Brazilian Journal of Probability and Statistics
2014, Vol. 28, No. 3, 376–380
DOI: 10.1214/12-BJPS213

The spectral decomposition and inverse of multinomial and negative multinomial covariances

Christopher S. Withers[a] and Saralees Nadarajah[b]

[a]*Industrial Research Limited*
[b]*University of Manchester*

Abstract. We give the spectral decomposition and inverse of multinomial and negative multinomial covariances and related matrices.

1 Introduction

Multinomial and negative multinomial distributions are two of the most popular models for multivariate discrete data (Johnson et al., 1997). Their applications have been widespread. We mention: models to cluster Internet traffic (Jorgensen, 2004), funding source and research report quality in nutrition practice-related research, crash-prediction models for multilane roads, pollen counts, changepoints in the north Atlantic tropical cyclone record, magazine and Internet exposure, genome analysis (Chang and Wang, 2011), fish diet compositions from multiple data sources, statistical alarm method for mobile gamma spectrometry, stylometric analyses, clinical trials (Ganju and Zhou, 2011), impacts of movie reviews on box office, amount individuals withdraw at cash machines, soil microbial community, longline hook selectivity for red tilefish Branchiostegus japonicus in the East China Sea (Yamashita et al., 2009), gambling by auctions, automatic image annotation, and probabilities for the first division Spanish soccer league (Diaz-Emparanza and Nunez-Anton, 2010).

One of the most important properties of any multivariate distribution is the structure of the inverse of its covariance. It can be used to determine independence or complete dependence among variables. It can also be used to estimate standard errors, construct confidence intervals and construct tests of hypotheses.

The aim of this short note is to derive explicit expressions for the structure of the inverse covariance for a class of distributions containing the multinomial and negative multinomial distributions. An explicit expression for the inverse covariance of the latter distribution has not been known.

The contents of this short note are organized as follows. In Section 2, we give the spectral decomposition and inverse of the nonsingular matrix $V \in \mathbb{C}^{s\times s}$ with (i, j) element

$$V_{ij} = p_i \delta_{ij} + \alpha q_i q_j,$$

Key words and phrases. Multinomial distribution, negative multinomial distribution, spectral decomposition.

Received August 2012; accepted November 2012.

where δ_{ij} is the Kronecker delta function. In Section 3, this is applied to the multinomial and negative multinomial covariances.

2 Main results

Let $\alpha, p_1, \ldots, p_s$ be nonzero.

Theorem 2.1. *The $s \times s$ matrix $V = (V_{ij})$, where*

$$V_{ij} = p_i \delta_{ij} + \alpha q_i q_j, \tag{2.1}$$

has inverse $V^{-1} = (V^{ij})$, where

$$V^{ij} = p_i^{-1} \delta_{ij} - (\alpha^{-1} + \nu)^{-1} \tilde{q}_i \tilde{q}_j, \tag{2.2}$$

where

$$\tilde{q}_i = p_i^{-1} q_i, \qquad \nu = \sum_{k=1}^{s} p_k^{-1} q_k^2, \tag{2.3}$$

provided that $\alpha^{-1} + \nu \neq 0$.

Proof. Note that

$$\begin{aligned}
\sum_{k=1}^{s} V_{ik} V_{kj} &= \sum_{k=1}^{s} (p_i \delta_{ik} + \alpha q_i q_k)\big(p_k^{-1} \delta_{kj} - (\alpha^{-1} + \nu)^{-1} \tilde{q}_k \tilde{q}_j\big) \\
&= p_i \sum_{k=1}^{s} \delta_{ik} p_k^{-1} \delta_{kj} - (\alpha^{-1} + \nu)^{-1} p_i \tilde{q}_j \sum_{k=1}^{s} \delta_{ik} \tilde{q}_k + \alpha q_i \sum_{k=1}^{s} q_k p_k^{-1} \delta_{kj} \\
&\quad - \alpha (\alpha^{-1} + \nu)^{-1} q_i \tilde{q}_j \sum_{k=1}^{s} q_k \tilde{q}_k \\
&= p_i \sum_{k=1}^{s} \delta_{ik} p_k^{-1} \delta_{kj} - (\alpha^{-1} + \nu)^{-1} p_i \tilde{q}_j \tilde{q}_i + \alpha q_i q_j p_j^{-1} \\
&\quad - \alpha (\alpha^{-1} + \nu)^{-1} q_i \tilde{q}_j \sum_{k=1}^{s} p_k^{-1} q_k^2 \\
&= p_i \sum_{k=1}^{s} \delta_{ik} p_k^{-1} \delta_{kj}.
\end{aligned}$$

Hence, the result. □

A different approach to finding the inverse is to adapt the method of Watson (1996).

Theorem 2.2. *The $s \times s$ matrix V of (2.1) has eigenvalues $\{\lambda_j\}$ equal to the roots of*

$$\sum_{i=1}^{s} q_i^2 (p_i - \lambda)^{-1} = -\alpha^{-1}. \tag{2.4}$$

Also the eigenvector corresponding to the eigenvalue $\lambda = \lambda_j$ is $x = x_j$ with ith component

$$\gamma_j q_i (\lambda - p_i)^{-1},$$

where

$$\gamma_j = \left[\sum_{i=1}^{s} q_i^2 (\lambda - p_i)^{-2} \right]^{-1/2}.$$

So,

$$x_i' x_j = \delta_{ij}, \qquad V = \sum_{j=1}^{s} \lambda_j x_j x_j', \qquad V^{-1} = \sum_{j=1}^{s} \lambda_j^{-1} x_j x_j'.$$

Proof. Set $P = \operatorname{diag}(p_i)$, $A = P - \lambda I_s$, $u = (-\alpha)^{1/2} q$. So, $A - uu' = V - \lambda I_s$ and

$$\det(V - \lambda I_s) = \det\bigl(A - uu'\bigr) = \bigl(1 - u' A^{-1} u\bigr) \det(A)$$

by page 32 of Rao (1973). So, for λ an eigenvalue of V not equal to any p_i,

$$1 = u' A^{-1} u = \sum_{i=1}^{s} u_i^2 (p_i - \lambda)^{-1},$$

giving (2.4).

For $\lambda = \lambda_j$ and $x = x_j$,

$$0 = (Vx - \lambda x)_i = (p_i - \lambda)(x)_i + \alpha q_i c,$$

where $c = q'x$ and $(z)_i$ represents the ith component of a vector z. So, $(x)_i = c\alpha q_i (\lambda - p_i)^{-1}$, where c is given by $1 = x'x$. □

Note that we can have some eigenvalues of V equal to some of the p_i. For a simple example, consider the multinomial covariance, where all p_i's are equal, that is $p_1 = \cdots = p_s = q_1 = \cdots = q_s = p$ and $\alpha = -1$. In this case, $V = pI_s - p^2 zz'$, where z is the unit vector of ones. If a vector y is orthogonal to z, then, $Vy = py - p^2 zz'y = py$, that is, y is an eigenvector with eigenvalue p. Hence, the hyperplane that is orthogonal to z is an eigenspace. In this case, (2.4) reduces to a linear equation in λ since we have now only one eigenvalue that is not equal to p.

A more general example is to consider the case for which $q = (q_1, \ldots, q_s)$ is an eigenvector of $P = \mathrm{diag}(p_1, \ldots, p_s)$. Remembering that q, in this case, is a member of an orthogonal basis of eigenvectors of P and that the eigenvalues of P are $p_1, \ldots, p_s$, we see that any eigenvector of P that is orthogonal to q will be an eigenvector of V associated to one of the p_i's.

Because α is nonzero, we cannot have all eigenvalues equal to the p_i's unless all q_i's are zero (this follows since $\mathrm{trace}(V) = \sum_{i=1}^s p_i + \alpha \sum_{i=1}^s q_i^2 = \sum_{i=1}^s \lambda_i$). Hence, (2.4) is really useful: all roots of (2.4) are really eigenvalues of V although not necessarily the converse.

3 Examples

Example 3.1. For $N \sim \mathrm{Multinomial}_s(n, p)$ with $\sum_{i=1}^s p_i < 1$, $0 < p_i$, $i = 1, \ldots, s$, $\mathrm{covar}(N) = nV$, where V has the form (2.1) with $\alpha = -1$, $q_i = p_i$. So, its Fisher information is $nI(\theta)$, where by (2.2), $I(\theta) = V^{-1}$ is given by

$$V^{ij} = p_i^{-1}\delta_{ij} + (1-\nu)^{-1}$$

and ν is given by (2.3). This form for V^{-1} is given on page 215 of Mood (1950). By (2.4), the eigenvalues of V are the roots of

$$\sum_{i=1}^{s} p_i^2 (p_i - \lambda)^{-1} = 1. \tag{3.1}$$

For a given eigenvalue λ, the corresponding eigenvector has its ith element equal to

$$p_i(\lambda - p_i)^{-1}\left[\sum_{k=1}^{s} p_k^2 (p_k - \lambda)^{-2}\right]^{-1/2}.$$

If $s = 1$ then the root of (3.1) is $\lambda = p_1(1-p_1)$. If $s = 2$ then the roots of (3.1) are $\lambda = \{-p_1(p_1-1) - p_2(p_2-1) \pm (p_1^4 - 2p_1^3 + 2p_1^2p_2^2 + 2p_1^2p_2 + p_1^2 + 2p_1p_2^2 - 2p_1p_2 + p_2^4 - 2p_2^3 + p_2^2)^{(1/2)}\}/2$.

Example 3.2. For $N \sim \mathrm{NegativeMultinomial}_s(n, p)$ with $0 < p_i$, $i = 1, \ldots, s$, $\mathrm{covar}(N) = nV$, where V has the form (2.1) with $\alpha = 1$, $q_i = p_i$. So, its Fisher information is $nI(\theta)$, where by (2.2), $I(\theta) = V^{-1}$ is given by

$$V^{ij} = p_i^{-1}\delta_{ij} - (1+\nu)^{-1}$$

and ν is given by (2.3). By (2.4), the eigenvalues of V are the roots of

$$\sum_{i=1}^{s} p_i^2 (p_i - \lambda)^{-1} = -1. \tag{3.2}$$

For a given eigenvalue λ, the corresponding eigenvector has its ith element equal to

$$p_i(\lambda - p_i)^{-1}\left[\sum_{k=1}^{s} p_k^2(p_k-\lambda)^{-2}\right]^{-1/2}.$$

If $s=1$ then the root of (3.2) is $\lambda = p_1(1+p_1)$. If $s=2$ then the roots of (3.2) are $\lambda = \{p_1(1+p_1)+p_2(1+p_2) \pm (p_1^4+2p_1^3+2p_1^2p_2^2-2p_1^2p_2+p_1^2-2p_1p_2^2-2p_1p_2+p_2^4+2p_2^3+p_2^2)^{(1/2)}\}/2$. These results appear to be new.

Tanabe and Sagae (1992) gave the Moore–Penrose inverse to the full multinomial covariance, that is, when $\sum_{i=1}^{s} p_i = 1$. That covariance is singular, unlike the matrices we consider here. Watson (1996) gave the spectral decomposition of this singular covariance, providing an alternative method for obtaining its generalized inverse.

Acknowledgments

The authors would like to thank the Editor and the referee for careful reading and for their comments which greatly improved the paper.

References

Chang, G. and Wang, T. (2011). Genome analysis with the conditional multinomial distribution profile. *Journal of Theoretical Biology* **271**, 44–50. MR2974870

Diaz-Emparanza, I. and Nunez-Anton, V. (2010). On the use of simulation methods to compute probabilities: Application to the first division Spanish soccer league. *SORT* **34**, 181–199. MR2797478

Ganju, J. and Zhou, K. F. (2011). The benefit of stratification in clinical trials revisited. *Statistics in Medicine* **30**, 2881–2889. MR2844690

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. New York: Wiley. MR1429617

Jorgensen, M. (2004). Using multinomial mixture models to cluster Internet traffic. *Australian and New Zealand Journal of Statistics* **46**, 205–218. MR2076392

Mood, A. M. (1950). *Introduction to the Theory of Statistics*. New York: McGraw-Hill. MR0033470

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. New York: Wiley. MR0346957

Tanabe, K. and Sagae, M. (1992). An exact Cholesky decomposition and the generalized inverse of the variance–covariance matrix of the multinomial distribution, with applications. *Journal of the Royal Statistical Society, Ser. B* **54**, 211–219. MR1157720

Watson, G. S. (1996). Spectral decomposition of the covariance matrix of a multinomial. *Journal of the Royal Statistical Society, Ser. B* **58**, 289–291. MR1379243

Yamashita, H., Shiode, D. and Tokai, T. (2009). Longline hook selectivity for red tilefish Branchiostegus japonicus in the East China Sea. *Fisheries Science* **75**, 863–874.

Applied Mathematics Group
Industrial Research Limited
Lower Hutt
New Zealand

School of Mathematics
University of Manchester
Manchester M13 9PL
UK