MATS442 Stochastic simulation — Problems 4, 7.2.2020

The problems marked with **C** have a part which needs to be implemented in a computer programming language. Please return a single code file with the answers to these exercises by email to ⟨santeri.j.karppinen(·at·)jyu.fi⟩.

1. Prove simplified version of Proposition 5.1 (law of total variance): Suppose that X and Y are any random numbers in $\mathbb{X} = \{1, \dots, m\}$, then

$$\mathrm{Var}(X) = \mathbb{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(\mathbb{E}[X \mid Y]). \tag{1}$$

Hints: If $(X, Y) \sim p$, that is, $\mathbb{P}(X = x, Y = y) = p(x, y)$, recall the decomposition $p(x, y) = p_Y(y)p_{X|Y}(x \mid y)$, where $p_Y(y) = \mathbb{P}(Y = y)$ is the marginal of Y and $p_{X|Y}(x \mid y) = \mathbb{P}(X = x \mid Y = y)$ is the conditional (for which you may assume for simplicity that $p_Y(y) > 0$ so that $p_{X|Y}$ is well-defined...)

Recall that for any function $f : \mathbb{X} \to \mathbb{R}$:

(i) $\mathbb{E}[f(X) \mid Y] = \sum_{x \in \mathbb{X}} f(x)p_{X|Y}(x \mid Y)$

(ii) $\mathbb{E}[\mathbb{E}[f(X) \mid Y]] = \sum_{y \in \mathbb{X}} p_Y(y)\mathbb{E}[f(X) \mid Y = y] = \mathbb{E}[f(X)]$

(iii) $\mathrm{Var}(X \mid Y) = \mathbb{E}[X^2 \mid Y] - (\mathbb{E}[X \mid Y])^2$

2. Prove Proposition 6.16: Show that the Metropolis-Hastings transition probability K is reversible with respect to p (in case $\mathbb{X}$ is countable).
Hints:

(i) Use Proposition 6.15 to start with.

(ii) Consider first the case $x = y$.

(iii) Continue then to the case $x \neq y$.

- Consider the case $q(x, y) > 0$ and $q(y, x) > 0$.
- Consider finally the case $q(x, y) = 0$ or $q(y, x) = 0$.

Remember the properties of $\min\{\,\cdot\,,\,\cdot\,\}$ proved in the last problems class!

3.R Recall that a $\mathrm{Binom}(m, \beta)$ p.m.f. is given as

$$p(k) = \binom{m}{k}\beta^k(1-\beta)^{m-k}, \qquad k \in \mathbb{X} := \{0, \dots, m\}.$$

(a) Implement a Metropolis-Hastings algorithm which targets p and uses $q(k, j) = (m+1)^{-1}$ for all $k, j \in \mathbb{X}$. Hint: You can use `logpdf(Binomial(m, beta), y)` from the `Distributions` package to calculate $\log p(y)$ and use it in your algorithm.

(b) Apply your algorithm to the case of $\mathrm{Binom}(100, 0.9)$, and construct an estimator for $\mathbb{E}_p[X]$ using 100000 samples.

(c) What was the *acceptance rate*, that is, the number of accepted proposals compared to steps of your algorithm?

4.R Let p be $\mathrm{Binom}(n, \beta)$ p.m.f. as above, but

(a) Implement now a Metropolis-Hastings algorithm where your proposals are generated as

$$Y_k = X_{k-1} + [Z_k],$$

where $[x]$ stands for rounding (`round` in Julia; nearest integer to x)[1], and $Z_k \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$, with $\sigma^2 = 5m\beta(1-\beta)$.
Hint: Do not try to calculate $q(k, j)$ but think what properties it has.

1. Do not care about borderline cases, because they happen with probability zero anyway.

(b) Use your algorithm in case of Binom(100, 0.9) and construct an estimator for $\mathbb{E}_p[X]$ using 100000 samples.

(c) Study empirically the mean and variance of your estimator and compare them against the variance of the estimator in Problem 3.

5.R Suppose your model is given as $V \sim U(0, 10)$ and $Y_i \mid V \overset{\text{i.i.d.}}{\sim} N(0, V)$ for $i = 1, \ldots, 10$, and you want to study the distribution $p(v)$, which is the conditional density $V \mid Y_1 = y_1, \ldots, Y_{10} = y_{10}$ where y_i are given as

$$y = [0.32, 0.27, 0.24, -0.09, -0.15, 0.43, -0.53, -0.38, -0.01, -0.04].$$

Note that $p(v) \propto p_V(v) \prod_{i=1}^{10} p_Y(y_i \mid v)$ where p_V is the $U(0, 10)$ density and $p_Y(\cdot \mid v)$ is the $N(0, v)$ density.

(a) Implement MCMC with proposals of the form

$$V_k' = V_{k-1} + W_k, \qquad W_k \overset{\text{i.i.d.}}{\sim} N(0, 0.1^2),$$

and estimate $\mathbb{E}_p[V]$ and $\mathrm{Var}_p[V]$.

(b) Implement MCMC with proposals of the form

$$\log(V_k') = \log(V_{k-1}) + W_k, \qquad W_k \overset{\text{i.i.d.}}{\sim} N(0, 1^2).$$

Check that you get the same result as in (a)!
Hint: Think carefully what the proposal density is. You may also think the problem using transformed target density…

6. Suppose that p_u is an unnormalised p.d.f. on $\mathbb{R}$, which is unimodal: p_u is continuous, increasing in $(-\infty, \beta]$ and decreasing in $[\beta, \infty)$ for some $\beta \in \mathbb{R}$.

Consider the following unnormalised density: $\tilde{p}_u(x, t) := \mathbf{1}(0 \le t \le p_u(x))$ on $\mathbb{R}^2$.

(a) Sketch the set $\{(x, t) \in \mathbb{R}^2 : 0 \le t \le p_u(x)\}$.

(b) Suppose that for any $0 \le p_* < p_u(\beta)$, you can calculate $x_- < x_+$ such that $p_u(x_-) = p_* = p_u(x_+)$. Describe a Gibbs sampling algorithm $(X_k, T_k)_{k \ge 1}$ targetting $\tilde{p}(x, t) \propto \tilde{p}_u(x, t)$.

(c) Suppose that $(X_k, T_k)_{k \ge 1}$ is the outcome of the Gibbs sampling algorithm. Where does

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k)$$

converge to?

The procedure described above is called *slice sampling*. Can you guess how the algorithm can be generalised to p_u on $\mathbb{R}^d$?