Exemple sheet for the class "Monte-Carlo inference"

Alexandra Carpentier A.CARPENTIER@STATSLAB.CAM.AC.UK

Editor:

Exercise 1: Explain how to implement a generator for random variables that follow an exponential distribution of parameter λ.

Exercise 2: Let $(f_i)_{i \leq n}$ be a collection of n densities on some domain $\mathcal{X}$ according to a measure μ. Assume that you can sample according to them. Explain how to implement a generator from a mixture density $\sum_{i=1}^k w_i f_i$ where the w_i are positive weights of sum 1.

Exercise 3: Explain how to generate a Bernoulli random variable from a uniform random variable by the method of inversion. Do you think it is efficient? Given a uniform random variable, how many Bernoulli random variable do you think you can generate?

Exercise 4: Let f be a density defined on $[0,1]$ and $L-$Lipschitz, and such that $f > a$ for some constant a.

1. Explain how you would generate a sample from f. What would you do if you do not have access to f but to Cf where C is an unknown constant?
2. Let f_u be the density of f restricted on $[0,u]$. Express it in function of f and u.
3. Explain how you would generate a sample from f_u by rejection sampling, using the best possible envelope knowing L and a.

Exercise 5: Consider the uniform measure μ on $[0,1]$. Consider two functions m and s defined on $[0,1]$, bounded in absolute value by M, and measurable with respect to μ. Assume additionally that $s \geq 0$. You wish to estimate the integral of m, and when you sample at a point x of $[0,1]$, you get a noisy sample

$$Y = m(x) + s(x)\epsilon,$$

where $\epsilon \sim \mathcal{N}(0,1)$. You wish to construct an estimate of the integral of m using n samples collected in this way. Consider a measurable partition $(\Omega_i)_{0 \leq i \leq K-1}$ of $[0,1]$.

1. What is the conditional distribution of sampling uniformly on $[0,1]$? What is the associated mean and variance?
2. What is the conditional distribution of sampling uniformly on Ω_i? What is the associated mean and variance?
3. What is the best allocation you can do if you do not know anything about f? What is the best allocation you can do if you know the variances of sampling in the Ω_i?
4. What happens if the functions m and s are Lipschitz, and if the Ω_i are actually the segments $[\frac{i}{K}, \frac{i+1}{K}]$, when K is very large?

Exercise 6:

- **1.** Explain how to generate a Gaussian distribution $\mathcal{N}(0,1)$ of mean 0 and variance 1 by the Box Muller method. Prove that it works.

Let $X_0 = 0$. We consider the following AR(1) model defined recursively for any $t \geq 0$:

$$X_{t+1} = c + aX_t + \epsilon_t,$$

where the ϵ_t are i.i.d. Gaussian $\mathcal{N}(0,1)$ of mean 0 and variance 1. We observe the chain $X_1, \ldots, X_n$ until some time $n \geq 1$. We would like to estimate (a,c) *using Bayesian inference*.

- **2.** Write the likelihood of the samples $X_1, \ldots, X_n$.
- **3.** In order to apply Bayesian inference to this chain, we set some priors to (a,c). We choose $a \sim \mathcal{N}(0,1)$ and $b \sim \mathcal{N}(0,1)$, with a, b independent of each other. What is the posterior distribution $\pi(a,c)$ of (a,c) knowing $X_1, \ldots, X_n$? What are the marginal distributions $\pi(.|a)$ and $\pi(.|c)$?
- **4.** Explain how you can use Gibbs sampler using these posterior distributions. What is the output of Gibbs sampler? How can you use this output for inferring (a,c)?

Exercise 7 (optional): Consider a distribution $\pi(x_1, x_2)$ defined on $\{1, \dots, M\}^2$ for $M \geq 2$ and such that $\pi > 0$ in any point of its domain. Assume that for any $x_2 \in \{1, \dots, M\}$, you can generate a sample X_1 according to the conditional distribution $\pi(.|x_2)$, and also that for any $x_1 \in \{1, \dots, M\}$, you can generate a sample X_2 according to the conditional distribution $\pi(.|x_1)$.

1. Explain how you would implement Gibbs sampler in this case.
2. Prove that the chain generated in this way has stationary distribution π.

Exercise 8 (optional): Let f be a density that is uniformly continuous according to the uniform measure on $[0, 1]$, and that is bounded by M. Let ϕ be a function defined on $[0, 1]$ such that $|\phi| \leq 1$. Let $\theta = \int_{[0,1]} \phi(x) f(x) dx$.

- **1.** Remind what is importance sampling for estimating θ. What is in this case the optimal distribution that minimises the variance of the importance sampling estimate? We write g^* for this distribution.

Assume that you dispose of n uniform on $[0, 1]$ and i.i.d. samples $U_1, \dots, U_n$.

- **2.** Propose a technique for sampling from f using these uniform samples. What is the expected number of samples from distribution f you obtain with this method? Recall the asymptotic distribution of the proportion associated to this number. Propose a confidence interval for n large enough. How can you use these samples for estimating θ?

- **3.** Propose a technique for sampling from g^* using these uniform samples. What is the expected number of samples from distribution g^* you obtain with this method? Recall the asymptotic distribution of the proportion associated to this number. Propose a confidence interval for n large enough. When proposing your method, you can only use punctual values of ϕ and f, the constant M, and the fact that $|\phi| \leq 1$.

- **4.** Can you use the samples from **3.** for estimating θ?

Exercice 9 (optional): Consider a density f with respect to the uniform measure μ on $[0, 1]$. We assume here that we know some constant $L > 0$ such that f is $L-$Lipshitz, i.e. such that $\forall (x, y) \in [0, 1]^2$, we have

$$|f(x) - f(y)| \leq L|x - y|.$$

We also assume that we know a constant $a > 0$ such that $\forall x \in [0, 1]$, we have $f(x) \geq a$.

- **1.** Explain how to generate with a computer a sample of density f with the rejection sampling method.

- **2.** Let us say that the necessary amount of time to simulate a uniform random variable on a computer is 1. What is the expected time needed for the method of question **1.** for generating one sample?

Let $I > 0$. Consider, for any integer $i \in \{0, \dots, I\}$, and any integer $j \in \{0, ..., 2^j - 1\}$, the interval $\mathcal{I}_{i,j} = [\frac{j}{2^i}, \frac{j+1}{2^i}]$. Let also $f_{i,j}$ be the density associated to the measure $\mu(.|\mathcal{I}_{i,j})$, i.e.

$$f_{i,j}(y) = f(y)\mathbf{1}\{y \in \mathcal{I}_{i,j}\} \times \frac{1}{\int_{\mathcal{I}_{i,j}} f(x)dx},$$

where $\mathbf{1}\{y \in \mathcal{I}_{i,j}\}$ is the indicator function that takes value 1 if $y \in \mathcal{I}_{i,j}$ and 0 otherwise. Let for any $i \in \{1, \dots, I\}$, and any $j \in \{0, ..., 2^j - 1\}$

$$p_{i,j} = \frac{\int_{\mathcal{I}_{i,j}} f(x)dx}{\int_{\mathcal{I}_{i-1,\lfloor j/2 \rfloor}} f(x)dx}.$$

Note that for i, j as above, and j even, it holds that $p_{i,j} + p_{i,j+1} = 1$. Consider the following simulation technique.

Initialize: $j = 0$
for $i = 1, \dots, I$ **do**
  Sample B_i according to a Bernoulli distribution of parameter $p_{i,j+1}$
  Set $j \leftarrow 2j + B_i$
end for
Output: $X \sim f_{I,j} d\mu$

- **3.** Prove that the density of a sample X generated by this algorithm is f.
- **4.** Explain how to generate with a computer a sample of density $f_{i,j}$ with the rejection sampling method. Choose the best possible enveloppe you can, knowing the constants a and L. Can you bound the expected number of uniform samples you will need to use in order to generate one sample from $f_{i,j}$?
- **5.** The necessary amount of time for simulating a Bernoulli random variable of parameter $1/2$ is b (and $b \leq 1$). Can you bound the expected time needed for the algorithm studied in question **2.**, if you simulate according to the $f_{i,j}$ as in question **3.**, for generating one sample? Use this bound to deduce an optimal number of iteration I^* (that minimizes this bound). Compare the computational costs of the procedures of question **1.** and **2.**.