J. R. Statist. Soc. B (1996)
58, *No.* 2, *pp.* 445–453

Monte Carlo Exact Conditional Tests for Log-linear and Logistic Models

By JONATHAN J. FORSTER, JOHN W. McDONALD and PETER W. F. SMITH†

University of Southampton, UK

[Received March 1994. Revised May 1995]

SUMMARY

The form of the exact conditional distribution of a sufficient statistic for the interest parameters, given a sufficient statistic for the nuisance parameters, is derived for a generalized linear model with canonical link. General results for log-linear and logistic models are given. A Gibbs sampling approach for generating from the conditional distribution is proposed, which enables Monte Carlo exact conditional tests to be performed. Examples include tests for goodness of fit of the all-two-way interaction model for a 2^8-table and of a simple logistic model. Tests against non-saturated alternatives are also considered.

Keywords: ESTIMATED *p*-VALUE; GENERALIZED LINEAR MODEL; GIBBS SAMPLER; LOGISTIC REGRESSION; LOG-LINEAR MODEL; MONTE CARLO EXACT CONDITIONAL TEST

1. INTRODUCTION

A common approach when testing a null hypothesis, in the presence of nuisance parameters, is to base the inference on the conditional distribution of a sufficient statistic for the interest parameters, given a sufficient statistic for the nuisance parameters. This results in an exact conditional test. Problems associated with this approach include obtaining the conditional distribution and calculating the *p*-value. These may be overcome by using an asymptotic approximation to the required distribution. However, for small sample sizes, sparse or unbalanced data, asymptotic results are often unreliable. An alternative solution is a Monte Carlo approach, where one generates a random sample from the required conditional distribution and estimates the *p*-value by using the generated empirical distribution of a test statistic (Agresti, 1992). Besag and Clifford (1989) advocated the use of Markov chain Monte Carlo (MCMC) methods. We extend their ideas to provide exact conditional tests for log-linear and logistic models, using Gibbs sampling.

2. EXACT CONDITIONAL DISTRIBUTIONS

We derive the form of the relevant conditional distributions for generalized linear models with canonical link. Following Firth (1991), section 3.3.2, if y_i, $i = 1, \ldots, n$, are independent, with each y_i from an exponential family distribution with parameters (θ_i, ϕ_i), the joint density is

†*Address for correspondence*: Department of Social Statistics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK.
E-mail: P.W.Smith@soton.ac.uk

 0035–9246/96/58445

$$f(y;\,\theta,\,\phi)=\exp\left[\sum_{i=1}^{n}\left\{\frac{y_i\theta_i-c(\theta_i)}{\phi_i}+h(y_i,\,\phi_i)\right\}\right], \tag{1}$$

where y, θ and ϕ are column vectors of length n. For the saturated regression model,

$$\theta_i=\theta_i(\beta)=x_i^{\mathrm{T}}\beta=\sum_{j=1}^{n}x_{ij}\beta_j, \qquad i=1,\,\ldots,\,n,$$

where $\beta=(\beta_1,\,\ldots,\,\beta_n)^{\mathrm{T}}$ is a vector of parameters corresponding to the vector of covariates $x_i^{\mathrm{T}}=(x_{i1},\,\ldots,\,x_{in})$, and T denotes the transpose. Then the joint density is

$$f(y;\,\beta,\,\phi)=\exp\left[\sum_{j=1}^{n}\beta_j\sum_{i=1}^{n}\frac{y_ix_{ij}}{\phi_i}-\sum_{i=1}^{n}\frac{c\{\theta_i(\beta)\}}{\phi_i}+\sum_{i=1}^{n}h(y_i,\,\phi_i)\right]. \tag{2}$$

If ϕ is known, then $\{z_j=\Sigma_{i=1}^{n}\,y_ix_{ij}/\phi_i;\,j=1,\,\ldots,\,n\}$ is a set of minimal sufficient statistics for $\{\beta_j;\,j=1,\,\ldots,\,n\}$. Note that $z=X^{\mathrm{T}}Wy$, where $X=(x_{ij})$ and W is an $n\times n$ diagonal matrix with elements $1/\phi_i$. Here the model matrix X is invertible and the density of the sufficient statistic is

$$f(z;\,\beta,\,\phi)\propto\exp\left[\sum_{j=1}^{n}z_j\beta_j+\sum_{i=1}^{n}h\{y_i(z),\,\phi_i\}\right], \tag{3}$$

where $y(z)=W^{-1}X^{-\mathrm{T}}z$, and $-\mathrm{T}$ denotes the inverse of the transposed matrix. The constant of proportionality does not depend on z.

Consider a reduced model defined by constraining r components of β. We partition β into β_R and $\beta_{\backslash R}$, where R and $\backslash R$ contain the indices of the restricted and unrestricted (nuisance) parameters respectively. The r-dimensional conditional distribution of z_R, the sufficient statistic for β_R, given $z_{\backslash R}$, the sufficient statistic for $\beta_{\backslash R}$, is

$$f(z_R|z_{\backslash R};\,\beta_R,\,\phi)\propto\exp\left[\sum_{j\in R}z_j\beta_j+\sum_{i=1}^{n}h\{y_i(z),\,\phi_i\}\right], \tag{4}$$

where the constant of proportionality does not depend on z_R. When $\beta_R=0$,

$$f(z_R|z_{\backslash R};\,\phi)\propto\exp\left[\sum_{i=1}^{n}h\{y_i(z),\,\phi_i\}\right]=\prod_{i=1}^{n}H\left(\phi_i\sum_{j=1}^{n}x^{ji}z_j,\,\phi_i\right), \tag{5}$$

where $H(\;,\;)=\exp\,h(\;,\;)$ and $X^{-\mathrm{T}}=(x^{ji})$.

This result is given in a similar form by Andersen (1980), p. 82; see also Lehmann (1986), p. 58, and Barndorff-Nielsen (1978), section 8.2. In most applications of exponential family theory $H\{y_i(z),\,\phi_i\}$ is dismissed as a constant of proportionality in the likelihood. However, for exact conditional tests $H\{y_i(z),\,\phi_i\}$ is fundamental.

Exact tests of goodness of fit of the reduced model, corresponding to H_0: $\beta_R=0$ against the saturated alternative, use the conditional distribution given by expression

(5). Tests of H_0 against a non-saturated alternative H_A: $\beta_A = 0$, where $A \subset R$, are based on the $(r-a)$-dimensional marginal distribution

$$f(z_{R\setminus A}|z_{\setminus R};\ \phi) \propto \int \prod_{i=1}^{n} H\left(\phi_i \sum_{j=1}^{n} x^{ji} z_j,\ \phi_i\right) \mathrm{d}z_A \tag{6}$$

obtained by integrating expression (5) over the sample space of each of the a components of z_A.

3. LOG-LINEAR MODEL

Consider a saturated log-linear model, where y_i are Poisson or (product) multinomial. Here, $H(y_i, \phi_i) = 1/y_i!$ and

$$\theta_i = \log E(Y_i) = \sum_{j=1}^{n} x_{ij}\beta_j, \qquad i = 1, \ldots, n.$$

The conditional distribution for testing H_0: $\beta_R = 0$ against the saturated alternative is, from expression (5),

$$f(z_R|z_{\setminus R}) \propto \left\{ \prod_{i=1}^{n} \left(\sum_{j=1}^{n} x^{ji} z_j \right)! \right\}^{-1}. \tag{7}$$

When testing for independence in a two-way table this distribution is a multi-dimensional hypergeometric distribution, so calculation or Monte Carlo estimation of the exact p-value is possible (Agresti, 1992). However, for arbitrary X and R, neither the normalizing constant nor the support of the distribution is available in closed form, and generating directly from this distribution is infeasible.

The exact conditional distribution for testing H_0 against the non-saturated alternative H_A is obtained by summing expression (7) over the sample space of each component of z_A.

4. GIBBS SAMPLING APPROACH

Gibbs sampling is an MCMC method which can be used to generate a realization from a multivariate distribution of interest when direct methods are not available (Smith and Roberts, 1993). Besag and Clifford (1989) proposed the use of MCMC methods to perform significance tests. We use Gibbs sampling to obtain realizations from the r-dimensional conditional distribution, given by expression (5), to perform Monte Carlo exact conditional tests. This involves sampling iteratively from univariate conditional distributions. Kolassa and Tanner (1994) implemented a Gibbs sampler for approximate conditional inference in exponential families. They used a double saddlepoint approximation to the cumulative distribution function for sampling from the univariate conditional distributions. For log-linear and logistic models, we sample from the univariate conditional distributions exactly.

For the log-linear model, the univariate distribution of a single component z_k of z_R, conditional on the other components of z, is, from expression (7),

$$f(z_k|z_{\backslash k}) \propto \left\{ \prod_{i=1}^{n} (c_i^k + x^{ki} z_k)! \right\}^{-1}, \qquad c_i^k = \sum_{j \neq k} x^{ji} z_j. \tag{8}$$

It is reasonably easy to find the support and to obtain the normalizing constant for these univariate conditional distributions by complete enumeration. Often, many of the x^{ki}-terms in expression (8) are 0, which results in a reduction from n to a much smaller number of factorial terms.

We estimate the exact conditional p-value by ranking the observed value of a test statistic among a random sample of values generated from the exact conditional distribution of the test statistic. The values generated are not independent and this must be taken into account when assessing the accuracy of the estimated p-value. Raftery and Lewis (1992) described a method for estimating the number of iterations of a Gibbs sampler required to evaluate a quantile to some prespecified accuracy. Their method can easily be modified to give approximate confidence intervals for an estimated p-value.

Monte Carlo tests against non-saturated alternatives are straightforward. The exact conditional test of H_0: $\beta_R = 0$ against H_A: $\beta_A = 0$ is based on the marginal distribution (6). For log-linear models, a sample from this distribution can be obtained directly by extracting $z_{R\backslash A}$ from a sample of z_R generated under H_0 as described above.

4.1. *Example: Log-linear Interaction Model*

For multiway tables, when the columns of X represent dummy variables for main effects and interactions, the model matrix can be written as a Kronecker product. The resulting model has been called a log-linear interaction model (Knuiman and Speed, 1988). The inverse of a Kronecker product of non-singular matrices is the Kronecker product of the corresponding inverse matrices. This facilitates the calculation of $X^{-\mathrm{T}}$.

For a 2^d-table, where the parameters corresponding to the second level are constrained to be 0 and the first component of β is the d-way interaction,

$$X = \bigotimes_{i=1}^{d} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad X^{-\mathrm{T}} = \bigotimes_{i=1}^{d} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

The exact conditional distribution for tests of goodness of fit can now be derived by substituting the elements of $X^{-\mathrm{T}}$ into expression (7). For example, for a test of no three-way interaction in a 2^3-table,

$$f(z_1|z_{\backslash 1}) \propto \{z_1!\,(z_2 - z_1)!\,(z_3 - z_1)!\,(z_4 - z_3 - z_2 + z_1)!\,(z_5 - z_1)!\,(z_6 - z_5 - z_2 + z_1)!\,(z_7 - z_5 - z_3 + z_1)!\,(z_8 - z_7 - z_6 + z_5 - z_4 + z_3 + z_2 - z_1)!\}^{-1}. \tag{9}$$

Tests for conditional independence and independence lead to multivariate conditional distributions. Generation from these distributions is possible by Gibbs sampling, although more direct approaches exist (Agresti, 1992). However, for $d \geqslant 4$ there are many models for which direct generation is infeasible.

We present two numerical examples: a 2^4-table where the exact results, obtained by complete enumeration, are available for comparison and a 2^8-table where the MCMC method provides the only feasible method for estimating the p-value. For each estimate we present an approximate 99% confidence interval, based on the method of Raftery and Lewis (1992).

Morgan and Blumenstein (1991) presented a 2^4-table with variables denoted by H, I, R and V and calculated the p-value for a test of the reduced model $HRV + HI + IR + IV$ against the alternative $HRV + HIR + IV$, using the probability of $z_{R\backslash A}$ as a test statistic. We generated a sample of 100000 realizations from the exact conditional distribution of z_R and extracted the sample of $z_{R\backslash A}$, the univariate sufficient statistic corresponding to the HIR interaction. The p-value, estimated by the generated empirical tail probabilities of $z_{R\backslash A}$, was 0.029. The approximate 99% confidence interval (0.026, 0.032) contains the exact p-value, 0.030. Morgan and Blumenstein (1991) also compared the fit of the model $HRV + HIR$ against the same alternative. The approximate 99% confidence interval (0.119, 0.131) for the exact two-sided p-value obtained from a further run of the Gibbs sampler was again in good agreement with the exact value, 0.126.

Whittaker (1990), p. 280, presented a 2^8-table with 665 observations and considered the all-two-way interaction model. The data are sparse, 165 of the 256 cells are empty and asymptotic goodness-of-fit tests are unreliable. We performed two exact conditional tests for this model by using the Gibbs sampler. Here, R contains the indices for the three- and higher way interactions and $r = 219$, the residual degrees of freedom for the all-two-way interaction model. We generated a sample of 100000 realizations of z_R from its exact conditional distribution. As the conditional probabilities of z_R up to a constant of proportionality are given by expression (7), they can be used as a test statistic for goodness of fit. The estimated p-value, using these probabilities as a test statistic, is 0.156 ± 0.005. Alternatively, the estimated p-value, using the likelihood ratio (LR) statistic, is 0.116 ± 0.004. Note that the observed LR statistic, 144.56, is in the right-hand tail of its exact distribution, but in the left-hand tail of its asymptotic distribution, χ^2_{219}.

For this table, Whittaker (1990) selected a reduced model with a further 14 parameters set to 0. Here, R also contains the indices for these 14 parameters. We performed an exact test of this reduced model against the all-two-way interaction model, using a sample from the distribution of $z_{R\backslash A}$, conditional on $z_{\backslash R}$, where A contains the indices for the three- and higher way interactions. The conditional probabilities of $z_{R\backslash A}$, given in the general case by expression (6), are not explicitly available, even up to a constant of proportionality. If $r - a$ is small, then these probabilities can be estimated by Monte Carlo integration, as in the 2^4-example where $r - a = 1$. However, in the current example, $r - a = 14$ and this is not feasible. A one-dimensional summary, such as the LR statistic, is necessary. For the LR statistic, the estimated exact p-value is 0.315 ± 0.014, based on a sample of 100000. As expected, the asymptotic p-value, 0.303, is much closer to the exact value than for the goodness-of-fit test.

5. LOGISTIC REGRESSION

The Gibbs sampler can also be used to perform Monte Carlo exact conditional tests for logistic regression models where $Y_i \sim \text{binomial}(m_i, \pi_i)$. Here, $\theta_i = \text{logit}(\pi_i) = x_i^T \beta$, and

$$H(y_i, \phi_i) \propto \binom{m_i}{y_i}, \qquad i = 1, \ldots, n. \tag{10}$$

To test hypotheses concerning a model with $p < n$ covariates, we construct the saturated model

$$\theta = \text{logit}(\pi) = X\beta = \begin{pmatrix} I & X_1 \\ \mathbf{0} & X_2 \end{pmatrix} \begin{pmatrix} \beta_R \\ \beta_{\backslash R} \end{pmatrix}. \tag{11}$$

Here $\beta_{\backslash R}$ is the vector of parameters corresponding to the p covariates (the nuisance parameters) and β_R is the vector of r additional parameters (set to 0 under H_0) which extends the model of interest to a saturated model. We assume, without loss of generality, that $R = \{1, \ldots, r\}$ and construct X as follows: the $n \times p$ matrix of covariates for the model of interest is partitioned into $(X_1, X_2)^{\mathrm{T}}$, where X_1 is an $r \times p$ matrix, X_2 is a $p \times p$ matrix, I is the $r \times r$ identity matrix and $\mathbf{0}$ is a $p \times r$ matrix of 0s.

The appropriate distribution for testing goodness of fit is conditional on the sufficient statistics for the regression parameters (McCullagh, 1986). In our framework, a goodness-of-fit test for the model of interest corresponds to the exact conditional test of H_0: $\beta_R = 0$. Here

$$X^{-\mathrm{T}} = \begin{pmatrix} I & \mathbf{0} \\ -X_2^{-\mathrm{T}} X_1^{\mathrm{T}} & X_2^{-\mathrm{T}} \end{pmatrix} \tag{12}$$

and the required exact conditional distribution is, from expressions (5) and (10),

$$f(z_R | z_{\backslash R}) \propto \prod_{i=1}^{n} \binom{m_i}{\sum_{j=1}^{n} x^{ji} z_j}. \tag{13}$$

When $m_i = 1$ for all i, this conditional distribution is uniform, since all the binomial coefficients equal 1. Hence, goodness-of-fit tests for pure binary data, based on this conditional distribution, are not sensible. For continuous covariates the conditional distribution of z_R is often degenerate as, conditional on $z_{\backslash R}$, only the observed z results in integer values of $\Sigma_{j=1}^{n} x^{ji} z_j$ within the required range $[0, m_i]$ for all i. However, when the covariate values are integer or evenly spaced, the exact conditional distribution is not usually degenerate.

In general, evaluation of the normalizing constant in expression (13) is infeasible, and direct generation from this r-dimensional distribution is not practical. However, we can generate from this distribution by using a Gibbs sampler by generating from the first r univariate conditionals,

$$f(z_k | z_{\backslash k}) \propto \prod_{i=1}^{n} \binom{m_i}{c_i^k + x^{ki} z_k}, \qquad c_i^k = \sum_{j \neq k} x^{ji} z_j. \tag{14}$$

For $i = 1, \ldots, r$, the ith row of $X^{-\mathrm{T}}$ is e_i^{T}, a vector of 0s except for a 1 in the ith position, so, for all k, $c_i^k + x^{ki} z_k = e_i^{\mathrm{T}} z = z_i$. Hence

$$f(z_k|z_{\backslash k}) \propto \prod_{i=1}^{r} \binom{m_i}{z_i} \prod_{i=r+1}^{n} \binom{m_i}{c_i^k + x^{ki} z_k}$$
$$\propto \begin{cases} \binom{m_k}{z_k} \prod_{i=r+1}^{n} \binom{m_i}{c_i^k + x^{ki} z_k} & \text{if } k = 1, \ldots, r, \\ \prod_{i=r+1}^{n} \binom{m_i}{c_i^k + x^{ki} z_k} & \text{if } k = r+1, \ldots, n \end{cases} \tag{15}$$

since the leading terms not involving z_k can be absorbed into the constant of proportionality. This reduction from n to at most $p+1$ terms simplifies the implementation of the Gibbs sampler.

Monte Carlo tests against non-saturated alternatives are again straightforward. The exact conditional test of H_0: $\beta_R = 0$ against H_A: $\beta_A = 0$ is based on the marginal distribution (6). For logistic regression models, a sample from this distribution can be obtained by extraction if the parameters $\beta_{R\backslash A}$ are included initially in the extension to the saturated model.

5.1. *Example: Simple Linear Logistic Regression*

Consider the model $\text{logit}(\pi_i) = \beta_1 + \beta_2 x_i$, $i = 1, \ldots, n$. Here $p = 2$, $r = n - 2$,

$$X_1 = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_{n-2} \end{pmatrix}, \quad X_2 = \begin{pmatrix} 1 & x_{n-1} \\ 1 & x_n \end{pmatrix}$$

and from equation (12)

$$X^{-\mathrm{T}} = \frac{1}{x_n - x_{n-1}} \begin{pmatrix} & (x_n - x_{n-1}) I & & \begin{matrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{matrix} \\ -(x_n - x_1) & \ldots & -(x_n - x_{n-2}) & x_n & -1 \\ x_{n-1} - x_1 & \ldots & x_{n-1} - x_{n-2} & -x_{n-1} & 1 \end{pmatrix}. \tag{16}$$

The exact conditional distribution for testing H_0: $\beta_R = 0$ is r dimensional and is given by expression (13). The first r univariate conditionals required for Gibbs sampling are, from expressions (15) and (16),

$$f(z_k|z_{\backslash k}) \propto \binom{m_k}{z_k} \binom{m_{n-1}}{c_{n-1}^k - \frac{x_n - x_k}{x_n - x_{n-1}} z_k} \binom{m_n}{c_n^k + \frac{x_{n-1} - x_k}{x_n - x_{n-1}} z_k}.$$

As a numerical example, we consider the dose–response data presented by Collett (1991), p. 75, in which 20 moths of each sex were exposed to one of six doses of cypermethrin, equally spaced on the $\log_2$-scale, and the number adversely affected was recorded. We tested the goodness of fit of the logistic model with $\log_2(\text{dose})$ as a covariate, but without a sex effect. For this model the goodness-of-fit LR statistic is 16.98 on 10 degrees of freedom with an asymptotic p-value of 0.075. The estimated Monte Carlo exact p-value, based on a sample of 100000 realizations, is 0.113 ± 0.006 with 99% confidence. Although the asymptotic result casts some doubt on the model without a sex effect, the estimated exact p-value indicates a more reasonable fit. The exact conditional distribution is far from degenerate as, in 100000 iterations, 56771 distinct values of the LR statistic were generated.

6. DISCUSSION

We have presented the theory for obtaining the form of the required conditional distribution for tests of nested hypotheses in log-linear and logistic regression models. We used a Gibbs sampling approach for generating from this conditional distribution to estimate an exact p-value. The implementation of this procedure does not depend on the choice of a test statistic, although the p-values may differ.

Diaconis and Sturmfels (1993) used a Markov chain algorithm for sampling from discrete exponential families conditionally on a sufficient statistic. They gave examples where certain chains are not connected, but they implicitly overcame lack of connectivity by reparameterization and overparameterization. Initial research suggests that a lack of connectivity occurs when some of the univariate conditional distributions are degenerate. Our implementation of the Gibbs sampler includes checks for this. In our experience, degeneracy may be overcome by reparameterization or by sampling from bivariate or trivariate conditional distributions. Useful reparameterizations include permuting parameters and saturating the model in a different way.

How best to implement the Gibbs sampler is still an area of much current research and vigorous debate (see Smith and Roberts (1993), Gelman and Rubin (1992) and the accompanying discussions). However, we have not encountered any numerical problems in generation from discrete univariate conditionals, or evidence of lack of convergence of the Gibbs sampler.

ACKNOWLEDGEMENT

The work of the last two authors was supported by Economic and Social Research Council award H519255005 as part of the Analysis of Large and Complex Datasets Programme.

REFERENCES

Agresti, A. (1992) A survey of exact inference for contingency tables (with discussion). *Statist. Sci.*, **7**, 131–177.

Andersen, E. B. (1980) *Discrete Statistical Models with Social Science Applications*. Amsterdam: North-Holland.

Barndorff-Nielsen, O. (1978) *Information in Exponential Families*. Chichester: Wiley.

Besag, J. and Clifford, P. (1989) Generalized Monte Carlo significance tests. *Biometrika*, **76**, 633–642.

Collett, D. (1991) *Modelling Binary Data*. London: Chapman and Hall.

Diaconis, P. and Sturmfels, B. (1993) Algebraic algorithms for sampling from conditional distributions. *Technical Report*. Department of Mathematics, Harvard University, Cambridge.

Firth, D. (1991) Generalized linear models. In *Statistical Theory and Modelling* (eds D. V. Hinkley, N. Reid and E. J. Snell), pp. 55–82. London: Chapman and Hall.

Gelman, A. and Rubin, D. B. (1992) Inference from iterative simulation using multiple sequences (with discussion). *Statist. Sci.*, **7**, 457–472.

Knuiman, M. W. and Speed, T. P. (1988) Incorporating prior information into the analysis of contingency tables. *Biometrics*, **44**, 1061–1071.

Kolassa, J. E. and Tanner, M. A. (1994) Approximate conditional inference in exponential families via the Gibbs sampler. *J. Am. Statist. Ass.*, **89**, 697–702.

Lehmann, E. L. (1986) *Testing Statistical Hypotheses*, 2nd edn. New York: Wiley.

McCullagh, P. (1986) The conditional distribution of the goodness-of-fit statistic for discrete data. *J. Am. Statist. Ass.*, **81**, 104–107.

Morgan, W. M. and Blumenstein, B. A. (1991) Exact conditional tests for hierarchical models in multidimensional contingency tables. *Appl. Statist.*, **40**, 435–442.

Raftery, A. E. and Lewis, S. (1992) How many iterations in the Gibbs sampler? In *Bayesian Statistics 4* (eds J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith), pp. 777–784. Oxford: Oxford University Press.

Smith, A. F. M. and Roberts, G. O. (1993) Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *J. R. Statist. Soc.* B, **55**, 3–23.

Whittaker, J. (1990) *Graphical Models in Applied Multivariate Statistics*. Chichester: Wiley.