Monte Carlo Inference, Example Sheet 1

The associated example class will be held on **Thursday 25 February, 9-10am, MR13**.

Exercise 1 (Converse of the inverse transform theorem for continuous CDF)**.** Let F be a continuous CDF, and $X \sim F$. Prove that $F(X) \sim \text{Unif}[0,1]$. Generalize the result to functions with countable number of discontinuities, i.e., when F jumps at $x_1, x_2, \ldots$.

Exercise 2. Given a random variable $U \sim \text{Unif}[0,1]$, How can you generate a random variable

(i) $X \sim \text{Ber}(p)$, $p \in (0,1)$.

(ii) $X \in \{1, \ldots, K\}$, with $\mathbb{P}(X = k) = w_k, k = 1, \ldots, K$, where the w_ks are positive and sum to one.

(iii) $X \sim \text{Unif}[a,b]$, where $a < b$.

(iv) $X \sim \text{Cauchy}(0,1)$, where the density of a Cauchy$(0,1)$ is given by $f(x) = \frac{1}{\pi(1+x^2)}, x \in \mathbb{R}$.

(v) $X \sim \text{GEV}(\mu, \sigma, \xi)$, $\mu \in \mathbb{R}, \sigma > 0, \xi \in \mathbb{R} \setminus \{0\}$, and the CDF of the GEV$(\mu, \sigma, \xi)$ distribution is given by

$$F(x) = \exp\left\{ -\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-1/\xi} \right\}, \quad x \in \mathbb{R} \text{ such that } 1 + \xi\frac{x-\mu}{\sigma} > 0.$$

(GEV stands for Generalized Extreme Value).

Exercise 3 (Ratio of uniforms)**.** Let $h(x)$ defined on $\mathbb{R}$ be a non-negative function with finite integral. Let

$$C_h = \left\{ (u,v) \in \mathbb{R}^2 : 0 \le u \le \sqrt{h(v/u)} \right\}.$$

(i) Show that C_h has a finite area by using an appropriate change of variables.

(ii) If $(U,V) \sim \text{Unif}(C_h)$, show that $X = V/U$ has density $h(x)/\int_{\mathbb{R}} h(z)dz$.

(iii) Suppose $h(x)$ and $x^2 h(x)$ are bounded. Devise an algorithm for generating according to h, using the results above.

(iv) Implement this method when

  (a) $h(x) = e^{-x}\mathbf{1}_{[0,\infty)}(x)$,

  (b) $h(x) = \exp(-x^2/2)$.

Exercise 4 (Alternative version of Box Muller)**.** Show that the algorithm

(i) Generate $U_1, U_2 \overset{\text{i.i.d.}}{\sim} \text{Unif}[-1,1]$ until $S = U_1^2 + U_2^2 \le 1$;

(ii) Define $Z = \sqrt{-2\log(S)/S}$ and set $X_1 = ZU_1, X_2 = ZU_2$;

generates $X_1, X_2 \overset{\text{i.i.d.}}{\sim} N(0,1)$.

Exercise 5. Let $\mu \in \mathbb{R}^d$, and Σ be a $d \times d$ non-negative symmetric matrix, i.e. $\Sigma = \Sigma^\top$ and $x^\top \Sigma x \ge 0, \forall x \in \mathbb{R}^d$. How can we generate a multivariate Gaussian random variable $N_d(\mu, \Sigma)$ from i.i.d. univariate $N(0,1)$?

Exercise 6. Suppose that $(Y_1, \ldots, Y_N)$ is a sample produced by an accept-reject method based on (f,g), where f, g are densities on $\mathbb{R}^d$, $f \le Mg$, and $M = \sup(f/g)$. Denote by $(X_1, \ldots, X_t)$ the accepted subsample and by $(Z_1, \ldots, Z_{N-t})$ the rejected subsample.

(i) Show that $\delta_1 = t^{-1}\sum_{i=1}^t h(X_i)$ and

$$\delta_2 = (N-t)^{-1} \sum_{i=1}^{N-t} h(Z_i) \frac{(M-1)f(Z_i)}{Mg(Z_i) - f(Z_i)},$$

are unbiased estimators of $\mu = \mathbb{E}_f(h(X))$ (conditional on $N > t \ge 1$), where $h : \mathbb{R}^d \to \mathbb{R}$ is such that $\mathbb{E}_f |h(X)| < \infty$.

(ii) Show that δ_1, δ_2 are independent.

(iii) Find the weight β^* that minimizes that variance of the aggregated estimator $\delta_3 = \beta\delta_1 + (1-\beta)\delta_2$ (conditional on $N > t \geq 1$).

Exercise 7. Suppose that f is an unnormalized density over $\mathbb{R}^d$ and g is another density over $\mathbb{R}^d$, from which we can simulate easily, and such that $f(x) \leq Mg(x), \forall x \in \mathbb{R}^d$, for some fixed $M \geq 0$.

(i) Show that $M \geq 1$ if f and g are normalized.

(ii) Show that the accept-reject algorithm works even though f is not normalized, and even if we do not know exactly M, provided we know $Mg(x)$ and we can sample from g.

(iii) What is the distribution of N, the number of variables generated with the distribution g until the first acceptance occurs in the accept-reject algorithm?

Exercise 8. We want to generate $N(0,1)$ random variables using the accept-reject method with $g_\lambda(x) = \frac{1}{2}\lambda\exp(-\lambda|x|), x \in \mathbb{R}, \lambda > 0$.

(i) Given $U_1, U_2 \overset{\text{i.i.d.}}{\sim} \text{Unif}[0,1]$, how can you simulate $Y \sim g_\lambda$?

(ii) Find the optimal $\lambda > 0$ (in terms of acceptance probabilities).

(iii) Describe the algorithm.

Exercise 9. You partially saw this example during the lecture. Let

$$p = \int_2^\infty \frac{1}{\pi(1+x^2)}dx. \tag{1}$$

(i) Show that

$$p = \int_0^{1/2} \frac{1}{\pi(1+y^2)}dy, \tag{2}$$

and

$$p = 1/2 - \int_0^2 \frac{1}{\pi(1+x^2)}dx. \tag{3}$$

(ii) Using (1), (2) and (3), construct Monte Carlo estimators $\hat{p}_{i,m}$, of p based on $X_1, \ldots, X_m \overset{\text{i.i.d.}}{\sim} f_i$, where

(a) f_1 is the density of a Cauchy$(0,1)$,
(b) f_2 is the density of a Unif$[0,1/2]$,
(c) f_3 is the density of a Unif$[0,2]$,

and compute their variances.

(iii) Someone uses the estimator $\hat{p}_{2,m}$ to estimate p, with sample size $m = 50$. You are using the estimator $\hat{p}_{1,m}$. What sample size m should you use to achieve the same accuracy (i.e. the same width of the confidence interval for a fixed level α). What if you are using $\hat{p}_{3,m}$?

Exercise 10. 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 $\theta = \int_{[0,1]} \phi(x)f(x)dx$.

(i) 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.

(ii) Propose a technique for sampling from f given n i.i.d. samples $U_1, \ldots, U_n \sim \text{Unif}[0,1]$. 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?

(iii) 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 when n is large enough. When proposing your method, you can only use punctual values of ϕ and f, the constant M, and the fact that $|\phi| \leq 1$.

(iv) Can you use the samples from (iii) for estimating θ?

Exercise 11. Let $h_n^{(s)}$ denote the Monte Carlo approximation of $\int h(x)f(x)dx$ by stratified sampling of T_i points from strata Ω_i with weights w_i, $i = \{1, \dots, k\}$.

(i) Prove that $\mathrm{Var}[h_n^{(s)}] = \sum_{i=1}^{K} \frac{w_i^2 \sigma_i^2}{T_i}$, where

$$\sigma_i^2 = \frac{1}{w_i^2} \int_{\Omega_i} (h(x) - \mu_i)^2 f(x)dx; \text{ and}$$

$$\mu_i = \frac{1}{w_i^2} \int_{\Omega_i} h(x)f(x)dx.$$

(ii) For the unform T_i, i.e. for $T_i = w_i n$, $n = \sum_{i=1}^{K} T_i$, show that $\mathrm{Var}[h_n^{(s)}] < \mathrm{Var}[h_n]$, where h_n is the classical Monte Carlo approximation of $\int h(x)f(x)dx$.