Appl. Statist. (1994)
43, *No.* 1, *pp.* 201–221

Bayesian Analysis of Linear and Non-linear Population Models by using the Gibbs Sampler

By J. C. WAKEFIELD and A. F. M. SMITH†,

Imperial College of Science, Technology and Medicine, London, UK

A. RACINE-POON

Ciba–Geigy AG, Basle, Switzerland

and A. E. GELFAND

University of Connecticut, Storrs, USA

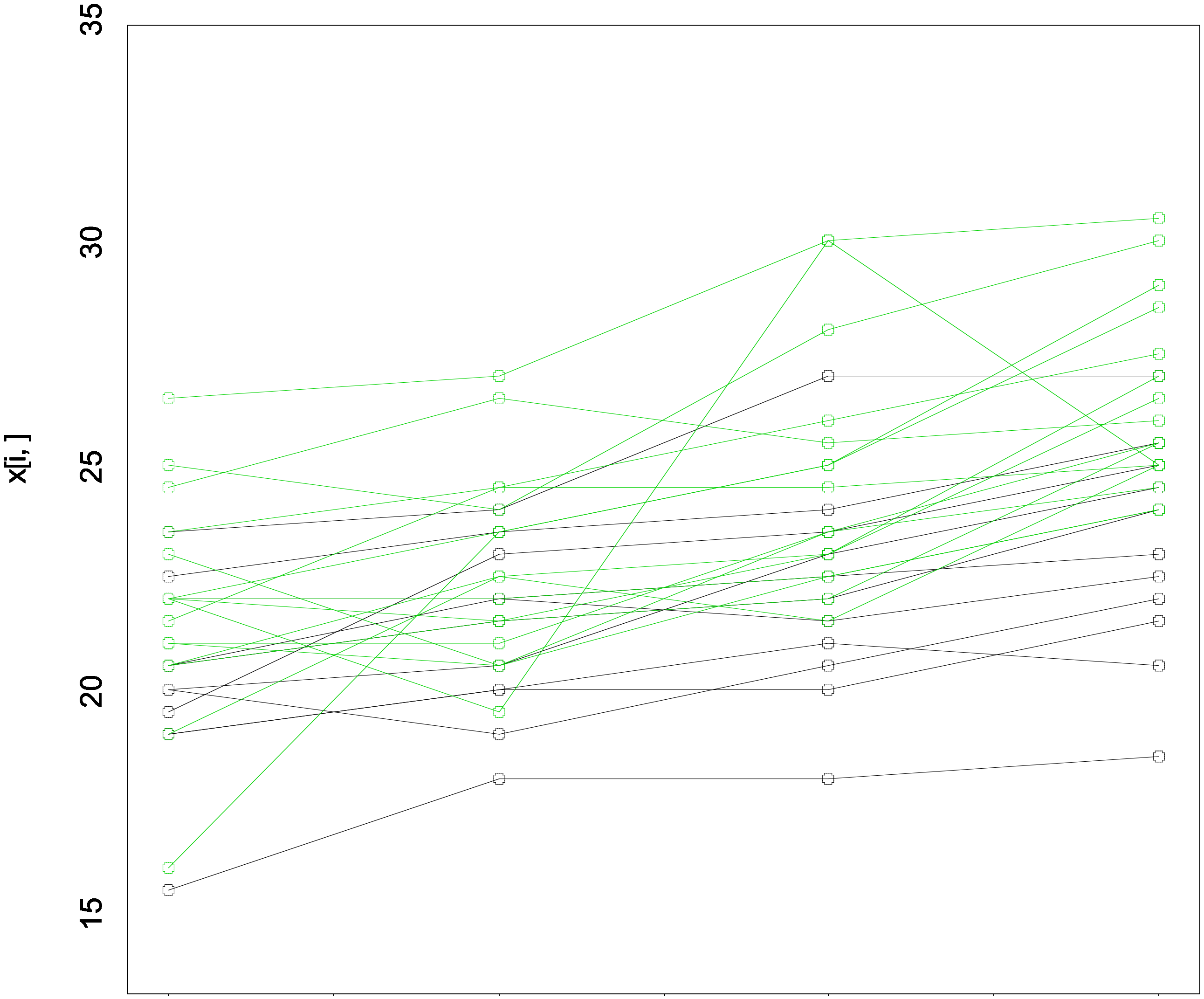
2.1. *Linear Population Biological Growth Example*

Table 1 records dental measurements of the distance (in millimetres) from the centre of the pituitary gland to the pteryo-maxillary fissure in 11 girls and 16 boys at the ages of 8, 10, 12 and 14 years. Both in the original analysis (Potthoff and

TABLE 1

Measurements on 11 girls and 16 boys, at four different ages

Individual	*Results for the following ages in years:* 8	10	12	14
Girls				
1	21	20	21.5	23
2	21	21.5	24	25.5
3	20.5	24	24.5	26
4	23.5	24.5	25	26.5
5	21.5	23	22.5	23.5
6	20	21	21	22.5
7	21.5	22.5	23	25
8	23	23	23.5	24
9	20	21	22	21.5
10	16.5	19	19	19.5
11	24.5	25	28	28
Boys				
12	26	25	29	31
13	21.5	22.5	23	26.5
14	23	22.5	24	27.5
15	25.5	27.5	26.5	27
16	20	23.5	22.5	26
17	24.5	25.5	27	28.5
18	22	22	24.5	26.5
19	24	21.5	24.5	25.5
20	23	20.5	31	26
21	27.5	28	31	31.5
22	23	23	23.5	25
23	21.5	23.5	24	28
24	17	24.5	26	29.5
25	22.5	25.5	25.5	26
26	23	24.5	26	30
27	22	21.5	23.5	25


x[i,]
15
20
25
30
35

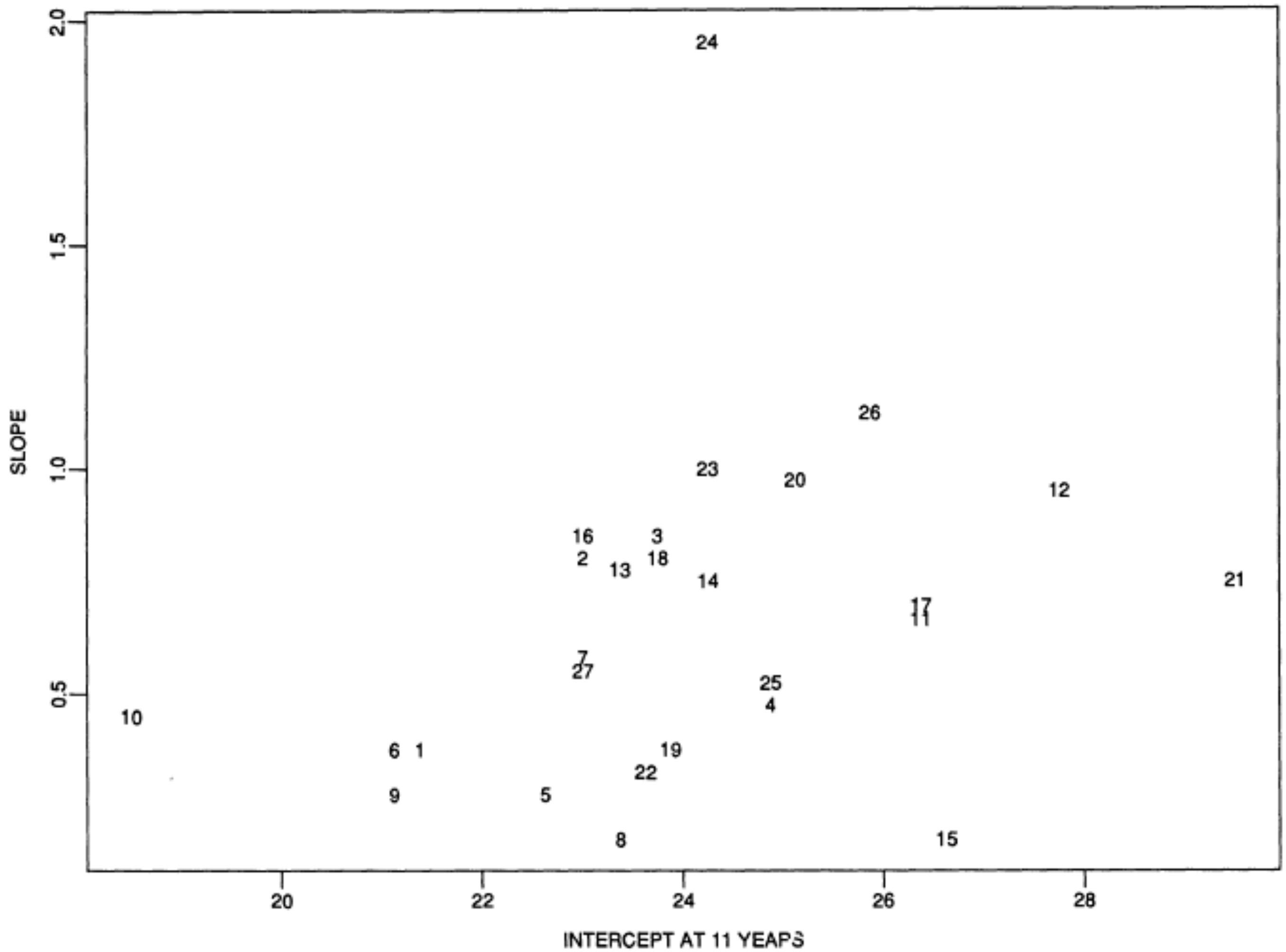


Fig. 1. Least squares estimates of intercepts and slopes (the plotting symbol is the individual's number)

Let y_{ij} denote the jth measurement for ith child

Let x_{ij} denote the time of the jth measurement for ith child

Consider a model with this likelihood:

$$\prod_i \prod_j [y_{ij} \mid \theta_i, \tau] = \prod_i \prod_j [y_{ij} \mid \alpha_i, \beta_i, \tau]$$

$$= \prod_i \prod_j N(y_{ij} \mid \alpha_i + \beta_i x_{ij}, \tau^{-1})$$

Bivariate normal prior for the θ_i's:

$$\prod_i [\boldsymbol{\theta}_i | \boldsymbol{\phi}] = \prod_i N(\boldsymbol{\theta}_i | \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

with $\boldsymbol{\phi} = (\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2)$, $E(\alpha_i) = \mu_1$ and $E(\beta_i) = \mu_2$, so that the individual straight line growth curves are, in effect, regarded as distributed around a 'mean' population growth curve, $\mu_1 + \mu_2 x$ with population variation described by the 2×2 covariance matrix $\boldsymbol{\Sigma}$.

$$[\tau] = G(\tau | \tfrac{1}{2}\nu_0, \tfrac{1}{2}\nu_0\tau_0)$$

Finally, suppose that the prior specification is completed by assuming ν_0 and τ_0 to be known and taking the third stage of the hierarchy to have the form

$$[\boldsymbol{\phi}] = [\boldsymbol{\mu}][\boldsymbol{\Sigma}^{-1}] = N(\boldsymbol{\mu} | \boldsymbol{\eta}, \mathbf{C})\, W\{\boldsymbol{\Sigma}^{-1} | (\rho\mathbf{R})^{-1}, \rho\},$$

Posterior:

$$[\theta_1, \ldots, \theta_{27}, \tau, \mu, \Sigma^{-1}|\mathbf{y}] \propto \prod_i \prod_j N(y_{ij}|\alpha_i, \beta_i, \tau) \times \prod_i N(\theta_i|\mu, \Sigma) G(\tau) N(\mu|\eta, C) W(\Sigma^{-1}|R^{-1}, \rho)$$

Full Conditionals:

$$[\boldsymbol{\theta}_i|\mathbf{y},\ \boldsymbol{\mu},\ \boldsymbol{\Sigma}^{-1},\ \tau,\ \boldsymbol{\theta}_j,\ j\neq i] = N\{\boldsymbol{\theta}_i|\mathbf{D}_i(\tau\mathbf{X}_i^{\mathrm{T}}y_i+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}),\ \mathbf{D}_i\}, \qquad i=1,\ \ldots,\ I,$$

$$[\boldsymbol{\mu}|\mathbf{y},\ \boldsymbol{\theta},\ \boldsymbol{\Sigma}^{-1},\ \tau] = N\{\boldsymbol{\mu}|\mathbf{V}(I\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{\theta}}+\mathbf{C}^{-1}\boldsymbol{\eta}),\ \mathbf{V}\},$$

$$[\boldsymbol{\Sigma}^{-1}|\mathbf{y},\ \boldsymbol{\theta},\ \boldsymbol{\mu},\ \tau] = W\left[\boldsymbol{\Sigma}^{-1}\middle|\left\{\sum_i(\boldsymbol{\theta}_i-\boldsymbol{\mu})(\boldsymbol{\theta}_i-\boldsymbol{\mu})^{\mathrm{T}}+\rho\mathbf{R}\right\}^{-1},\ I+\rho\right],$$

$$[\tau|\mathbf{y},\ \boldsymbol{\theta},\ \boldsymbol{\mu},\ \boldsymbol{\Sigma}^{-1}] = G\left[\tau\middle|\tfrac{1}{2}(\nu_0+n),\ \tfrac{1}{2}\left\{\sum_i(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\theta}_i)^{\mathrm{T}}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\theta}_i)+\nu_0\tau_0\right\}\right].$$

$$\bar{\boldsymbol{\theta}} = I^{-1}\Sigma_i\boldsymbol{\theta}_i,\ \mathbf{D}_i^{-1} = \tau\mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i + \boldsymbol{\Sigma}^{-1},\ \ \mathbf{V}^{-1} = I\boldsymbol{\Sigma}^{-1} + \mathbf{C}^{-1}$$

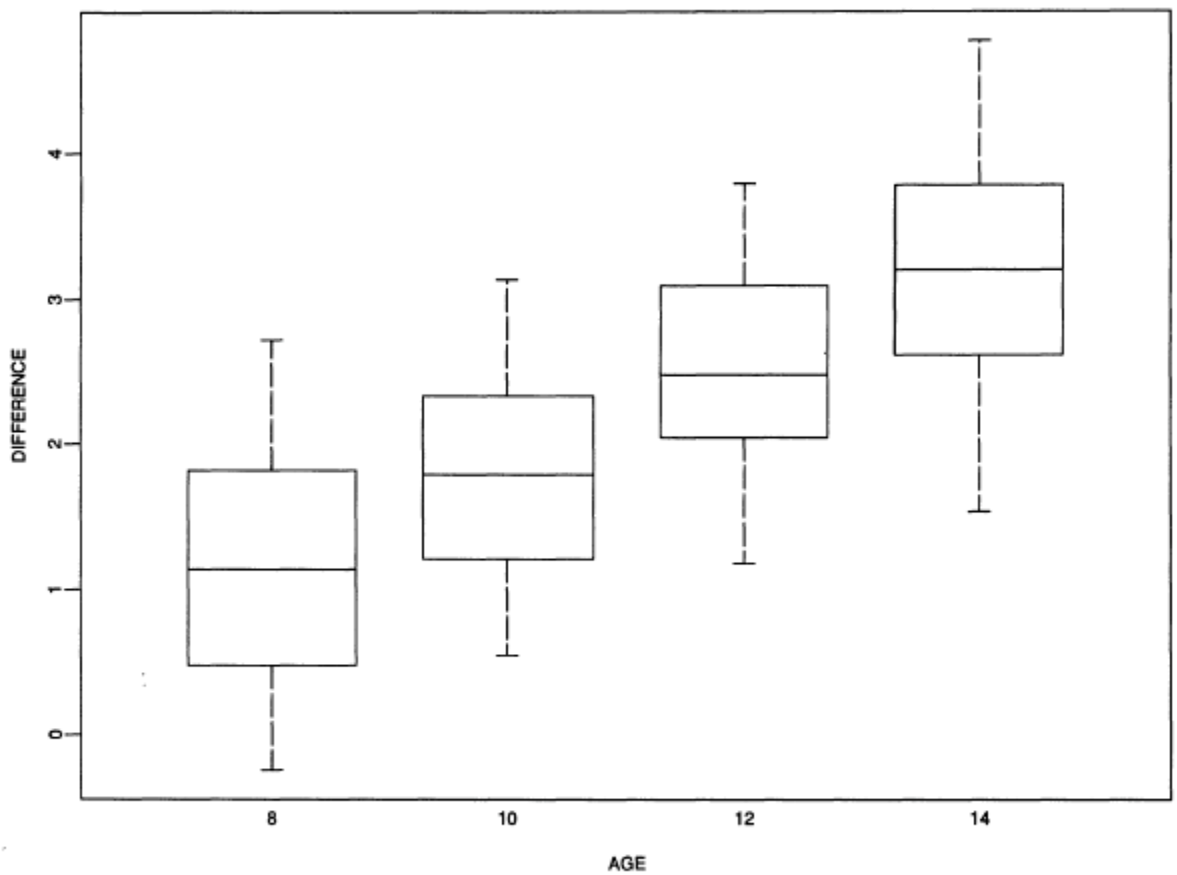


Fig. 3. Difference between growth of boys and girls *versus* age

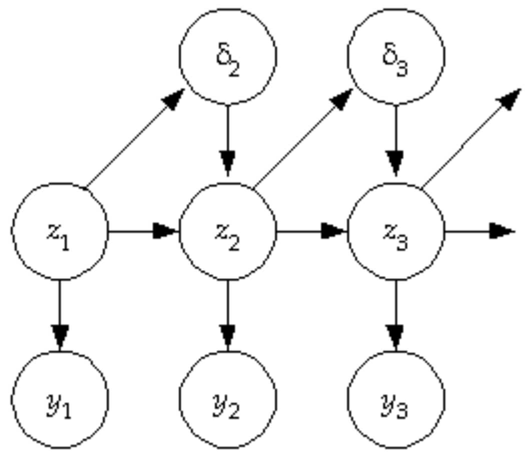


Figure 2: *Random observation time hidden Markov model.*

$$z_t \in \{0, \ldots, K-1\}$$

$$[y_t | z_t = i] \sim \mathcal{N}(0, \sigma_i^2), i = 0, \ldots, K-1$$

$$[\delta_t | z_{t-1} = i] \sim geometric(p_i), i = 0, \ldots, K-1$$

Assume that *z* stays in state *i* for an exponential(λ_i) amount of time

Prior distributions for $\mathbf{p}$, $\boldsymbol{\lambda}$, $\boldsymbol{\beta}$, and $\boldsymbol{\sigma}$ complete the model specification:

$$\begin{aligned}
\left[\sigma_i^2\right] &\sim \text{Inv}-\chi^2(\nu_0, \sigma_0^2), i = 0, \ldots, K-1 \\
[p_i] &\sim \text{Beta}(\alpha_p, \beta_p), i = 0, \ldots, K-1 \\
[\lambda_i] &\sim \Gamma(\alpha_\lambda, \beta_\lambda), i = 0, \ldots, K-1 \\
[\beta_i] &\sim \text{Beta}(\alpha_\beta, \beta_\beta), i = 1, \ldots, K-2
\end{aligned}$$

$$[\mathbf{z}, \mathbf{y}, \boldsymbol{\delta}, \mathbf{p}, \boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\sigma}] = [\mathbf{p}][\boldsymbol{\lambda}][\boldsymbol{\beta}][\boldsymbol{\sigma}] \left\{ \prod_{i=1}^{n} [y_i | z_i, \boldsymbol{\sigma}] \right\} \left\{ \prod_{i=2}^{n} [z_i | z_{i-1}, \delta_i, \boldsymbol{\lambda}, \boldsymbol{\beta}][\delta_i | z_{i-1}, \mathbf{p}] \right\}$$

For K=2:

$$p(z_t = 0 | z_{t-1} = 0, \delta_t = t) = \frac{\lambda_1}{\lambda_0 + \lambda_1} + \frac{\lambda_0}{\lambda_0 + \lambda_1} \exp^{-(\lambda_0 + \lambda_1)t}$$

$$p(z_t = 0 | z_{t-1} = 1, \delta_t = t) = \frac{\lambda_1}{\lambda_0 + \lambda_1} - \frac{\lambda_1}{\lambda_0 + \lambda_1} \exp^{-(\lambda_0 + \lambda_1)t}$$

$$
\begin{aligned}
[z_t|-] &\propto [z_t|z_{t-1}, \delta_t, \lambda_0, \lambda_1]\, [z_{t+1}|z_t, \delta_{t+1}, \lambda_0, \lambda_1]\, [\delta_{t+1}|z_t, p_0, p_1]\, [y_t|z_t, \sigma_0, \sigma_1] \\
\left[\sigma_i^2|-\right] &\propto [\sigma_i] \prod_{\substack{t=1 \\ z_t=i}}^{n} [y_t|z_t, \sigma_i] \\
[p_i|-] &\propto [p_i] \sum_{\substack{t=2 \\ z_{t-1}=i}}^{n} [\delta_t|z_{t-1}, p_i] \\
[\boldsymbol{\lambda}, \boldsymbol{\beta}|-] &\propto [\boldsymbol{\lambda}]\, [\boldsymbol{\beta}] \prod_{t=2}^{n} [z_t|z_{t-1}, \delta_t, \boldsymbol{\lambda}, \boldsymbol{\beta}]
\end{aligned}
$$

First one is easy requiring *K* calculations and a normalization step…

Next two are conjugate…

Last one not so simple…

We use a Metropolis step to sample values of $\mathbf{\Lambda} \equiv (\boldsymbol{\lambda}, \boldsymbol{\beta})$. Specifically, let $q(\mathbf{\Lambda}, \mathbf{\Lambda}')$ denote the proposal density for the transition from $\mathbf{\Lambda}$ to $\mathbf{\Lambda}'$. This proposal density may depend on any or all of $\mathbf{z}, \mathbf{y}, \boldsymbol{\delta}, \boldsymbol{\sigma}$ and $\mathbf{p}$. Then accept $\mathbf{\Lambda}'$ with probability:

$$\alpha(\mathbf{\Lambda}, \mathbf{\Lambda}' | \mathbf{z}, \boldsymbol{\delta}) = \min\{1, \frac{[\mathbf{\Lambda}'] \prod_{t=2}^{n} [z_t | z_{t-1}, \delta_t, \mathbf{\Lambda}']}{[\mathbf{\Lambda}] \prod_{t=2}^{n} [z_t | z_{t-1}, \delta_t, \mathbf{\Lambda}]} \times \frac{q(\mathbf{\Lambda}', \mathbf{\Lambda})}{q(\mathbf{\Lambda}, \mathbf{\Lambda}')}\}.$$

An independence Metropolis sampler is straightforward to implement and is the approach we adopt. For $i = 0, \ldots, K-1$, sample a candidate λ_i' uniformly in the interval $[\lambda_i - c_\lambda, \lambda_i + c_\lambda]$ where c_λ is chosen experimentally.