MATS442 Stochastic simulation — Problems 5, 14.2.2020

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. Suppose that p and q are probability densities on $\mathbb{R}$, satisfying $q(x) = 0 \implies p(x) = 0$ and you have implemented self-normalised importance sampling, that is,

$$X_1, \ldots, X_n \sim q(\,\cdot\,), \qquad W_k^{(n)} = \frac{w_u(X_k)}{\sum_{j=1}^n w_u(X_j)}, \qquad w_u(x) = c_w \frac{p(x)}{q(x)}.$$

(a) How can you estimate $\mathrm{Var}_p(X)$ using $(X_1, \ldots, X_n)$ and $(W_1^{(n)}, \ldots, W_n^{(n)})$? (Hint: Express $\mathrm{Var}_p(X)$ in terms of expectations.)

(b) Show that your estimator is consistent.

(c) How about quantiles, that is, for a given $\beta \in (0,1)$, how can you estimate the number $x_\beta \in \mathbb{R}$ such that $\mathbb{P}(X \le x_\beta) = \beta$ where $X \sim p$? (Hint: Think first what might be an estimator of $F(x) = \mathbb{P}(X \le x)$.)

(d) (*) Can you show that the quantile estimator is consistent if F is strictly increasing?

2. Consider Algorithm 6.13 (Metropolis-Hastings).

(a) Observe that $(X_k, Y_{k+1})_{k\ge1}$ is a Markov chain.

(b) Show that $(X_k, Y_{k+1})_{k\ge1}$ admits $\tilde{p}(x,y) = p(x)q(x,y)$ as invariant distribution. (Hint: Use first the fact that $X_{k-1} \sim p \implies X_k \sim p$, and then think about a conditional distribution of $Y_{k+1} \ldots$)

(c) Where does the following average converge to:

$$I_{p,q,\mathrm{WR-MH}}^{(n)}(f) := \frac{1}{n}\sum_{k=1}^n \big[\alpha(X_k, Y_{k+1}) f(Y_{k+1}) + \big(1 - \alpha(X_k, Y_{k+1})\big) f(X_k)\,?$$

(d) (*) Can you relate the k:th term in the sum above to the conditional expectation of $f(X_{k+1})$ given X_k and Y_{k+1}? What does this mean regarding the variance of the k:th term? What about the variance of the whole 'waste recycling' estimator $I_{p,q,\mathrm{WR-MH}}^{(n)}(f)$ compared to $I_{p,q,\mathrm{MH}}^{(n)}(f)$?

3.C & 4.C (worth 2 points). Consider a Metropolis-Hastings algorithm targetting the following unnormalised distribution in $\mathbb{R}^2$:

$$p_u([x,y]) = \exp\bigg(-\frac{1}{2}\big(x^2 + y^2\big)\bigg) + \exp\bigg(-\frac{1}{2}\big((x-7)^2 + y^2\big)\bigg),$$

You may use the following Julia function to calculate values of $\log p_u([x,y])$:

```
function log_p_u(x, m1=[0.0,0.0], m2=[7.0,0.0])
  q1 = -.5(x-m1)'*(x-m1); q2 = -.5(x-m2)'*(x-m2)
  q0 = max(q1, q2)  # Avoid underflow of both exp's below
  return(q0 + log(exp(q1-q0) + exp(q2-q0)))
end
```

Use the proposal distribution $q(x,y) = \tilde{q}(y-x)$, where $\tilde{q}$ corresponds to $N(0, 4I_2)$, normal distribution in $\mathbb{R}^2$ with diagonal variance 4.

(a) Implement this Metropolis-Hastings algorithm.

(b) Run the algorithm for 100000 iterations and with suitable burn-in, and calculate estimators for

$$\mathbb{E}_p[f_i(X)] \qquad \text{and} \qquad \mathrm{Var}_p\big(f_i(X)\big),$$

where $f_i(x) := x^{(i)}$, that is, $f_1(x, y) := x$ and $f_2(x, y) := y$.

(c) Calculate and inspect sample autocorrelations for $f_i(X_1), \ldots, f_i(X_{100000})$ (using function `autocor` of `StatsBase` package).

(d) Calculate effective sample size by summing autocorrelations up to some truncation.

(e) Use the functions `estimateBM` and `estimateSV` of the code provided, which give estimates of σ^2_{MH}. Calculate the corresponding effective sample sizes, and compare to your result.

(f) Build up 95% confidence intervals to your estimators, using the estimated variance $\mathrm{Var}_p\big(f_i(X)\big)$ and effective sample sizes.

(g) (* optional) Repeat your experiments 1000 times, and check how many of your confidence intervals for $\mathbb{E}_p[f_i(X)]$ covered the true value (what are the true values?).

5. Let p_a stand for the uniform distribution on the set

$$C_0 \cup C_a, \qquad \text{where} \qquad C_t := \{(x, y) \in \mathbb{R}^2 \,:\, x \in [t, t+1],\, y \in [t, t+1]\}.$$

(a) Draw a picture of the set with $a = 1/2$ and $a = 1$.

(b) Describe a Gibbs sampler for p_a with $a \in [0, 1]$.

(c) Is the Gibbs sampler irreducible with $a = 1$? How about $a \in (0, 1)$?

(Precise proof is not necessary here, but explain your reasoning...)

6. Suppose p_u is an unnormalised p.m.f. on $\mathbb{X}$ and $q(x, \cdot)$ a 'proposal' p.m.f.. Further assume that $h(x, \cdot)$ is a p.m.f. on non-negative integers $\mathbb{N}_0 := \{0, 1, \ldots\}$ defined for each $x \in \mathbb{X}$, with the property

$$\sum_{z \in \mathbb{N}} z h(x, z) = p_u(x) \qquad \text{for each } x \in \mathbb{X}.$$

Consider a Markov chain (X_k, Z_k) defined as follows: Start from $X_0 \in \mathbb{X}$, $Z_0 = 1$ and for $k = 1, 2, \ldots$

- Draw $X_k^* \sim q(X_{k-1}, \cdot)$ and then $Z_k^* \sim h(X_k^*, \cdot)$.
- Draw $U_k \sim U(0, 1)$ and set

$$(X_k, Z_k) = \begin{cases} (X_k^*, Z_k^*), & \text{if } U_k \le \frac{Z_k^*}{Z_{k-1}} \frac{q(X_k^*, X_{k-1})}{q(X_{k-1}, X_k^*)}, \\ (X_{k-1}, Z_{k-1}), & \text{otherwise.} \end{cases}$$

(a) Write down the transition probability

$$K\big((x, z), (x', z')\big) = \mathbb{P}\big((X_n, Z_n) = (x', z') \mid (X_{n-1}, Z_{n-1}) = (x, z)\big)$$

of this Markov chain.
(Hint: Write down first the 'compound proposal' $\tilde{q}\big((x, z), (x^*, z^*)\big) = \mathbb{P}(X_k^* = x^*, Z_k^* = z^* \mid X_{k-1} = x, Z_k = z)$.)

(b) Observe that $h(x, z)z$ is an unnormalised distribution on $\mathbb{X} \times \mathbb{N}_0$ (why?), and show that K is reversible with respect to $\pi(x, z) \propto h(x, z)z$.

(c) Suppose you know that the Markov chain $(X_k, Z_k)_{k \ge 1}$ is irreducible. What can you say about

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k)\,?$$