ELSEVIER

Statistics & Probability Letters 36 (1997) 127–134

STATISTICS & PROBABILITY LETTERS

Posterior mean identifies the prior distribution in NB and related models

H. Papageorgiou[a], Jacek Wesołowski[b,*]

[a] Department of Mathematics, University of Athens, Panepistemiopolis, 15784 Athens, Greece
[b] Mathematical Institute, Warsaw University of Technology, Plac Politechniki 1, 00-661 Warsaw, Poland

Received 1 October 1996; revised 1 March 1997

Abstract

Bayes negative binomial models under two different parameterizations are shown to be completely identifiable by the form of the Bayes estimates of the parameter. Also power series mixtures are briefly treated. © 1997 Elsevier Science B.V.

AMS classification: 62H05; 62E10

Keywords: Conditional specifications; Regression function; Posterior mean; Identifiability of mixtures; Negative binomial mixtures; Power series mixtures

1. Introduction

For a mixture model (X, θ) denote by $\mu_{X|\theta}$ the conditional distribution of the random variable (rv) X given θ, where θ is a random parameter. The classical problem of identifiability of a mixture model defined by $\mu_{X|\theta}$, see Teicher (1961) or Seshadri and Patil (1964), is connected with the question of unique determination of the prior distribution of θ by the distribution of X. Here we are concerned with another approach involving posterior mean $E(\theta|X)$. Identifiability problems in such a setting have been considered beginning with Korwar (1975). More recent contributions involving different types of mixtures include: Kyriakoussis and Papageorgiou (1991), Arnold et al. (1993), Wesołowski (1995a, b), Gupta and Wesołowski (1997). Here we are interested in negative binomial (nb) mixtures.

Denote by $\mathrm{nb}_1(r, p)$ the nb distribution with the probability mass function (pmf)

$$p_k = \binom{r+k-1}{k} p^r (1-p)^k, \quad k \in \mathbb{N} = \{0, 1, 2, \ldots\},$$

* Corresponding author. E-mail: wesolo@alpha.im.pw.edu.pl.

PII S0167-7152(97)00055-2

where $r>0$, $0<p<1$ and by $\mathrm{nb}_2(r,q)$ the nb distribution with the pmf

$$p_k = \binom{r+k-1}{k} q^k(1+q)^{-(r+k)}, \quad k\in\mathbb{N},$$

$0<q<\infty$. Obviously, nb_2 is just a reparameterization of nb_1, i.e. $\mathrm{nb}_2(r,q)=\mathrm{nb}_1(r,1/(1+q))$. Different random chance mechanisms producing negative binomial distribution with different natural parameterizations, including both the above examples, can be found in Boswell and Patil (1970).

In the context of identifiability via posterior mean $E(\theta|X)$, the nb type mixtures with respect to the parameter $\theta=r$ were considered first. In Cacoullos and Papageorgiou (1982) a result for $\mu_{X|\theta}=\mathrm{nb}_2(\theta,q)$, where θ is a natural valued rv, was announced. Then in Cacoullos and Papageorgiou (1983) $\mu_{X|\theta}=\mathrm{nb}_1(\theta,p)$ was shown to be identifiable for a natural valued rv θ. Also positive solutions for the related conditional distributions $\mu_{X-\theta|\theta}=\mathrm{nb}_1(\theta,p)$ and $\mu_{X-\theta|\theta}=\mathrm{nb}_1(n,p)$ were given in that paper. A mixture with respect to the second parameter was studied in Papageorgiou (1984). It was proved there that the model $\mu_{X|q}=\mathrm{nb}_2(r,q)$, where q is a positive rv, is identifiable via the posterior mean under the restrictive assumption of uniqueness of the solution of the moment problem for both the rv's q and X.

The aim of this note is to consider identifiability of both types of negative binomial mixtures with respect to the second parameter, without any additional conditions of a technical nature. Our approach does not rely on the classical identifiability (in Teicher sense) results, as it was done in the three papers on nb mixtures mentioned above. Thus instead of seeking for the distribution of X and then applying classical identifiability result to conclude that the prior distribution is unique, we try to obtain directly the prior distribution. In Section 2 the nb_1 model is studied. Except of the posterior mean, also an indication on a possibility of using higher posterior moments is given. An extension towards power series mixtures is also considered. Section 3 is devoted to nb_2 mixtures and includes a straight forward extension of Papageorgiou (1984) result, mentioned above. Some applications of the results are presented in Section 4.

2. nb_1 mixtures

In Wesołowski (1995b) it is proved that for the geometric mixture, i.e. in the case $\mu_{X|\theta}=\mathrm{nb}_1(1,1-\theta)$, where the distribution of θ is concentrated on the discrete set $\{1, 1/2, 1/3, \ldots\}$, the posterior mean determines the joint distribution. It is obvious that the analoguous result holds for the mixture $\mu_{X|\theta}=\mathrm{nb}_1(1,\theta)$. Here we begin with a generalization of that result by considering a general nb mixture of the first kind, and with no additional restriction on the support of the mixing parameter.

Theorem 1. *Assume that (X,θ) is a mixture model with*

$$\mu_{X|\theta}=\mathrm{nb}_1(r,\theta), \tag{1}$$

where $r>0$ and $S_\theta=\mathrm{supp}(\theta)\subset[0,1]$. Then the prior distribution of θ is uniquely determined by the posterior mean $E(\theta|X)$.

Proof. Denote by $F_{\theta|X}$ and F_θ, respectively, the distribution functions (df's) of the posterior and prior distributions. By the Bayes formula we have for any $k\in\mathbb{N}$ and any $t\in S_\theta$

$$P(X=k)\,\mathrm{d}F_{\theta|X=k}(t)=P(X=k|\theta=t)\,\mathrm{d}F_\theta(t), \tag{2}$$

where $P(X=k|\theta=t)$ is defined by the relation

$$P(X=k)=\int_{S_\theta} P(X=k|\theta=t)\,\mathrm{d}F_\theta(t), \quad k\in\mathbb{N}$$

(we choose its regular version). Additionally, denote $m(k)=E(\theta|X=k)$, $k\in\mathbb{N}$. Then we have

$$m(k)=\int_{S_\theta} t\,\mathrm{d}F_{\theta|X=k}(t)$$

and by (2) for any $k\in\mathbb{N}$

$$m(k)\int_{S_\theta} P(X=k|\theta=t)\,\mathrm{d}F_\theta(t)=\int_{S_\theta} tP(X=k|\theta=t)\,\mathrm{d}F_\theta(t). \tag{3}$$

Hence (1) implies

$$m(k)\int_{S_\theta} t^r(1-t)^k\,\mathrm{d}F_\theta(t)=\int_{S_\theta} t^{r+1}(1-t)^k\,\mathrm{d}F_\theta(t), \tag{4}$$

$k\in\mathbb{N}$. Define now a new df

$$\mathrm{d}G(t)=c^{-1}t^r\,\mathrm{d}F_\theta(t),$$

where $c=E(\theta^r)$. If Z is a rv with df G then (4) implies that

$$(1-m(k))E(U^k)=E(U^{k+1}),\quad k\in\mathbb{N},$$

where $U=1-Z$. From the above recurrence we have

$$E(U^k)=\prod_{j=0}^{k-1}(1-m(j)),\quad k\in\mathbb{N}$$

and the distributions of U and, consequently, Z are uniquely determined by the function m (all the moments exist and uniquely identify the distributions since U and Z have bounded supports). Hence G is also uniquely determined by m. Since $c=[E(Z^{-r})]^{-1}$ then, by the definition of G, also the df F_θ is unique. □

It is easy to see that the result also holds if instead of (1) we consider

$$\mu_{X|\theta}=\mathrm{nb}_1(r,1-\theta).$$

Then the mixture is a special case of the power series model, i.e. $\mu_{X|\theta}=\mathrm{PSD}(a,\theta)$, where a power series distribution, $\mathrm{PSD}(a,\theta)$, is defined by the pmf

$$p_\theta(k)=a(k)\theta^k/f(\theta),\quad k\in S\subset\mathbb{N},$$

where $a\geqslant 0$ is a coefficient function and $f>0$ is a series function ($f(\theta)=\sum_{k\in S}a(k)\theta^k$). Such a general problem in the case of bivariate discrete measures was considered in Wesołowski (1995b). Similar argument to that given in the proof above implies that $\mathrm{PSD}(a,\theta)$ mixtures are also identifiable:

Theorem 2. *Let* (X,θ) *be a mixture model with*

$$\mu_{X|\theta}=\mathrm{PSD}(a,\theta), \tag{5}$$

with $\mathrm{supp}(X)=\mathbb{N}$. *If* $\mathrm{supp}(\theta)$ *is not bounded additionally assume*

$$\sum_{k=0}^{\infty}\sqrt[2k]{a(k)}=\infty. \tag{6}$$

Then the prior distribution of θ *is uniquely determined by the posterior mean* $E(\theta|X)$.

Proof. Similarly as in the proof of Theorem 1 we have from (5) and the Bayes formula

$$m(k)E(Z^k)=E(Z^{k+1}), \quad k\in\mathbb{N}, \tag{7}$$

where Z is a rv with df G defined by

$$\mathrm{d}G(t)=c\,\mathrm{d}F_\theta(t)/f(t), \quad t\in \operatorname{supp}(\theta)=S_\theta,$$

$c^{-1}=\int_{S_\theta} 1/f(t)\,\mathrm{d}F_\theta(t)$. Observe that for unbounded S_θ

$$E(Z^k)=\int_{S_\theta} c\frac{t^k}{f(t)}\,\mathrm{d}F_\theta(t)=\frac{c}{a(k)}\int_{S_\theta}\frac{a(k)t^k}{f(t)}\,\mathrm{d}F_\theta(t)\leqslant\frac{c}{a(k)}$$

for any $k\in\mathbb{N}$. Hence Z is infinitely integrable and (7) is really valid. Also by (6)

$$\sum_{k=1}^{\infty}\frac{1}{\sqrt[2k]{E(Z^k)}}\geqslant\sum_{k=1}^{\infty}\sqrt[2k]{\frac{a(k)}{c}}=\infty$$

and by the Carleman criterion the distribution of Z is uniquely determined by the sequence of moments. Consequently G is uniquely determined by m. And also the distribution of θ is identifiable. □

Similar problem has been considered also in Sapatinas (1995), where instead of the condition on the coefficient function a a restriction on the regression function m of the following form

$$\sum_{x=1}^{\infty}\left(\prod_{i=0}^{x-1} m(i)\right)^{-1/2x}=\infty \tag{8}$$

was imposed and an approach via the classical identifiability method was proposed.

Assume that instead of the posterior mean $E(\theta|X)$ a form of the second posterior moment $E(\theta^2|X)$ is known. Then, similarly as in the proof of Theorem 1 (with its notation), we have

$$E(U^{k+2})=2E(U^{k+1})-(n(k)+1)E(U^k), \quad k\in\mathbb{N},$$

where $n(k)=E(\theta^2|X=k)$, $k\in\mathbb{N}$, which is not sufficient for determining the distribution of U since $E(U)$ remains unknown. This can be handled by additional assumption that $E(\theta|X=0)=m(0)$ is known. Consequently, we have

Theorem 3. *Assume that the nb mixture model (X,θ) is defined by the relation* (1). *Then the prior distribution of θ is uniquely determined by $E(\theta^2|X)$ and $E(\theta|X=0)$.*

Similarly, it follows easily that the mixture (1) is identifiable by a higher posterior moment $E(\theta^n|X)$ and the collection $E(\theta^i|X=0)$, $i=1,\dots,n-1$. It is possible also to consider $E(\theta^n|X)$ with some other combinations of $E(\theta^i|X=j)$, $i\in\{1,\dots,n-1\}$, $j\in\{0,1,\dots,n-1\}$. If we consider $\mu_{X|\theta}=\mathrm{nb}_1(r,1-\theta)$, the identifiability results also hold (even with much simpler recurrence relations). Also similar extensions towards $\mathrm{PSD}(a,\theta)$ mixtures can be obtained.

3. nb_2 mixtures

Here we consider the second type of a nb mixture. It was proved in Papageorgiou (1984) that such a mixture is identifiable by a regression function under assumptions that all the moments of marginals exist and uniquely determine their distributions. The method of the proof was based on the classical identifiability of the mixture by the marginal. We prove that the both type of restrictions are superfluous.

Theorem 4. *Assume that* (X,θ) *is a nb mixture model defined by*

$$\mu_{X|\theta} = \mathrm{nb}_2(r,\theta), \tag{9}$$

where $r>0$ *and* $\mathrm{supp}(\theta)\subset(0,\infty)$. *Then the prior distribution of* θ *is uniquely determined by the posterior mean* $E(\theta|X)$.

Proof. Similarly as in the proof of Theorem 1, via the Bayes formula and (9), we obtain the following equation for any $k\in\mathbb{N}$

$$m(k)\int_0^\infty \left(\frac{t}{1+t}\right)^k \frac{\mathrm{d}F_\theta(t)}{(1+t)^r} = \int_0^\infty t\left(\frac{t}{1+t}\right)^k \frac{\mathrm{d}F_\theta(t)}{(1+t)^r}, \tag{10}$$

where F_θ is a df of θ. Denote, now,

$$K(i,s) = \int_0^\infty \left(\frac{t}{1+t}\right)^i \frac{\mathrm{d}F_\theta(t)}{(1+t)^s},$$

for $i=0,1,\ldots,$ and $s=r-1,r$. Then (10) implies

$$(m(k)+1)K(k,r) = K(k,r-1), \quad k\in\mathbb{N}.$$

On the other hand, from the definition of the function K it follows that

$$K(k+1,r-1) = K(k,r-1) - K(k,r), \quad k\in\mathbb{N}.$$

Combining the above two equations we have the reccurence formula

$$m(k)K(k,r-1) = (m(k)+1)K(k+1,r-1), \quad k\in\mathbb{N}. \tag{11}$$

Define now a new df G by

$$\mathrm{d}G(t) = \frac{\mathrm{d}F_\theta(t)}{c(1+t)^{r-1}}, \quad t\in S_\theta,$$

where $c=K(0,r-1)$. Consequently, if Z is a rv with the df G then for $U=Z/(Z+1)$ we have

$$cE(U^k) = K(k,r-1), \quad k\in\mathbb{N}$$

and, hence (11) yields

$$E(U^{k+1}) = \frac{m(k)}{m(k)+1}E(U^k) = \prod_{j=0}^{k}\frac{m(j)}{m(j)+1},$$

$k\in\mathbb{N}$. Since U is bounded a.s. then its distribution is uniquely determined by the sequence of moments. Thus the function m determines the distribution of U, which gives the distribution of Z. And finally G is uniquely determined. Now $c^{-1}=E(1+Z)^{r-1}$ and, by the definition of G the df F_θ is also unique. $\square$

4. Applications

A rv X is said to have generalized hypergeometric type IV distribution ($\mathrm{GH_{IV}}(A,B,C)$) if its pmf has the form

$$P(X=k)=\frac{\binom{A}{k}\binom{B}{C-k}}{\binom{A+B}{C}}, \quad k=0,1,\ldots,$$

where A, B, C are real numbers fulfilling $A<0$, $C<0$, $0<A+B+1$. The distribution was introduced in Kemp and Kemp (1956), where it was derived by mixing negative binomial with beta distributions in two ways:

(1) If $\mu_{X|\theta}=\mathrm{nb}_1(r,\theta)$ and θ has the beta distribution of the first kind ($\mathrm{B_I}(a,b)$) with the density

$$f(t)=\frac{t^{a-1}(1-t)^{b-1}}{B(a,b)}I_{(0,1)}(t),$$

where $a>0$, $b>0$, then X is a $\mathrm{GH_{IV}}(-b,a+b-1,-r)$ rv.

(2) If $\mu_{X|\theta}=\mathrm{nb}_2(r,\theta)$ and θ has the beta distribution of the second kind ($\mathrm{B_{II}}(a,b)$) with the density

$$f(t)=\frac{t^{a-1}}{B(a,b)(1+t)^{a+b}}I_{(0,\infty)}(t),$$

where $a>0$, $b>0$, then X is a $\mathrm{GH_{IV}}(-a,a+b-1,-r)$ rv.

Since our previous results assure identifiability of nb mixtures by posterior moments then to characterize both the models it suffices to compute the respective quantities. Hence we easily get the following three results.

Proposition 1. *If* $\mu_{X|\theta}=\mathrm{nb}_1(r,\theta)$ *and*

$$E(\theta|X)=\frac{c}{d+X}, \quad 0<r<c<d,$$

then the prior distribution is $\mathrm{B_I}(c-r,d-c)$ *and* X *is a* $\mathrm{GH_{IV}}(c-d,d-r-1,-r)$ *rv.*

Proposition 2. *If* $\mu_{X|\theta}=\mathrm{nb}_1(r,\theta)$ *and*

$$E(\theta^2|X)=\frac{c(c+1)}{(d+X)(d+X+1)}, \qquad E(\theta|X=0)=c/d, \quad 0<r<c<d,$$

then the prior distribution is $\mathrm{B_I}(c-r,d-c)$ *and* X *is a* $\mathrm{GH_{IV}}(c-d,d-r-1,-r)$ *rv.*

Proposition 3. *If* $\mu_{X|\theta}=\mathrm{nb}_2(r,\theta)$ *and*

$$E(\theta|X)=cX+d, \quad 0<r<1/d, \ 0<c,$$

then the prior distribution is $\mathrm{B_{II}}(c/d,1/d-r)$ *and* X *is a* $\mathrm{GH_{IV}}(-c/d,(1+c)/d-r-1,-r)$ *rv.*

The results of Propositions 1 and 2 can be easily extended to negative binomial-generalized beta models, as proposed in Holla (1968), or to the Pascal $\mathrm{Pas}(r,\theta)$ models defined by the pmf

$$p(k)=\binom{k-1}{r-1}\theta^r(1-\theta)^{k-r}, \quad k=r,r+1,\ldots,$$

$0<\theta<1$, $r\in\{1,2,\dots\}$, which are just shifted negative binomial models. In particular the Pascal-beta model $\mu_{X|\theta}=\mathrm{Pas}(r,\theta)$ introduced in Dubey (1966) may be characterized by the posterior mean of the form

$$E(\theta|X)=\frac{a}{b+X},\quad 0<r<a<b+r.$$

Denote by $\mathscr{P}(\theta)$ the Poisson mixture model defined by the pmf

$$p(k)=\mathrm{e}^{-\theta}\frac{\theta^k}{k!},\quad k=0,1,\dots$$

In Wesołowski (1996), it was shown that $\mu_{X|\theta}=\mathscr{P}(ab^\theta)$, $a>0$, $0<b\leqslant 1$ and $E(\theta|X)=cb^X$ characterize the bivariate Poisson conditionals distribution (see Arnold and Strauss, 1991). Here we consider the Poisson mixture $\mu_{X|\theta}=\mathscr{P}(\lambda\theta)=\mathrm{PSD}(a,\theta)$ with $a(k)=\lambda^k/(k!)$ and $f(\theta)=e^{\lambda\theta}$, $\lambda>0$. Since

$$\sum_{k=0}^{\infty}\sqrt[2k]{a(k)}\geqslant\sqrt{\lambda}\sum_{k=1}^{\infty}\frac{1}{k}=\infty$$

then by Theorem 2, it follows that the Poisson mixture model is identifiable by the posterior mean $E(\theta|X)$. This result was originally obtained in Cacoullos and Papageorgiou (1982) (see also Cacoullos and Papageorgiou, 1983). Consequently for the Poisson mixture there is no need to check the condition (8) of Sapatinas (1995) to conclude about its identifiability.

The above observation can be used to identify by the posterior mean numerous Poisson mixture models known in the literature as for example: Poisson-gamma of Greenwood and Yule (1920), Poisson-uniform of Bhattacharya and Holla (1965), Poisson-power function of Rai (1971), Poisson–Lindley of Sankaran (1970), Poisson-generalized inverse Gaussian of Sichel (1975) or Poisson-beta of Holla and Bhattacharya (1965).

Observe that $\mathrm{nb}_1(r,1-\theta)$ is $\mathrm{PSD}(a,\theta)$ with

$$a(k)=\binom{r+k-1}{k}\quad\text{and}\quad f(\theta)=(1-\theta)^{-r}.$$

Since

$$\sum_{k=0}^{\infty}\sqrt[2k]{a(k)}>\sum_{k=1}^{\infty}\frac{1}{k}=\infty,$$

then Theorem 2 implies that any mixture of the form $\mu_{X|\theta}=\mathrm{nb}_1(r,1-\theta)$, $r>0$, is identifiable by the posterior mean. This can be also derived directly from Theorem 1. Consequently, the correct version of Corollary 3 of Sapatinas (1995), is

Proposition 4. *If* $\mu_{X|\theta}=\mathrm{nb}_1(r,1-\theta)$ *and*

$$E(\theta|X)=\frac{X+a}{X+a+b+r},\quad a,b,r>0,$$

then the prior distribution is $\mathrm{B_I}(a,b)$ *and* X *is a* $\mathrm{GH_{IV}}(-a,a+b-1,-r)$ *rv.*

Denote by $\mathrm{LSD}(\theta)$ the logarithmic series model defined by the pmf

$$p(k)=-\frac{\theta^k}{k\log(1-\theta)},\quad k=1,2,\dots,$$

where $\theta\in(0,1)$. Observe that $\mathrm{LSD}(\theta)=\mathrm{PSD}(a,\theta)$ with $a(k)=1/k$ and $f(\theta)=-\log(1-\theta)$. Since the condition (6) is obviously fulfilled then Theorem 2 implies that for the mixture $\mu_{X|\theta}=\mathrm{LSD}(\theta)$ the prior distribution of θ is uniquely determined by the posterior mean $E(\theta|X)$, which again allows to forget about checking the condition (8) of Sapatinas (1995) to conclude about its identifiability.

Acknowledgements

The authors would like to thank the referee for helpful suggestions which lead to improvements in the presentation of the material.

References

Arnold, B.C., Castillo, E., Sarabia, J.M., 1993. Conditionally specified models: structure and inference. In: Coudras, C.M., Rao, C.R. (Eds.), Multivariate Analysis: Future Directions 2. Elsevier, Amsterdam, pp. 441–450.

Arnold, B.C., Strauss, D.J., 1991. Bivariate distributions with conditionals in prescribed exponential families. J. Roy. Statist. Soc. B 53, 365–375.

Bhattacharya, S.K., Holla, M.S., 1965. On a discrete distribution with special reference to the theory of accident proneness. J. Amer. Statist. Assoc. 60, 1060–1066.

Boswell, M.T., Patil, G.P., 1970. Chance mechanisms generating the negative binomial distributions. In: Patil, G.P. (Ed.), Random Counts in Scientific Work, vol. 1: Random Counts in Models and Structures, The Pennsylvania State University Press, University Park, London, pp. 3–22.

Cacoullos, T., Papageorgiou, H., 1982. Characterizing the negative binomial distribution. J. Appl. Probab. 19, 742–743.

Cacoullos, T., Papageorgiou, H., 1983. Characterizations of discrete distributions by a conditional distribution and a regression function. Ann. Inst. Statist. Math. 35, 95–103.

Dubey, S.D., 1966. Compound Pascal distributions. Ann. Inst. Statist. Math. 18, 357–365.

Greenwood, M., Yule, G.U., 1920. An inquiry into the nature of frequency distribution representative of multiple happenings with particular reference to the outcome of multiple attacks of disease or of repeated accidents. J. Roy. Statist. Soc. 83, 255–279.

Gupta, A.K., Wesołowski, J., 1997. Uniform mixtures via posterior means. Ann. Inst. Statist. Math. 49, 171–180.

Holla, M.S., 1968. Discrete distributions with prior information. Ann. Inst. Statist. Math. 20, 151–157.

Holla, M.S., Bhattacharya, S.K., 1965. On discrete compound distribution. Ann. Inst. Statist. Math. 15, 377–384.

Kemp, C.D., Kemp, A.W., 1956. Generalized hypergeometric distributions. J. Roy. Statist. Soc. B 18, 202–211.

Korwar, R.M., 1975. On characterizing some discrete distributions by linear regression. Commun. Statist. 4, 1133–1147.

Kyriakoussis, A., Papageorgiou, H., 1991. Characterizations of logarithmic series distributions. Statistica Neerlandica 45, 1–8.

Papageorgiou, H., 1984. Characterizations of continuous binomial and negative binomial mixtures. Biometrical J. 26, 795–798.

Rai, G., 1971. A mathematical model of accident proneness. Trabajos de Estadistica 22, 207–212.

Sankaran, M., 1970. The discrete Poisson–Lindley distribution. Biometrics 26, 145–149.

Sapatinas, T., 1995. Identifiability of mixtures of power series distributions and related characterizations. Ann. Inst. Statist. Math. 47, 447–459.

Seshadri, V., Patil, G.P., 1964. A characterization of a bivariate distribution by the marginal and conditional distributions of the same component. Ann. Inst. Statist. Math. 15, 215–221.

Sichel, H.S., 1975. On a distribution law for word frequencies. J. Amer. Statist. Assoc. 70, 542–547.

Teicher, H., 1961. Identifiability of mixtures. Ann. Math. Statist. 32, 244–248.

Wesołowski, J., 1995a. Bivariate distributions via a Pareto conditional distribution and a regression function. Ann. Inst. Statist. Math. 47, 177–183.

Wesołowski, J., 1995b. Bivariate discrete measures via a power series conditional distribution and a regression function. J. Multivariate Anal. 55, 219–229.

Wesołowski, J., 1996. A new conditional specification of the bivariate Poisson conditionals distribution. Statistica Neerlandica 50, 390–393.