Matematiska vetenskaper
Chalmers tekniska högskola

Exam: 2014-10-27, Time: 14:00
Computer Intensive Statistical methods
(MSA100/MVE186)

Examiner: Jonas Wallin, tel 772 53 23.

Allowed aids: None.

Correct, well motivated solution gives points indicated within the parentheses at each exercise. Total number of points is 42. The grade limits are 3:18, 4:24 och 5:32 points.

---

1. Show that for the bivariate distribution $f(x_1, x_2)$, the two stage Gibbs sampler has f as stationary distribution. (5p)

2. Show that Metropolis-Hastings algorithm satisfies the global balance condition. (5p)

3. Using only standard normal ($\mathcal{N}(0,1)$) and uniform ($U[0,1]$) random variables, write down algorithms that draws random numbers from:

   (a) The exponential distribution, $Exp(\lambda)$ (2p)

   (b) The raised cosine distribution, $Rcos(\mu)$. (2p)

   (c) Truncated normal, $\mathcal{N}(x;0,1)\mathcal{I}(x \in [c,d])$. (2p)

4. In the following exercise f is a density taking values in $\mathbb{R}$, and $h : \mathbb{R} \to \mathbb{R}$.

   (a) Define the importance sampler estimator τ_n of $\tau = \mathbb{E}_f[h(x)]$ with instrumental density g. Also show that the importance sampler is unbiased. (2p)

   (b) Derive the variance of the importance sampler. (2p)

   (c) For the function $h(x) = x$ and the density $f(x) \propto e^{-x^2-|x|}$. Which is best of the two instrumental densities $g_1(x) \propto e^{-|x|}$ and $g_2(x) \propto e^{-x^2}$? (2p)

5. To handle a data set,$\mathbf{y}$, with clear bimodal density, and where $y_i \in \mathbb{N}$, the following model is suggested:

$$\pi(\lambda_i) = \Gamma(\lambda_i, 1, 0.1)\,, i = 1,2,$$
$$\pi(p) \propto 1.$$
$$f(\mathbf{y}|\lambda_1,\lambda_2,p) = \prod_{i=1}^{n} \left(pPois(y_i;\lambda_1) + (1-p)Pois(y_i;\lambda_2)\right)$$

   The goal of this exercise is to generate the posterior distribution using MCMC.

   (a) Suggest auxiliary variables $\mathbf{z}$ which makes it possible to Gibbs sample the posterior. Derive the joint distribution $f(\mathbf{z},\mathbf{y},\lambda_1,\lambda_2,p)$. (4p)

   b) Derive the conditional distribution for each variable in its parametric form. (4p)

   c) It turns out that the bi-modality could be explained by a covariate $\mathbf{x}$. A suitable model is then given by

$$\pi(\beta_i) = \mathcal{N}(0,10)\,, i = 1,2,$$
$$f(\mathbf{y}|\lambda_1,\lambda_2,p) = \prod_{i=1}^{n} Pois(y_i; e^{\beta_0+\beta_1 x_i})$$

   To sample from the posterior distribution a Metropolis-Hastings random walk is suggested. Derive the acceptance probability in the Metropolis-Hastings algorithm. (4p)

6. Derive why the Markov chain generated by the Gibbs sampler below does not converge to a distribution?

   given (x^*, y^*)
   draw $X \sim f_{X|Y}(x|y^*) = e^{-xy^*}$.
   draw $Y \sim f_{Y|X}(y|X) = e^{-yX}$.
   return (X, Y)

   Why? motivate your answer (8p)

- Normal distribution.

Notation	$\mathcal{N}(x;\mu,\sigma^2)$
Suport	$x \in \mathbb{R}$
parameter	$\sigma^2 > 0, \mu \in \mathbb{R}$
pdf	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

- Beta distribution.

Notation	$\mathcal{B}(x;\alpha,\beta)$
Suport	$x \in [0,1]$
parameter	$\alpha > 0, \beta > 0$
pdf	$\frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$

- Gamma distribution.

Notation	$\Gamma(x;\alpha,\beta)$
Suport	$x \in [0,\infty)$
parameter	$\alpha > 0, \beta > 0$
pdf	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$

- inverse Gamma distribution.

Notation	$IG(x;\alpha,\beta)$
Suport	$x \in [0,\infty)$
parameter	$\alpha > 0, \beta > 0$
pdf	$\frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta x^{-1}}$

- Exponential distribution.

Notation	$exp(x;\lambda)$
Suport	$x \in [0,\infty)$
parameter	$\lambda > 0$
pdf	$\lambda e^{-\lambda x}$
cdf	$1 - e^{-\lambda x}$

- Raised cosine

Notation	$Rcos(x;\mu)$
Suport	$x \in [\mu - 1, \mu + 1]$
parameter	$\mu \in \mathbb{R}$
pdf	$\frac{1}{2}[1 + \cos(\pi(x-\mu))]$

- Poisson

Notation	$Pois(x;\lambda)$
Suport	$x \in \mathbb{N}$
parameter	$\lambda > 0$
pdf	$\frac{\lambda^x}{x!}e^{-\lambda}$

- Binomial

Notation	$Bin(x;n,p)$
Suport	$x \in \{0,1,\ldots,n\}$
parameter	$p \in [0,1], n \in \mathbb{N}$
pdf	$\binom{n}{x}(1-p)^{n-x}p^x$

LÖSNINGAR!

Matematiska vetenskaper
Chalmers tekniska högskola

Exam: 2014-10-27, Time: 14:00
Computer Intensive Statistical methods
(MSA100/MVE186)

Examiner: Jonas Wallin, tel 772 53 23.

Allowed aids: None.

Correct, well motivated solution gives points indicated within the parentheses at each exercise. Total number of points is 42. The grade limits are 3:18, 4:24 och 5:32 points.

---

1. See lecture 8, slide 24. (5p)
2. See lecture 9, slide 25.
3. Using only standard normal ($\mathcal{N}(0,1)$) and uniform ($U[0,1]$) random variables, write down algorithms that draws random numbers from:
   (a) Draw $U \sim U[0,1]$, set $X = -log(-U)/\lambda$
   (b) An accept reject algorithm where the proposal distribution is $U[0,1]$, and constant $M = 1$.
   (c) Repeat $X \sim \mathcal{N}(0,1)$ until $X \in [c,d]$.
4. In the following exercise f is a density taking values in $\mathbb{R}$, and $h : \mathbb{R} \to \mathbb{R}$.
   (a) $\tau_n = \frac{1}{n}\sum_{i=1}^{n} h(X_i)\frac{f(X_i)}{g(X_i)}$ where $X_i \sim g$ and $f(x) > 0 \to g(x) > 0$. The estimator is unbiased since

$$\mathbb{E}_g[\tau_n] = \frac{1}{n}\sum_{i=1}^{n}\int h(x_i)\frac{f(x_i)}{g(x_i)}\,g(x_i)\,dx_i = \int h(x)f(x)\,dx = \tau.$$

(2p)

   (b)

$$\mathbb{V}_g[\tau_n] = \frac{1}{n^2}\left(\int h^2(x)\frac{f^2(x)}{g(x)}\,dx - \tau^2\right)$$

(2p)

   (c) $g_1(x)$ is better then $g_2(x)$ since its tails covers the tail of $f(x)$.
5. To handle a data set,$\mathbf{y}$, with clear bimodal density, and where $y_i \in \mathbb{N}$, the following model is suggested:

$$\pi(\lambda_i) = \Gamma(\lambda_i, 1, 0.1)\,, i = 1, 2,$$
$$\pi(p) \propto 1.$$
$$f(\mathbf{y}|\lambda_1, \lambda_2, p) = \prod_{i=1}^{n}\left(pPois(y_i;\lambda_1) + (1-p)Pois(y_i;\lambda_2)\right)$$

The goal of this exercise is to generate the posterior distribution using MCMC.

   (a) A suitable auxiliary variables is $z_i \sim Bin(1,p)$ and $\xi_i \sim Pois(\lambda_i)$ since then

$$y_i = \xi_{2-z_i}$$

   in distribution. The full joint distribution is then

$$f(\mathbf{z}, \mathbf{y}, \lambda_1, \lambda_2, p) = \Gamma(\lambda_1, 1, 0.1)\Gamma(\lambda_2, 1, 0.1)\prod_{i=1}^{n} p^{z_i}(1-p)^{1-z_i}Pois(y_i;\lambda_{2-z_i}).$$

(4p)

b) Since $\mathbb{P}(z_i = 0) \propto pPois(y_i; \lambda_1)$ and $\mathbb{P}(z_i = 1) \propto (1-p)Pois(y_i; \lambda_2)$:

$$z_i|\ldots \sim Bin(1, \frac{pPois(y_i|\lambda_1)}{pPois(y_i;\lambda_1) + (1-p)Pois(y_i;\lambda_2)})$$

For λ_1

$$f(\mathbf{z}, \mathbf{y}, \lambda_1, \lambda_2, p) \propto e^{-0.1\lambda_1} \prod_{i=1}^{N} \lambda_1^{z_i y_i} e^{-z_i \lambda_1} = \lambda_1^{\sum_{i=1}^n z_i y_i} e^{-(\sum_{i=1}^n z_i + 0.1)\lambda_1}$$

Thus

$$\lambda_1|\ldots \sim \Gamma(1 + \sum_{i=1}^{n} z_i y_i, \sum_{i=1}^{n} z_i + 0.1).$$

The distribution of λ_2 follows by same reasoning.
Finally, for p $f(\mathbf{z}, \mathbf{y}, \lambda_1, \lambda_2, p) \propto p^{\sum z_i}(1-p)^{n-\sum z_i}$ thus

$$p|\ldots \sim Beta(\sum z_i + 1, n - \sum z_i + 1).$$

c)

The acceptance rate of MH algorithm is given by

$$\alpha(\beta^*, \beta) = \frac{f(\beta^*)q(\beta|\beta^*)}{f(\beta)q(\beta^*|\beta)},$$

where $f(\beta)$ is the density $\propto f(y|\beta)\pi(\beta)$ and $q(.|.)$ the proposal kernel. Since the proposal kernel is symmetric for a random walk MH the acceptance rate is

$$\frac{f(\beta^*)}{f(\beta)} = \frac{e^{-\frac{(\beta_0^*)^2 + (\beta_1^*)^2}{20} + \sum_{i=1}^n y_i(\beta_0^* + \beta_1^* x_i) - e^{\beta_0^* + \beta_1^* x_i}}}{e^{-\frac{\beta_0^2 - \beta_1^2}{20} + \sum_{i=1}^n y_i(\beta_0 + \beta_1 x_i) - e^{\beta_0 + \beta_1 x_i}}}$$

6. The issue is that $f(x, y)$ is not a density. Since

$$f(x,y) \propto \frac{f_{X|Y}(x|\xi_2) f_{Y|X}(y|x)}{f_{Y|X}(\xi_1|\xi_2) f_{Y|X}(\xi_2|x)} \propto e^{-xy},$$

and $\int e^{-xy} dx dy = \int \frac{1}{x} dx = \infty$ and thus not a density.