interest in the Sarmanov family of distributions to other applications that require bivariate distributions with known marginals, but where correlation between the marginal random variables is anticipated.

[Received March 2005. Revised June 2005.]

REFERENCES

Altham, P. M. E. (1978), "Two Generalizations of the Binomial Distribution," *Applied Statistics*, 27, 162–167.

Backas, N. (2003), "Awakening New Breakfast Ideas." Available online at http://www.foodproductdesign.com/archive/2003/0603FFOC.html.

Bell, D. R., and Lattin, J. M. (1998), "Shopping Behavior and Consumer Preference for Store Price Format: Why 'Large Basket' Shoppers Prefer EDLP," *Marketing Science*, 17, 66–88.

Cole, B. F., Lee, M.-L. T., Whitmore, G. A., and Zaslavsky, A. M. (1995), "An Empirical Bayes Model for Markov-Dependent Binary Sequences With Randomly Missing Observations," *Journal of the American Statistical Association*, 90, 1364–1372.

Danaher, P. J. (1989), "Markov Mixture Model for Magazine Exposure," *Journal of the American Statistical Association*, 84, 922–926.

——— (1991), "A Canonical Expansion Model for Multivariate Media Exposure Distributions: A Generalization of the 'Duplication of Viewing Law'," *Journal of Marketing Research*, 28, 361–367.

——— (1992), "Some Statistical Modeling Problems in the Advertising Industry: A Look at Media Exposure Distributions," *The American Statistician*, 46, 254–260.

Gong, G., and Samaniego, F. J. (1981), "Psuedo Maximum Likelihood Estimation: Theory and Applications," *Annals of Mathematical Statistics*, 9, 861–869.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1997), *Discrete Multivariate Distributions*, New York: Wiley.

Kocherlakota, S., and Kocherlakota, K. (1992), *Bivariate Discrete Distributions*, New York: Marcel Dekker.

Kupper, L. L., and Haseman, J. K. (1978), "The Use of a Correlated Binomial Model for the Analysis of Certain Toxicological Experiments," *Biometrics*, 34, 69–76.

Lancaster, H. O. (1969), *The Chi-Squared Distribution*, New York: Wiley.

Lee, M.-L. T. (1996), "Properties and Applications of the Sarmanov Family of Bivariate Distributions," *Communications in Statistics—Theory and Methods*, 25, 1207–1222.

Park, Y.-H., and Fader, P. S. (2004), "Modeling Browsing Behavior at Multiple Websites," *Marketing Science*, 23, 280–303.

Rust, R. T. (1986), *Advertising Media Models: A Practical Guide*, Lexington, MA: Lexington Books.

Sarmanov, O. V. (1966), "Generalized Normal Correlation and Two-Dimensional Frechet Classes," *Doklady (Soviet Mathematics)*, 168, 596–599.

Shubina, M., and Lee, M. T. (2004), "On Maximal Attainable Correlation and Other Measures of Dependence for the Sarmanov Family of Distributions," *Communications in Statistics—Theory and Methods*, 33, 1031–1052.