Stat 535 C - Statistical Computing & Monte Carlo Methods

Lecture 18 - 16th March 2006

Arnaud Doucet

Email: arnaud@cs.ubc.ca

1.1– Outline

- Trans-dimensional Markov chain Monte Carlo.
- Bayesian model for autoregressions.
- Bayesian analysis of finite mixture of Gaussians.

2.1– Metropolis-Hastings

- The standard MH algorithm where $\mathcal{X} \subset \mathbb{R}^d$ corresponds to

$$K\left(x, dx'\right) = \alpha\left(x, x'\right) q\left(x, dx'\right) + \left(1 - \int \alpha\left(x, z\right) q\left(x, dz\right)\right) \delta_x\left(dx'\right)$$

where

$$\alpha\left(x, x'\right) = \min\left\{1, \frac{\pi\left(x'\right) q\left(x', x\right)}{\pi\left(x\right) q\left(x, x'\right)}\right\}$$

- You should think of

$$\frac{\pi\left(x'\right) q\left(x', x\right)}{\pi\left(x\right) q\left(x, x'\right)}$$

not as just a “number”!

2.1– Metropolis-Hastings

- The acceptance ratio corresponds to a ratio of probability measures -importance weight- defined on the same spaces

$$\frac{\pi\left(dx'\right) q\left(x', dx\right)}{\pi\left(dx\right) q\left(x, dx'\right)} = \frac{\pi\left(x'\right) dx' q\left(x', x\right) dx}{\pi\left(x\right) dx q\left(x, x'\right) dx'} = \frac{\pi\left(x'\right) q\left(x', x\right)}{\pi\left(x\right) q\left(x, x'\right)}.$$

- You can only compared points defined on the same joint space. If you have $x = \left(x_1, x_2\right)$ and $\pi_1\left(dx_1\right) = \pi_1\left(x_1\right) dx_1$, $\pi_2\left(dx_1, dx_2\right) = \pi_2\left(x_1, x_2\right) dx_1 dx_2$, you can compute numerically

$$\frac{\pi_2\left(x_1, x_2\right)}{\pi_1\left(x_1\right)}$$

but it means *nothing* as the measures π_1 and π_2 are not defined on the same space. You CANNOT compare a surface to a volume!

2.3– Example: Birth/Death Moves

• The acceptance probability for the associated "death move" is

$$\min\left(1,\frac{\pi\left(1,\theta\right)}{\pi\left(2,\theta_1,\theta_2\right)}g\left(u\right)\left|\frac{\partial\left(\theta,u\right)}{\partial\left(\theta_1,\theta_2\right)}\right|\right)=\min\left(1,\frac{\pi\left(1,\theta\right)g\left(u\right)}{\pi\left(2,\theta,u\right)}\right)$$

• Once the birth move is defined then the death move follows automatically. In the death move, we do not simulate from g but its expression still appear in the acceptance probability.

2.3– Example: Birth/Death Moves

• Assume we have a distribution defined on $\{1\}\times\mathbb{R}\cup\{2\}\times\mathbb{R}\times\mathbb{R}$. We want to propose some moves to go from $(1,\theta)$ to $(2,\theta_1,\theta_2)$.

• We can propose $u\sim g\in\mathbb{R}$ and set

$$\left(\theta_1,\theta_2\right)=h\left(\theta,u\right)=\left(\theta,u\right),$$

i.e. we do not need to introduce a variable u'. Its inverse is given by

$$\left(\theta,u\right)=h'\left(\theta_1,\theta_2\right)=\left(\theta_1,\theta_2\right).$$

• The acceptance probability for this "birth" move is given by

$$\min\left(1,\frac{\pi\left(2,\theta_1,\theta_2\right)}{\pi\left(1,\theta\right)}\frac{1}{g\left(u\right)}\left|\frac{\partial\left(\theta_1,\theta_2\right)}{\partial\left(\theta,u\right)}\right|\right)=\min\left(1,\frac{\pi\left(2,\theta_1,\theta_2\right)}{\pi\left(1,\theta_1\right)g\left(\theta_2\right)}\right).$$

2.2– Designing trans-dimensional moves

• We can rewrite formally

$$\pi\left(dx\right)q\left(x,\left(dx',du'\right)\right)=\pi\left(x\right)g\left(u\right)dxdu$$

and

$$\pi\left(dx'\right)q\left(x',\left(dx,du\right)\right)=\pi\left(x'\right)g'\left(u'\right)dx'du'.$$

• An acceptance ratio ensuring $\pi-$reversibility of this trans-dimensional move is given by

$$\frac{\pi\left(dx'\right)q\left(x',\left(dx,du\right)\right)}{\pi\left(dx\right)q\left(x,\left(dx',du'\right)\right)}=\frac{\pi\left(x'\right)g'\left(u'\right)}{\pi\left(x'\right)g\left(u\right)}\left|\frac{\partial\left(x',u'\right)}{\partial\left(x,u\right)}\right|.$$

• In this respect, the RJMCMC is an extension of standard MH as you need introduce auxiliary variables u and u'.

2.2– Designing trans-dimensional moves

• In the general case where $\mathcal{X}$ is a union of subspaces of different dimensions, you might want to move from $x\in\mathbb{R}^d$ to $x'\in\mathbb{R}^d$.

• To construct this move, you can use $u\in\mathbb{R}^r$ and $u'\in\mathbb{R}^{r'}$ and a *one-to-one differentiable* mapping $h{:}\mathbb{R}^d\times\mathbb{R}^r\rightarrow\mathbb{R}^{d'}\times\mathbb{R}^{r'}$

$$\left(x',u'\right)=h\left(x,u\right)\ \text{ where } u\sim g$$

and

$$\left(x,u\right)=h^{-1}\left(x',u'\right)\ \text{ where } u\sim g'.$$

• We need $d+r=d'+r'$ and typically, if $d<d'$, then $r'=0$ and $r=d'-r$, that is in most case the variable u' is not introduced.

2.5– Example: Split/Merge Moves

• The acceptance probability for the associated "merge move" is

$$\min\left(1, \frac{\pi(1,\theta)}{\pi(2,\theta_1,\theta_2)} g(u) \left|\frac{\partial(\theta,u)}{\partial(\theta_1,\theta_2)}\right|\right) = \min\left(1, \frac{\pi(1,\theta)}{\pi(2,\theta-u,\theta+u)} \frac{g(u)}{2}\right)$$

• Once the split move is defined then the merge move follows automatically. In the merge move, we do not simulate from g but its expression still appear in the acceptance probability.

2.5– Example: Split/Merge Moves

• Assume we have a distribution defined on $\{1\} \times \mathbb{R} \cup \{2\} \times \mathbb{R} \times \mathbb{R}$. We want to propose some moves to go from $(1,\theta)$ to $(2,\theta_1,\theta_2)$.

• We can propose $u \sim g \in \mathbb{R}$ and set

$$(\theta_1,\theta_2) = h(\theta,u) = (\theta-u, \theta+u).$$

Its inverse is given by

$$(\theta,u) = h'(\theta_1,\theta_2) = \left(\frac{\theta_1+\theta_2}{2}, \frac{\theta_2-\theta_1}{2}\right).$$

• The acceptance probability for this "split" move is given by

$$\min\left(1, \frac{\pi(2,\theta_1,\theta_2)}{\pi(1,\theta)} \frac{1}{g(u)} \left|\frac{\partial(\theta_1,\theta_2)}{\partial(\theta,u)}\right|\right) = \min\left(1, \frac{\pi(2,\theta_1,\theta_2)}{\pi\left(1,\frac{\theta_1+\theta_2}{2}\right)} \frac{2}{g\left(\frac{\theta_2-\theta_1}{2}\right)}\right).$$

2.4– Example: Birth/Death Moves

• The acceptance probability for the associated "death move" is

$$\min\left(1, \frac{\pi(1,\theta)}{\pi(2,\theta_1,\theta_2)} g(u|\theta) \left|\frac{\partial(\theta,u)}{\partial(\theta_1,\theta_2)}\right|\right) = \min\left(1, \frac{\pi(1,\theta)\, g(u|\theta)}{\pi(2,\theta,u)}\right)$$

• Once the birth move is defined then the death move follows automatically. In the death move, we do not simulate from g but its expression still appears in the acceptance probability.

• Clearly if we have $g(\theta_2|\theta_1) = \pi(\theta_2|2,\theta_1)$ then the expressions simplify

$$\begin{aligned} \min\left(1, \frac{\pi(2,\theta_1,\theta_2)}{\pi(1,\theta_1)\, g(\theta_2|\theta_1)}\right) &= \min\left(1, \frac{\pi(2,\theta_1)}{\pi(1,\theta_1)}\right), \\ \min\left(1, \frac{\pi(1,\theta)\, g(u|\theta)}{\pi(2,\theta,u)}\right) &= \min\left(1, \frac{\pi(1,\theta)}{\pi(2,\theta)}\right). \end{aligned}$$

2.4– Example: Birth/Death Moves

• To simplify notation -as in Green (1995) & Robert (2004)-, we don't emphasize that actually we can have the proposal g which is a function of the current point θ but it is possible!

• We can propose $u \sim g(\cdot|\theta) \in \mathbb{R}$ and set

$$(\theta_1,\theta_2) = h(\theta,u) = (\theta,u).$$

Its inverse is given by

$$(\theta,u) = h'(\theta_1,\theta_2) = (\theta_1,\theta_2).$$

• The acceptance probability for this "birth" move is given by

$$\min\left(1, \frac{\pi(2,\theta_1,\theta_2)}{\pi(1,\theta)} \frac{1}{g(u|\theta)} \left|\frac{\partial(\theta_1,\theta_2)}{\partial(\theta,u)}\right|\right) = \min\left(1, \frac{\pi(2,\theta_1,\theta_2)}{\pi(1,\theta_1)\, g(\theta_2|\theta_1)}\right).$$

2.6– Mixture of Moves

• In practice, the algorithm is based on a combination of moves to move from $x=(k,\theta_k)$ to $x'=(k',\theta_{k'})$ indexed by $i\in\mathcal{M}$ and in this case we just need to have

$$\int_{(x,x')\in A\times B}\pi(dx)\,\alpha_i(x,x')\,q_i(x,dx')=\int_{(x,x')\in A\times B}\pi(dx')\,\alpha_i(x',x)\,q_i(x',dx)$$

to ensure that the kernel $P(x,B)$ defined for $x\notin B$

$$P(x,B)=\frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}}\alpha_i(x,x')\,q_i(x,dx')$$

is π-reversible.

• In practice, we would like to have

$$P(x,B)=\sum_{i\in\mathcal{M}}j_i(x)\,\alpha_i(x,x')\,q_i(x,dx')$$

where $j_i(x)$ is the probability of selecting the move i once we are in x and $\sum_{i\in\mathcal{M}}j_i(x)=1$.

2.6– Mixture of Moves

• In this case reversibility is ensured if

$$\int_{(x,x')\in A\times B}\pi(dx)\,j_i(x)\,\alpha_i(x,x')\,q_i(x,dx')$$

$$=\int_{(x,x')\in A\times B}\pi(dx')\,j_i(x')\,\alpha_i(x',x)\,q_i(x',dx)$$

which is satisfied if

$$\alpha_i(x,x')=\min\left(1,\frac{\pi(x')\,j_i(x')\,g_i'(u')}{\pi(x)\,j_i(x)\,g_i(u)}\left|\frac{\partial(x',u')}{\partial(x,u)}\right|\right).$$

• In practice, we will only have a limited number of moves possible from each point x.

2.7– Summary

• For each point $x=(k,\theta_k)$, we define a collection of potential moves selected randomly with probability $j_i(x)$ where $i\in\mathcal{M}$.

• To move from $x=(k,\theta_k)$ to $x'=(k',\theta_{k'})$, we build one (or several) deterministic differentiable and inversible mapping(s)

$$(\theta_{k'},u_{k',k})=T_{k,k'}(\theta_k,u_{k,k'})$$

where $u_{k,k'}\sim g_{k,k'}$ and $u_{k',k}\sim g_{k',k}$ and we accept the move with proba

$$\min\left(1,\frac{\pi(k',\theta_{k'})\,j_i(k',\theta_{k'})\,g_{k',k}(u_{k',k})}{\pi(k,\theta_k)\,j_i(k,\theta_k)\,g_{k,k'}(u_{k,k'})}\left|\frac{\partial T_{k,k'}(\theta_k,u_{k,k'})}{\partial(\theta_k,u_{k,k'})}\right|\right).$$

2.8– One minute break

• This brilliant idea is due to P.J. Green, *Reversible Jump MCMC and Bayesian Model Determination*, Biometrika, 1995 although special cases had appeared earlier in physics.

• This is one of the top ten most cited paper in maths and is used nowadays in numerous applications including genetics, econometrics, computer graphics, ecology, etc.

2.9– Example: Bayesian Model for Autoregressions

• The performance are obviously very dependent on the selection of the proposal distribution. We select whenever possible the full conditional distribution, i.e. we have $u = a_{k+1,k+1} \sim p\left(a_{k+1,k+1} \middle| y_{1:T}, a_{k,1:k}, \sigma_k^2, k+1\right)$ and

$$\begin{aligned} & \min\left(1, \frac{p\left(a_{k,1:k}, u, \sigma_k^2, k+1 \middle| y_{1:T}\right) d_{k+1}}{p\left(a_{k,1:k}, \sigma_k^2, k \middle| y_{1:T}\right) b_k p\left(u \middle| y_{1:T}, a_{k,1:k}, \sigma_k^2, k+1\right)}\right) \\ = & \min\left(1, \frac{p\left(a_{k,1:k}, \sigma_k^2, k+1 \middle| y_{1:T}\right) d_{k+1}}{p\left(a_{k,1:k}, \sigma_k^2, k \middle| y_{1:T}\right) b_k}\right). \end{aligned}$$

• In such cases, it is actually possible to reject a candidate before sampling it!

2.9– Example: Bayesian Model for Autoregressions

• *Birth move*: We propose to move from k to $k+1$

$$\left(a_{k+1,1:k}, a_{k+1,k+1}, \sigma_{k+1}^2\right) = \left(a_{k,1:k}, u, \sigma_k^2\right) \text{ where } u \sim g_{k,k+1}$$

and the acceptance probability is

$$\min\left(1, \frac{p\left(a_{k,1:k}, u, \sigma_k^2, k+1 \middle| y_{1:T}\right) d_{k+1}}{p\left(a_{k,1:k}, \sigma_k^2, k \middle| y_{1:T}\right) b_k g_{k,k+1}(u)}\right).$$

• *Death move*: We propose to move from k to $k-1$

$$\left(a_{k-1,1:k-1}, u, \sigma_{k-1}^2\right) = \left(a_{k,1:k-1}, a_{k,k}, \sigma_k^2\right)$$

and the acceptance probability is

$$\min\left(1, \frac{p\left(a_{k,1:k-1}, \sigma_k^2, k-1 \middle| y_{1:T}\right) b_{k-1} g_{k-1,k}(a_{k,k})}{p\left(a_{k,1:k}, \sigma_k^2, k \middle| y_{1:T}\right) d_k}\right)$$

2.9– Example: Bayesian Model for Autoregressions

• We propose the following moves. If we have $\left(k, a_{1:k}, \sigma_k^2\right)$ then with probability b_k we propose a birth move if $k \leq k_{\max}$, with proba u_k we propose an update move and with proba $d_k = 1 - b_k - u_k$ we propose a death move.

• We have $d_1 = 0$ and $b_{k\max} = 0$.

• The *update move* can simply done in a Gibbs step as

$$p\left(\theta_k \middle| y_{1:T}, k\right) = \mathcal{N}\left(a_{k,1:k}; m_k, \sigma^2 \Sigma_k\right) \mathcal{IG}\left(\sigma^2; \frac{\nu_k}{2}, \frac{\gamma_k}{2}\right)$$

2.9– Example: Bayesian Model for Autoregressions

• The model $k \in \mathcal{K} = \{1, ..., k_{\max}\}$ is given by an AR of order k

$$Y_n = \sum_{i=1}^{k} a_i Y_{n-i} + \sigma V_n \text{ where } V_n \sim \mathcal{N}(0,1)$$

and we have $\theta_k = \left(a_{k,1:k}, \sigma_k^2\right) \in \mathbb{R}^k \times \mathbb{R}^+$ where

$$\begin{aligned} p(k) &= k_{\max}^{-1} \text{ for } k \in \mathcal{K}, \\ p(\theta_k | k) &= \mathcal{N}\left(a_{k,1:k}; 0, \sigma_k^2 \delta^2 I_k\right) \mathcal{IG}\left(\sigma^2; \frac{\nu_0}{2}, \frac{\gamma_0}{2}\right). \end{aligned}$$

• For sake of simplicity, we assume here that the initial conditions $y_{1-k_{\max}:0} = (0, ..., 0)$ are known and we want to sample from

$$p\left(\theta_k, k \middle| y_{1:T}\right).$$

• Note that this is not very clever as $p\left(k \middle| y_{1:T}\right)$ is known up to a normalizing constant!

2.9– Example: Bayesian Model for Autoregressions

- We simulate 200 data with $k=5$ and use 10,000 iterations of RJMCMC.

- The algorithm output is $\left(k^{(i)},\theta_k^{(i)}\right)\sim p\left(\theta_k,k\mid y\right)$ (asymptotically).

- The histogram of $\left(k^{(i)}\right)$ yields an estimate of $p\left(k\mid y\right)$.

- Histograms of $\left(\theta_k^{(i)}\right)$ for which $k^{(i)}=k_0$ yields estimates of $p\left(\theta_{k_0}\mid y,k_0\right)$.

- The algorithm provides us with an estimate of $p\left(k\mid y\right)$ which matches analytical expressions.

2.10– Finite Mixture of Gaussians

- The model $k\in\mathcal{K}=\{1,...,k_{\max}\}$ is given by a mixture of k Gaussians

$$Y_n\sim\sum_{i=1}^{k}\pi_i\mathcal{N}\left(\mu_i,\sigma_i^2\right).$$

and we have $\theta_k=\left(\pi_{1:k},\mu_{1:k},\sigma_{1:k}^2\right)\in S_k\times\mathbb{R}^k\times\left(\mathbb{R}^+\right)^k$.

- We need to defined a prior $p\left(k,\theta_k\right)=p\left(k\right)p\left(\theta_k\mid k\right)$, say

$$\begin{aligned}p\left(k\right)&=k_{\max}^{-1}\text{ for }\in\mathcal{K}\\ p\left(\theta_k\mid k\right)&=\mathcal{D}\left(\pi_{k,1:k};1,...,1\right)\prod_{i=1}^{k}\mathcal{N}\left(\mu_{k,i};\alpha,\beta\right)\mathcal{IG}\left(\sigma_{k,i}^2;\frac{\nu_0}{2},\frac{\gamma_0}{2}\right).\end{aligned}$$

- Given T data, we are interested in $\pi\left(k,\theta_k\mid y_{1:T}\right)$.

2.11– Trans-dimensional MCMC

- When k is fixed, we will use Gibbs steps to sample from $\pi\left(\theta_k,z_{1:T}\mid y_{1:T},k\right)$ where $z_{1:T}$ are the discrete latent variables such that $\Pr\left(z_n=i\mid k,\theta_k\right)=\pi_{k,i}$.

- To allow to move in the model space, we define a birth and death move. The birth and death moves use as a target $\pi\left(\theta_k\mid y_{1:T},k\right)$ and not $\pi\left(\theta_k,z_{1:T}\mid y_{1:T},k\right)$.

$\Rightarrow$ Reduced dimensionality, easier to design moves.

2.12– Birth Move

- We propose a naive move to go from $k\to k+1$ where $j\sim\mathcal{U}_{\{1,...,k+1\}}$

$$\begin{aligned}\mu_{k+1,-j}&=\mu_{k,1:k},\ \sigma_{k+1,-j}^2=\sigma_{k,1:k}^2,\\ \pi_{k+1,-j}&=\left(1-\pi_{k+1,j}\right)\pi_{k,-j},\end{aligned}$$

where $\left(\pi_{k+1,j},\mu_{k+1,j},\sigma_{k+1,j}^2\right)\sim g_{k,k+1}$ (prior distribution in practice).

- The Jacobian of the transformation is $\left(1-\pi_{k+1,j}\right)^{k-1}$ (only $k-1$ “true” variables for $\pi_{k,-j}$)

- Now one has to be careful when considering the reverse death move. Assume the death move going from $k+1\to k$ by removing the component j.

2.13– Split and Merge Moves

- It follows that the acceptance probability of the split move with $j \sim \mathcal{U}_{\{1,\ldots,k\}}$ is $\min(1, A)$ where

$$A = \frac{\pi\left(k+1, \pi_{k+1,1:k+1}, \mu_{k+1,1:k+1}, \sigma^2_{k+1,1:k+1} \,\middle|\, y_{1:T}\right)}{\pi\left(k, \pi_{k,1:k}, \mu_{k,1:k}, \sigma^2_{k,1:k} \,\middle|\, y_{1:T}\right)} \times \frac{(m_{k+1,k}/k)}{(s_{k,k+1}/k)} \times \frac{\pi^3_{k,j}}{(1-u_1)^2} \left|\mu_{k,j}\right| \sigma^2_{k,j}.$$

- You should think of the split move as a mixture of k split moves and you have k associated merge moves.

2.13– Split and Merge Moves

- The associated merge move is

$$\begin{aligned}
\pi_{k,j} &= \pi_{k+1,j} + \pi_{k+1,j+1}, \\
\pi_{k,j}\mu_{k,j} &= \pi_{k+1,j}\mu_{k+1,j} + \pi_{k+1,j+1}\mu_{k+1,j+1}, \\
\pi_{k,j}\sigma^2_{k,j} &= \pi_{k+1,j}\sigma^2_{k+1,j} + \pi_{k+1,j+1}\sigma^2_{k+1,j+1}.
\end{aligned}$$

- The Jacobian of the transformation of the split is given by

$$\left|\frac{\partial\left(\pi_{k+1,1:k+1}, \mu_{k+1,1:k+1}, \sigma^2_{k+1,1:k+1}\right)}{\partial\left(\pi_{k,1:k}, \mu_{k,1:k}, \sigma^2_{k,1:k}, u_1, u_2, u_3\right)}\right| = \frac{\pi^3_{k,j}}{(1-u_1)^2} \left|\mu_{k,j}\right| \sigma^2_{k,j}.$$

2.13– Split and Merge Moves

- To move from $k \to k+1$, one can also select a split move of the component $j \sim \mathcal{U}_{\{1,\ldots,k\}}$

$$\begin{aligned}
\pi_{k+1,j} &= u_1\pi_{k,j},\ \pi_{k+1,j+1} = (1-u_1)\,\pi_{k,j}, \\
\mu_{k+1,j} &= u_2\mu_{k,j},\ \mu_{k+1,j+1} = \frac{\pi_{k,j} - \pi_{k+1,j}u_2}{\pi_{k,j} - \pi_{k+1,j}}\mu_{k,j}, \\
\sigma^2_{k+1,j} &= u_3\sigma^2_{k,j},\ \sigma^2_{k+1,j+1} = \frac{\pi_{k,j} - \pi_{k+1,j}u_3}{\pi_{k,j} - \pi_{k+1,j}}\sigma^2_{k,j}
\end{aligned}$$

with $u_1, u_2, u_3 \sim \mathcal{U}(0,1)$.

2.12– Birth Move

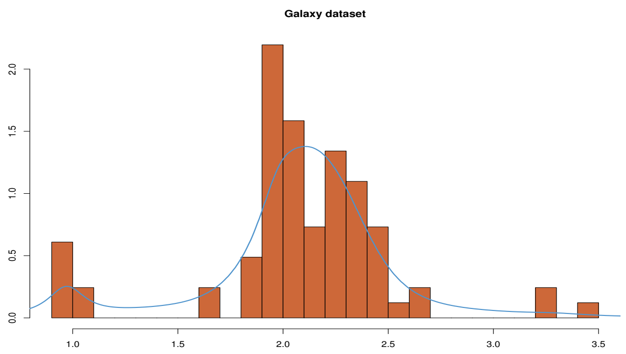
- The acceptance probability of the birth move is given by $\min(1, A)$ where

$$A = \frac{\pi\left(k+1, \pi_{k+1,1:k+1}, \mu_{k+1,1:k+1}, \sigma^2_{k+1,1:k+1} \,\middle|\, y_{1:T}\right)}{\pi\left(k, \pi_{k,1:k}, \mu_{k,1:k}, \sigma^2_{k,1:k} \,\middle|\, y_{1:T}\right)} \times \frac{\left(d_{k+1,k}/\left(k+1\right)\right)\left(1-\pi_{k+1,j}\right)^{k-1}}{\left(b_{k,k+1}/\left(k+1\right)\right) g_{k,k+1}\left(\pi_{k+1,j}, \mu_{k+1,j}, \sigma^2_j\right)}.$$

- This move will work properly if the prior is not too diffuse. Otherwise the acceptance probability will be small.

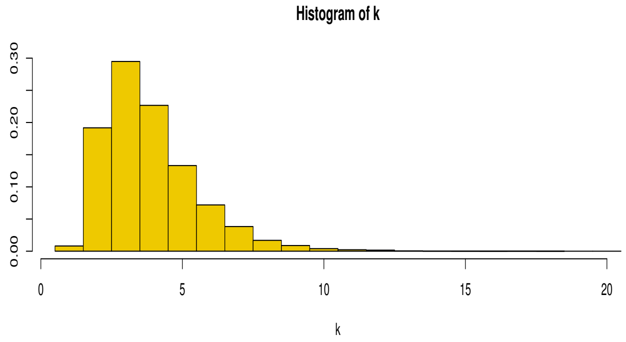
- We have $(k+1)$ birth moves to move from $k \to k+1$ and $k+1$ associated death moves.

2.14– Application to the Galaxy Dataset



Estimation of $E\left[f\left(y \mid k, \theta_k \right) \mid y_{1:T} \right]$

2.14– Application to the Galaxy Dataset

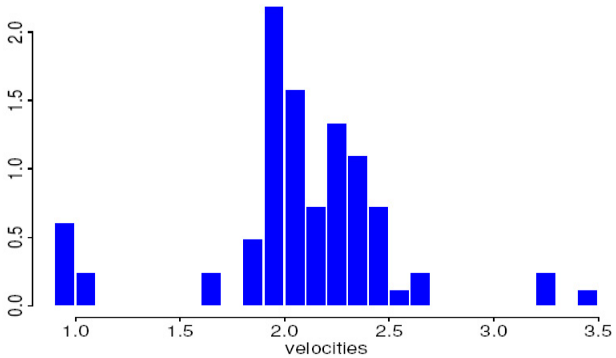


Estimation of the marginal posterior distribution $p\left(k \mid y_{1:T} \right)$.

2.14– Application to the Galaxy Dataset

- We set $k_{\max} = 20$ and we select (rather) informative priors following Green & Richardson (1999). In practice, it is worth using a hierarchical prior.
- We run the algorithm for over 1,000,000 iterations.
- We set additional constraints on the mean $\mu_{k,1} < \mu_{k,2} < .... < \mu_{k,k}$.
- The cumulative averages stabilize very quickly.

2.14– Application to the Galaxy Dataset



Velocity (km/sc) of galaxies in the Corona Borealis Region

2.15– Summary

• Trans-dimensional MCMC allows us to implement numerically problems with Bayesian model uncertainty.

• Practical implementation is relatively easy, theory behind not so easy...

• Designing efficient trans-dimensional MCMC algorithms is still a research problem.