



ACADEMIC  
PRESS

Available at  
**WWW.MATHEMATICSWEB.ORG**  
POWERED BY SCIENCE @ DIRECT®

Journal of Multivariate Analysis 87 (2003) 356–369

---

---

Journal of  
**Multivariate  
Analysis**

---

---

<http://www.elsevier.com/locate/jmva>

## Factor models for multivariate count data

Michel Wedel,<sup>a,\*</sup> Ulf Böckenholt,<sup>b</sup> and Wagner A. Kamakura<sup>c</sup>

<sup>a</sup> *University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109, USA*

<sup>b</sup> *Faculty of Management, McGill University, 1001 Sherbrooke St. West, Montreal, Quebec, Canada H3A 1G5*

<sup>c</sup> *Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708-0120, USA*

Received 1 November 2001

---

### Abstract

We develop a general class of factor-analytic models for the analysis of multivariate (truncated) count data. Dependencies in multivariate counts are of interest in many applications, but few approaches have been proposed for their analysis. Our model class allows for a variety of distributions of the factors in the exponential family. The proposed framework includes a large number of previously proposed factor and random effect models as special cases and leads to many new models that have not been considered so far. Whereas previously these models were proposed separately as different cases, our framework unifies these models and enables one to study them simultaneously. We estimate the Poisson factor models with the method of simulated maximum likelihood. A Monte-Carlo study investigates the performance of this approach in terms of estimation bias and precision. We illustrate the approach in an analysis of TV channels data.

© 2003 Elsevier Science (USA). All rights reserved.

*AMS 2000 subject classifications:* 62H25

*Keywords:* Poisson distribution; Simulated likelihood; Link function

---

### 1. Introduction

Multivariate cross-sectional count data occur commonly in cases where incidences of several related events are measured by counting them. If such multivariate counts are available, interest is often in the investigation of dependencies among them, but

---

\*Corresponding author.

*E-mail address:* [wedel@umich.edu](mailto:wedel@umich.edu) (M. Wedel).

applications of methods for the analysis of multivariate count data are relatively rare. This has probably been caused by computational intractability [11, p. 252], even though recent developments in simulation methods for estimation enable such analyses, as illustrated in the work of Chib and Winkelmann [12].

Although information on clusters of variables that are highly correlated, can also be obtained from direct inspection of the estimated covariance matrix as in [12], a powerful tool for the investigation of dependencies in multivariate data is factor analysis. Factor-analytic representations offer insights into dependencies among observed variables based on a reduction of the data to a few important underlying factors and graphical displays of the estimated factor weights that often have interesting substantive interpretation. Factor analysis is probably one of the most frequently used multivariate analysis tools. However, direct applications of standard factor analysis to the analysis of counts neglect the discrete nature of count data. Lawley [17] first accommodated assumptions on the distribution of observed variables and thereby introduced factor analysis to the area of statistical modeling. Whereas he developed maximum likelihood estimation for normally distributed variables, Bartholomew [6] later proposed maximum likelihood factor analysis for binary variables.

Here, we present factor models to explore the dependencies in multivariate count data. We develop a general class of factor-analytic models that provides a parsimonious and easy to interpret representation of multivariate dependencies in counts. The models we provide extend the general linear latent variable model given in Bartholomew and Knott [17, p. 19, 25, 159] and Wedel and Kamakura [27]. Our contribution is that we allow for a more general class of distributions and link functions and propose a new method for estimating these models that can be used even for large numbers of factors. Other relevant work in this area is that of Arminger and Küsters [5], who provide simultaneous equation models for observed variables of any measurement level and with metric latent endogenous variables. Their work is restricted to one-factor models, however, without considering full maximum likelihood estimation of their models. Our work does include multiple factors and employs maximization of the (simulated) likelihood function for parameter estimation, but does not consider structural relations among the latent variables.

By accommodating a variety of different distributional assumptions for the factors, we develop a general framework that includes many both previously proposed and new models as well as compound distributions. Whereas so far these models needed to be studied separately as a collection of special cases, our framework enables one to study them simultaneously. The approach that we take fits in the econometric tradition in of studying unobserved heterogeneity due to omitted variables. Although the full covariance matrix of the latent errors can be estimated without imposing any restriction [12], the factor approach can provide a more parsimonious specification and easier interpretation. We show that count data factor models can be estimated effectively by the method of simulated likelihood [13].

We proceed by describing the Poisson factor models, along with a host of previously published compound distributions arising as special cases. We then

describe estimation by the method of simulated likelihood and discuss the selection of the number of factors. A Monte-Carlo study is conducted to investigate the performance of the simulated likelihood estimator compared to numerical integration. The paper concludes with an illustrative application to TV channel viewing data.

**2. Poisson factor models for count data**

*2.1. Model description*

Consider an  $I \times J$  matrix of counts  $Y=(y_{ij})$  classified by respondents  $i$  and variables  $j$ . Next to the standard Poisson distribution, we accommodate left-, right- and doubly truncated Poisson distributions that are useful in several applications [15, p. 181–186], and for which the probability functions are, respectively

$$\begin{aligned} \Pr(Y = y_{ij}) &= \frac{\exp(-\theta_{ij})\theta_{ij}^{y_{ij}}}{y_{ij}!} \left[ 1 - \exp(-\theta_{ij}) \sum_{k=0}^L \frac{\theta_{ij}^k}{k!} \right]^{-1}, \\ \Pr(Y = y_{ij}) &= \frac{\theta_{ij}^{y_{ij}}}{y_{ij}!} \left[ \sum_{k=0}^R \frac{\theta_{ij}^k}{k!} \right]^{-1}, \\ \Pr(Y = y_{ij}) &= \frac{\theta_{ij}^{y_{ij}}}{y_{ij}!} \left[ \sum_{k=L}^R \frac{\theta_{ij}^k}{k!} \right]^{-1}, \end{aligned} \tag{2.1}$$

for left and right truncation points  $L$  and  $R$ . To link the mean for subject  $i$ ,  $\theta_i=(\theta_{ij})$ , to factors, we parameterize it as

$$g(\theta_i) = \mu + x_i\Gamma', \tag{2.2}$$

where  $\mu=(\mu_j)$  is a vector of intercepts,  $x_{ip}$  represents the value of respondent  $i$  on the  $p$ th latent factors,  $x_i$  is a  $P$ -dimensional row vector, and  $\Gamma=(\gamma_{jp})$  is a  $(J \times P)$  matrix of fixed factor weights. Although many other distributions are easily accommodated, we currently restrict ourselves to the following two distributions for the factors  $x_{ip}$ :

- (1)  $x_{ip} \sim G(v_p, \lambda_p)$ , a Gamma distribution with mean  $v_p$  and variance  $\lambda_p$ , and
- (2)  $x_{ip} \sim N(v_p, \lambda_p)$ , a Normal distribution with mean  $v_p$  and variance  $\lambda_p$ .

We consider a range of possible factor models through various structures on the weight matrix  $\Gamma$ , i.e., we allow for both common and unique factor models and for exploratory and confirmatory factor models. For example, if  $\Gamma^1=(\gamma_{jp})$  is a full matrix of rank  $P$ , we have an exploratory factor model with  $P$  factors. For  $\Gamma^0 = \text{diag}(\gamma_j)$  of rank  $J$ , the model has unique factors only (i.e., an independent factor for each variable). For  $\Gamma'=(\Gamma^{0'}, \Gamma^{1'})$ , the model has both common and unique factors. Confirmatory factor models with  $P$  hypothesized factors are intermediate

between the common and unique factors model and arise when some elements of  $\Gamma$  are constrained to zero (often, for example, only one element in each row is free) according to prior hypotheses. The link function is  $g(\cdot)$ . The two most commonly used link functions are the identity link and the log link, but other links, such as the square root, are also possible. In our framework, any link function  $g(\cdot)$ , can be used, in contrast to Bartholomew and Knott [7], who use the canonical link function. In choosing the link function in (2.2) it is important to apply a function that constrains the expectation of the conditional Poisson distribution to be positive, which makes certain link functions, such as the identity link, less suitable.

The marginal distribution of  $y_i$  is obtained by integrating over the distribution of the unobserved factors:

$$f(y_i|\Theta) = \int \prod_{j=1}^J f(y_{ij}|x_i; \Theta) f(x_i|\Xi) dx_i, \tag{2.3}$$

$\Xi$  are the parameters characterizing the distribution of  $x_i$  ( $v_p$  and  $\lambda_p$ ), and  $\Theta$  contains all parameters of interest. Note that the multivariate distributions generated by (2.3) have univariate marginals in the same family [11, p. 258]. Conditional upon  $x_i$ , the attractive properties of the Poisson model apply, i.e.:  $f(y_{i*}|x_i; \Theta)$  is Poisson and  $f(y_{ij}|x_i, y_{i*}; \Theta)$  is multinomial, where  $y_{i*} = \sum_j y_{ij}$ .

### 2.2. Invariance properties

The fit of a factor model with Normal factors is not affected by changes in the location and scale of the variables as well as orthogonal rotations of the factor weights. Location invariance arises since under the parameterization  $\mu_j^* = \mu_j + m$ , and  $v_p^* = v_p - m$ , the distribution of  $y_{ij}$ ,  $f(y_{ij}; \Theta^*)$ , is the same as that under the original parameterization  $f(y_{ij}; \Theta)$ . Location invariance is alleviated by fixing  $\mu$  or  $v$ . Similarly, scale invariance holds since with  $\Gamma^* = \Gamma C$ , with  $C = \text{diag}(c_p)$ , and  $A = A C^{-1}$  with  $A = \text{diag}(\lambda_p)$  the distribution of  $y_{ij}$  is the same as under the original parameterization. Fixing  $\Gamma$  or  $A$  alleviates scale invariance. Rotation invariance (cf. [7,9]), can be shown by introducing an orthogonal ( $P \times P$ ) rotation matrix  $R$  with  $\Gamma^* = \Gamma R$ . For  $x_i^* = x_i R^{-1}$ , the distribution of  $x_i^*$  is equal to  $f_x(x_i^*, R)$ , since the Jacobean is 1 only for  $R$  orthogonal and  $x_i$  normal the unconditional distribution of  $y_{ij}$  under the new parameterization,  $f(y_{ij}; \Theta^*)$ , is the same as that under the original parameterization,  $f(y_{ij}; \Theta)$ . By specifying the distribution of the factors as normal with a zero mean vector and identity covariance matrix, we obtain an identified model. If the distribution of  $x_i$  is Gamma, we both have location and scale invariance; however, the unconditional distribution arising from parameterizations with and without rotation is not the same. This arises because the distribution of  $x_i R^{-1}$  is different from  $x_i$  for  $R$  orthogonal and  $x_i$  Gamma [16]. In other words, a factor-analytic model with Gamma distributed factors is *not* rotation invariant.

Factor models are invariant under re-labeling of the factors. We resolve that problem by reporting one particular sequence of the factors. Unless stated otherwise, we fix  $v$  and  $A$ , specifying  $v=0$  and  $A=I$ , to achieve identification. Local

identification of all models can be checked by inspecting the eigenvalues of the Information matrix [9]. The identifiability of factor models improves when constraints are imposed on the factor weight matrix, as is the case in confirmatory factor models.

### 2.3. Special cases

By specifying the matrix of weights  $\Gamma$ , link functions, and distributions of the factors, interesting special cases of (2.1)–(2.3) arise, some of which have been previously described in the literature. These special cases arise because for known  $\Gamma$  (but  $v_p$  and  $\lambda_p$  free), the factor model becomes a multivariate random intercept model that allows for overdispersion and a restrictive covariance structure of the multivariate counts. We consider a number of cases of interest, displayed in Table 1 (except when explicitly mentioned we set  $\mu=0$ ). In particular, compound Poisson distributions described in [15], such as the Negative Binomial, the Negative Multinomial, the Poisson-normal, the Poisson-lognormal, and the Poisson-inverse Gaussian arise.

Fig. 1 provides illustrative plots of the probability function for two distribution functions of the factors (Normal and Gamma), several means and variances of the factors, and a variety of different link-functions (the Log-link for the Normally distributed factors, and the Identity, Inverse and Log-links for the Gamma distributed factors). It can be seen from Fig. 1 that the proposed framework accommodates a wide variety of different probability functions. Skewed probability functions that are strongly peaked at zero, and unimodal probability functions that have both high and low kurtosis are revealed. Within each of the four panels, it is apparent that the choice of the link function dramatically affects the form of the joint probability density. Fig. 2 displays contour plots of illustrative bivariate probability functions, computed for  $J=P=2$ , both for Gamma (Identity link) and Normal (Log-link) factors, as well as several choices of the factor weights. It shows that the covariance is strongly affected by the factor weights. Positive weights for both factors result in elongated contour plots that stretch into the first quadrant at an angle determined by the factor weights. Negative weights result in contours of the joint probability function that stretch along the axis of the first quadrant, and large negative weights result in contours that shrink towards zero. This illustrates that quite a variety of dependencies of count data can be accommodated within our framework.

### 2.4. Simulated likelihood estimation

The Poisson factor models can be estimated by maximum likelihood methods. The observed data likelihood is obtained by taking the product of (2.3) over the observations. However, when using standard numerical integration based on, for example, Gaussian quadrature, the estimation of  $\Theta$  is not feasible for many models, given the high-dimensional integration involved in the likelihood. In particular, this problem has severely constrained the estimation of the Poisson-lognormal model

Table 1  
Selected special cases of the compound Poisson factor model

$P$	Link, score distribution	Mean	Weights	Remarks
$P=1$	Identity, Gamma	$\mu$	$\Gamma = \gamma 1_J$	The marginal distribution of $y_i$ is negative multinomial [14]; the marginal distribution $f(y_{i*} \Theta)$ is negative binomial (NBD) and the conditional distribution $f(y_{ij} y_{i*};\Theta)$ is multinomial.
$P=J$	Identity, Gamma	$\mu_j$	$\Gamma = \text{diag}(\gamma_j)$	The marginal distribution of $y_i$ is NBD [15]. The marginal distribution $f(y_{i*} \Theta)$ is also negative binomial, and $f(y_{ij} y_{i*};\Theta)$ is the Dirichlet Multinomial distribution [14].
$P=1$	Identity, Normal	$\mu$	$\Gamma = \gamma 1_J$	Poisson-normal (PN) or Hermite distribution [14]. If the score distribution is truncated at zero the Poisson-zero-truncated normal distribution is obtained [14].
$P \leq J$	Identity, Normal	$\mu_j$	$\Gamma = \text{diag}(\gamma_p)$	For $P=J$ , a multivariate extension of the PN arises (mPN; [26]). For $P < J$ , and $\mu$ and $\Gamma$ unconstrained, the expectation of the Poisson distribution $\theta_i \sim N(\mu, \Gamma \Gamma')$ is multivariate normal and the marginal distribution of $y_i$ is also mPN [14].
$P=1$	Log, Normal	$\mu$	$\Gamma = \gamma 1$	The Poisson log-normal (PLN) distribution [23]. This distribution has also been called the discrete log-normal [14].
$P \leq J$	Log, Normal	$\mu_j$	$\Gamma = \text{diag}(\gamma_p)$	(Multivariate) Poisson-lognormal (mPLN; [1]).
$P < J$	Log, Normal	$\mu_j$	$\Gamma$	The expectation of the Poisson distribution $\theta_i \sim LN(\mu, \Gamma \Gamma')$ is multivariate lognormal. The marginal distribution of $y_i$ is in this case also mPLN.
$P=1$	Identity, Inverse Gaussian	$\mu$	$\Gamma = \gamma 1$	The Sichel, or Poisson-inverse Gaussian distribution [14,25].

proposed by Aitchison and Ho [1]. However, simulated likelihood (SML) estimation provides a feasible and easy to implement alternative to the evaluation of such integrals [13]. In SML the likelihood function is approximated based on  $S$  draws from the distribution of the factors:

$$\tilde{L}(\Theta|y) = \prod_{i=1}^I \sum_{s=1}^S \prod_{j=1}^J f(y_{ij}|z_i^s; \Theta) / S, \tag{2.4}$$

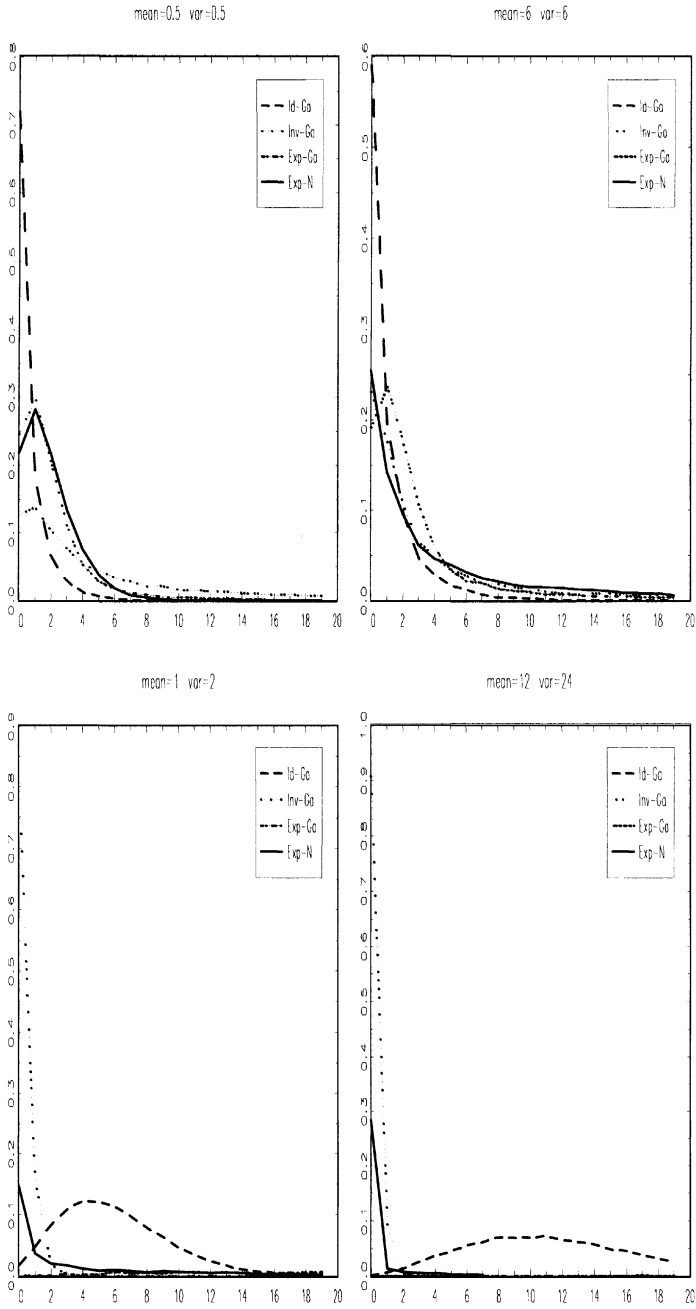


Fig. 1. Plots of the distribution functions of illustrative forms of univariate compound probability functions as a function the link function, and the form and mean and variance of the factor distribution.

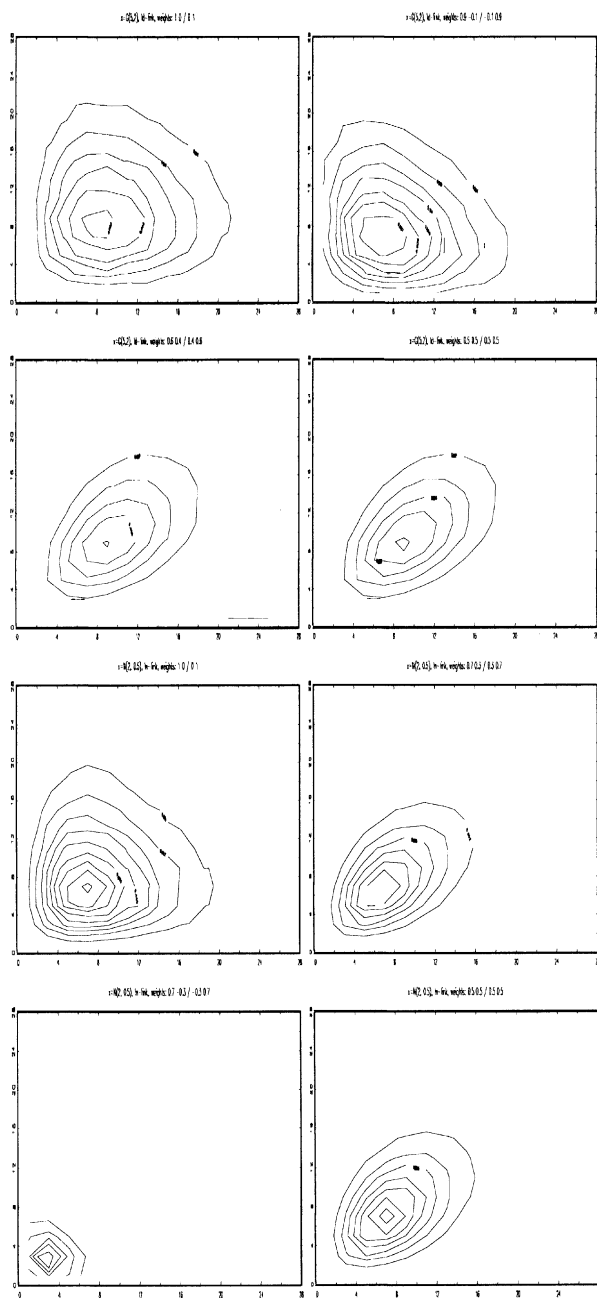


Fig. 2. Contour plots of illustrative forms of bivariate compound Poisson probability functions as a function of the link function, the factor distribution and the factor weights.



where,  $z_i^s$  presents the  $s$ th draw from the distribution of  $x_i$ . The particular distributions of  $x_i$  considered here, the Normal and Gamma distributions, can be easily sampled from, given a preliminary estimate of the parameter values  $\Xi$  characterizing their distribution, where such an estimate is available at each stage of the iterative estimation algorithm. Gouriéroux and Monfort [13] show that the SML estimator is both consistent if  $\sqrt{I}/S \rightarrow 0$  as  $S \rightarrow \infty$  and asymptotically equivalent to the MLE. Simulation studies by Lee [18] (Lee, 1997) have shown that SML has good properties for values of  $S$  exceeding 200, but performance depends on the dimensionality of the integral. We further examine the properties of SML in a Monte-Carlo simulation study in the context of the Poisson factor models below. An appealing aspect of SML estimation is that the simulated likelihood function (2.3) is twice differentiable, simplifying optimization with the Newton-type algorithms that we apply. The optimization of the simulated likelihood function is performed in GAUSS [4], using the Broyden, Fletcher, Goldfarb and Shanno algorithm, with numerical approximation of the derivatives. The draws from the factor distribution are kept constant throughout the optimization.

Anderson and Rubin [3] have investigated consistency and asymptotic normality of ML estimators in factor analysis. For the models considered here such properties follow readily from likelihood theory, since standard regularity conditions hold including that the parameter space is compact and the parameter estimates are unique. Thus, the SML estimators have an asymptotic Normal distribution:  $\hat{\Theta}_{\text{SML}} \sim N(\Theta, H^{-1}UH^{-1})$ . Here,  $H$  is the observed information matrix estimated with the second derivatives of the simulated log likelihood function, and  $U$  is the expected information matrix estimated by the cross product of the first derivatives [13].

### 2.5. Choosing the value of $P$

In some cases, prior constraints may be available to guide the choice of the number of latent factors,  $P$ . For example, if the aim is a graphical display of the dependency structure of the  $J$  variables,  $P=2$  is a convenient choice. However, in many cases the value of  $P$  needs to be determined from the data. Akaike [2] developed his information criterion  $\text{AIC} = -2 \ln L + 2Q$  precisely for this problem where  $Q$  denotes the number of parameters in the model. He showed that AIC can be viewed as an approximate estimate of the expected relative Kullback–Liebler distance between the model and the assumed true one. The standard LR test cannot be applied to test for the number of factors because the  $P$ -factor model is not nested in the  $P+1$ -factor model. An often used procedure is to test the  $P < J$  factor model against the saturated  $P=J$  model. However, at commonly used critical values this LR test is oversensitive to small departures from the null hypothesis due to the large number of degrees of freedom [2]. Although Akaike [2] argued for the AIC statistic to compare models with a different number of factors, it should be taken into account that the AIC statistic does not asymptotically indicate the true model among a set of candidate models. In response, several authors have proposed

dimension-consistent criteria, such as  $CAIC = -2 \ln L + Q(\ln(I) + 1)$  [10] and  $BIC = -2 \ln L + Q \ln(I)$  [24] criteria. Since the criteria lead to similar conclusions, we use one of them only, CAIC. Based on the assumptions that model dimensionality is fixed as  $I \rightarrow \infty$  and that the true model is among the set of candidate models, this statistic indicates the true model with probability one, asymptotically.

### 3. A Monte-Carlo simulation study

To investigate the properties of the SML estimators of Poisson factor model parameters for small- and medium-sized samples, we conduct a simulation study that varies the sample size ( $I=100$  and  $300$ ), the distribution of the factors (Normal and Gamma), and the dimensionality of the model ( $P=1$  and  $2$ ). All models were based on  $J=6$  variables. For each of the  $2^3$  cells, 100 data sets were generated. For the models with normal factors, all intercepts and factor weights were set to  $0.5$  or  $-0.5$ . For the models with Gamma distributed factors, the intercepts were set to  $2$  or  $3$ , the weights to  $0.5$  and  $-0.5$ . The models were estimated both with maximum likelihood using a numerical approximation of the integrals (ML), as well as using simulations to approximate the integrals (SML), with both  $S=50$  and  $250$  draws. The log-link function was used in the estimation. We confine ourselves to one- and two-dimensional models since for those, the numerical integration that we use as a benchmark is accurate. We note that to obtain satisfactory accuracy we used about 600 quadrature points. As measures of performance we compute bias and precision, assessed as the mean deviation of the estimates from the true value of the parameters across 100 replications, and the standard deviation of the estimates across 100 replications, respectively. Table 2 reports the estimated bias and precision based on the mean of the absolute values for the intercepts and factor weights (in order not to let positive and negative bias cancel each other out).

Both, the ML and SML estimates show some bias. However as expected the bias decreases with the sample size and an increasing number of draws for SML. The bias also depends on the dimensionality of the model, being somewhat higher for the  $P=2$  models, as compared to  $P=1$ . For the Gamma factor models, SML estimates are more biased than for normal ones, but again this bias decreases when sample size and number of draws increase. Note that the bias for the intercepts is higher than for the weights, since the values of those parameters are larger. With  $S=250$  draws from the factor distribution, the bias of SML estimates are comparable to that of the ML estimates obtained with numerical integration. The precision of SML, as judged by the standard errors in Table 2, is less than that of ML for  $S=50$ . This holds in particular for two-dimensional integration. But, precision improves if the number of draws increases: SML is as precise as ML or better for  $S=250$  draws, in particular when the sample size is larger ( $n=300$ ) and for the models with Gamma distributed factors.

These results are consistent with earlier simulation studies on SML estimation [18]. SML estimates are biased, but the bias decreases with the number of draws and the number of observations, and increases with the dimensionality of the integration.

Table 2  
Bias and precision of ML and SML of Poisson factor model estimates

Model	Bias						Precision					
	$n=100$			$n=300$			$n=100$			$n=300$		
	ML	$S=50$	$S=250$	ML	$S=50$	$S=250$	ML	$S=50$	$S=250$	ML	$S=50$	$S=250$
<i>Normal, P=1</i>												
$\mu$	0.02	0.01	0.02	0.01	0.01	0.01	0.12	0.13	0.12	0.13	0.07	0.07
$\gamma$	0.02	0.01	0.01	0.01	0.00	0.01	0.12	0.14	0.13	0.13	0.08	0.08
<i>Normal, P=2</i>												
$\mu$	0.02	0.02	0.02	0.01	0.03	0.01	0.13	0.15	0.15	0.14	0.08	0.08
$\gamma$	0.02	0.04	0.06	0.01	0.03	0.03	0.16	0.26	0.23	0.16	0.23	0.19
<i>Gamma, P=1</i>												
$\mu$	0.01	0.03	0.02	0.01	0.04	0.02	0.20	0.21	0.20	0.20	0.13	0.11
$\gamma$	0.02	0.03	0.01	0.01	0.05	0.02	0.13	0.14	0.12	0.11	0.11	0.08
<i>Gamma, P=2</i>												
$\mu$	0.02	0.15	0.02	0.03	0.18	0.03	0.32	0.41	0.30	0.20	0.29	0.20
$\gamma$	0.02	0.02	0.01	0.02	0.05	0.03	0.28	0.27	0.24	0.20	0.25	0.15

The bias is relatively small, in general less than 5%, and may even be negligible for as little as 50 draws. The precision of SML estimates tends to be less than that of ML estimates for a small numbers of draws, but is smaller or equal to that of ML estimates for a larger number of draws ( $S=250$ ). Precision improves with increasing numbers of observations and draws, and may even exceed that of ML estimates, especially when the number of quadrature points is too small. We conclude that for our models SML performs well for  $S=250$  and we use that number of draws in the illustration provided next.

#### 4. Empirical illustration

We illustrate our factor model and analyze data on TV channel viewing, which presents a simple application of our general method but serves the purpose of illustrating it. The number of TV channels has increased dramatically, in the US with the introduction of Cable TV, and in Europe with the advent of commercial TV stations. This has led to a substantial increase in the choice options for viewers and increasing competition among TV channels, which has important implications for TV networks, who are interested in identifying typical patterns of television viewing and in understanding the propensities of consumers to view different channels to enhance their competitive scheduling.

We apply our count-data factor models to analyze daily reach of TV channels in the Netherlands. In media research, “Reach” refers to the number of times a person

Table 3

Estimates and standard errors of the  $P=2$  factor model for the TV station data

	Intercept		Factor 1		Factor 2	
	Estimate	SE	Estimate	SE	Estimate	SE
NL1	0.486	0.156	0.916	0.088	-0.563	0.122
NL2	0.672	0.150	0.937	0.094	-0.730	0.112
NL3	0.469	0.158	0.938	0.092	-0.488	0.116
RTL4	1.854	0.203	0.436	0.145	-1.317	0.150
RTL5	1.416	0.173	0.428	0.134	-1.610	0.161
Veronica	1.513	0.190	0.399	0.140	-1.925	0.194
SBS6	1.624	0.196	0.349	0.171	-2.153	0.223

is exposed to a certain channel or program, during a certain time interval [8, p. 589]. For TV, a relevant interval is often 1 week. The analysis of reach is instrumental in media vehicle selection, as well as for competitive programming of channels [20]. For our application, we use a random sample of 276 households from the representative panel of the market research agency *Intomart*. Panel members agree to have an electronic device, called the people-meter, connected to their TV-set [19]. This device records which channel is viewed at what time, for every 5 s for 24 h a day. We use daily reach data on the seven major TV channels in the Netherlands, NL1, NL2, NL3, RTL4, RTL5, Veronica and SBS6, during a 1-week period. The first three channels are public-, the last four are commercial channels. To investigate the structure of the dependencies in viewing behavior, we fitted several Poisson factor models. We used the standard canonical log-link function throughout. Since the daily reach data are truncated at 7 days, we use a right-truncated Poisson distribution to describe them. An alternative model to describe these data is a binomial factor model (cf. [27]), which we fit as well. The two-factor model with normal factors fits the data best ( $P=1$ : CAIC=6818.6;  $P=2$ : CAIC=6737.1;  $P=3$ : CAIC=6777.2). The ( $P=2$ ) common factor model with Gamma factors provides improved fit (CAIC=6726.1), while the (also  $P=2$ ) binomial factor models with a Gamma factor distribution (CAIC=6818.69) or with a normal factor distribution (CAIC=6733.14) produce a worse fit. For the  $P=2$  Poisson factor model, the deviance residuals did not indicate any systematic lack of fit [21]: the 95% coverage interval of the residuals is  $(-1.971, 1.807)$ . None of the residuals are large in their absolute values (min.  $-2.070$ , max.  $2.180$ ). The pseudo  $R^2=0.242$ .

Table 3 displays the factor weights of the ( $P=2$ ) solution. Note that the solution is not rotated, since the model with Gamma distributed latent variables is not rotation invariant. Nevertheless, the interpretation of the estimates of the weights is clear. All coefficient estimates are significant. The table reveals that the public channels (NL1, NL2 and NL3) have high weights for the first factor. Correspondingly, for the second factor we obtain high weights for the commercial channels. In Table 3, NL2 is the public channel loading highest on the “commercial” factor. This channel is designed to compete with the commercial channels through a higher frequency of

entertainment programs. Thus, our analysis reveals that competitive programming has been effective for audience retention. Conversely, the commercial channels with the highest level of information and news, RTL4 and RTL5, targeted at more affluent and higher educated households, have larger weights for the public channel factor. However, the factor weights show that these public and commercial channels compete in terms of reach only to a limited extent. It shows that the commercial channels tapped into specific entertainment needs that are more difficult to meet through competitive programming of the public channels.

## 5. Conclusion

The initial interest in the Poisson distribution [22] to describe counts focused on univariate and i.i.d. counts, on which a vast literature in statistics and econometrics has emerged [11]. Yet, there is considerable consensus that the Poisson distribution as such, while being a most useful benchmark, has limited practical use, since count data as a rule display overdispersion and contagion. Moreover, associations among multiple event categories are rarely represented well by a simple random-effects structure. Few models have been proposed for the analysis of multivariate count data, which can be accredited to the analytical complications and numerical intractability arising in those investigations. Only recently have researchers begun to systematically study multivariate and contagious count data, but applications of these methods were hampered by numerical problems, in particular, as a result of numerical integration. We expect that the Poisson factor models proposed in this paper will constitute an important tool for the investigation of associations in multivariate count data because they provide a straightforward approach to describe and display the latent dependency structure.

## References

- [1] J. Aitchison, C.H. Ho, The multivariate poisson-log normal distribution, *Biometrika* 76 (1989) 643–653.
- [2] H. Akaike, Factor analysis and AIC, *Psychometrika* 52 (1987) 317–332.
- [3] T.W. Anderson, H. Rubin, Statistical inference in factor analysis, *Proceedings of the Third Berkeley Symposium in Mathematical Statistics and Probability*, Vol. 5, 1956, pp. 111–150.
- [4] Aptech Systems, *GAUSS Maximum Likelihood Manual*, Maple Valley, WA 98038, USA, 1995.
- [5] G. Arminger, U. Küsters, Latent trait models with indicators of mixed measurement level, in: R. Langeheine, J. Rost (Eds.), *Latent Trait and Latent Class Models*, Plenum Press, New York, 1988, pp. 1–76.
- [6] D.J. Bartholomew, Factor analysis for categorical data, *J. Roy. Statist. Soc. B* 42 (1980) 293–321.
- [7] D.J. Bartholomew, M. Knott, *Latent Variable Models and Factor Analysis*, Kendall's Library of Statistics, Edward Arnold, London, 1999.
- [8] R. Batra, J.G. Meyers, D.A. Aaker, *Advertising Management*, Prentice-Hall, New York, 1996.
- [9] P. Bekker, A. Merckens, T.J. Wansbeek, *Identification, Equivalent Models and Computer Algebra*, Academic Press Inc, New York, 1994.

- [10] H. Bozdogan, Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions, *Psychometrika* 52 (1987) 345–370.
- [11] A.C. Cameron, P.K. Trivedi, *Regression Analysis of Count Data*, Cambridge University Press, Cambridge, 1998.
- [12] S. Chib, R. Winkelmann, Markov chain Monte Carlo analysis of correlated count data, *J. Business Econ. Statist.* 19 (2001) 428–435.
- [13] C. Gourieroux, A. Monfort, *Simulation Based Econometric Methods*, Oxford University Press, Oxford, 1997.
- [14] N.L. Johnson, S. Kotz, N. Balakrishnan, *Multivariate Discrete Distributions*, Wiley, New York, 1997.
- [15] N.L. Johnson, S. Kotz, A.W. Kemp, *Univariate Discrete Distributions*, Wiley, New York, 1992.
- [16] H.O. Lancaster, Traces and cumulants of quadratic forms in normal variables, *J. Roy. Statist. Soc. B* 16 (1954) 247–254.
- [17] D.N. Lawley, The estimation of factor loadings by the method of maximum likelihood, *Proc. Roy. Soc. Edinburgh* 61 (1940) 176–185.
- [18] L.F. Lee, Asymptotic bias in simulated maximum likelihood estimation of discrete choice models, *Econ. Theory* (1995) 437–483.
- [19] D. Lu, D.A. Kiewit, Passive people meters: a first step, *J. Advertising Res.* 27 (1987) 9–14.
- [20] G.B. Murray, J.R.G. Jenkins, The concept of effective reach in advertising, *J. Advertising Res.* 32 (1992) 34–42.
- [21] D.A. Pierce, D.W. Schafer, Residuals in generalized linear models, *J. Amer. Statist. Assoc.* 81 (1986) 977–986.
- [22] S.D. Poisson, *Recherches sur la Probabilité des jugements en Matière Criminelle et en Matière Civile*, Paris, Bachelier, 1837.
- [23] F.W. Preston, The commonness and rarity of species, *Ecology* 29 (1948) 254–283.
- [24] G. Schwartz, Estimating the dimension of a model, *Ann. Statist.* 6 (1987) 461–464.
- [25] H.S. Sichel, On a family of discrete distributions particularly suited to represent long-tailed frequency data, in: N.F. Laubscher (Ed.), *Proceedings of the Third Symposium on Mathematical Statistics*, CSIR, Pretoria, 1971, pp. 51–97.
- [26] H.S. Steyn, On the multivariate Poisson normal distribution, *J. Amer. Statist. Assoc.* 71 (1976) 233–236.
- [27] M. Wedel, W.A. Kamakura, Factor models with (mixed) observed and latent variables in the exponential family, *Psychometrika* 66 (2001) 515–530.