Sequential importance sampling

1. The following exercise is an illustration of the weight degeneracy for the sequential importance sampling (SIS) algorithm. For simplicity, we consider the (admittedly artificial) case where we sample, using SIS, from a sequence $(f_n)_{n\geq 0}$ of densities

$$f_n(x_{0:n}) = \prod_{k=0}^{n} f(x_k), \quad x_{0:n} \in \mathbb{R}^{n+1},$$

where

$$f(x) = \frac{z(x)}{c}, \quad x \in \mathbb{R},$$

is a density that is known up to the normalizing constant c only. For this purpose, we use the proposal density

$$g_n(x_{n+1} \mid x_{0:n}) = g(x_{n+1}), \quad x_{0:n} \in \mathbb{R}^{n+1},$$

where the density g on $\mathbb{R}$ is such that $g(x) = 0 \Rightarrow f(x) = 0$ and $\mathbb{E}_g(\varrho^2(X)) < \infty$, with $\varrho(x) = z(x)/g(x)$. The algorithm is initialized with $g_0 = g$.

(a) Describe, using a pseudo-code, how the SIS algorithm updates a particle sample $(X^i_{0:n}, \omega^i_n)_{i=1}^N$ approximating f_n to another sample $(X^i_{0:n+1}, \omega^i_{n+1})_{i=1}^N$ approximating f_{n+1} in this case.

(b) Let ϕ be a bounded objective function on $\mathbb{R}$ and consider the estimator

$$\sum_{i=1}^{N} \frac{\omega^i_n}{\sum_{\ell=1}^N \omega^\ell_n} \phi(X^i_n)$$

of $\tau = \mathbb{E}_{f_n}(\phi(X_n)) = \mathbb{E}_f(\phi(X))$. (This is of course a rather convoluted and very inefficient way of constructing an estimate of τ but still constitutes a valid instance of the SIS approach in a very particular case.) Establish the CLT

$$\sqrt{N}\left(\sum_{i=1}^{N} \frac{\omega^i_n}{\sum_{\ell=1}^N \omega^\ell_n} \phi(X^i_n) - \tau\right) \xrightarrow{\text{d.}} N(0, \sigma^2_n(\phi)) \quad \text{as } N \to \infty,$$

where

$$\sigma^2_n(\phi) = \frac{\mathbb{E}(\varrho^2(X))^n \, \mathbb{E}(\varrho^2(X)\{\phi(X) - \tau\}^2)}{c^{2(n+1)}}.$$

(c) Show, using Jensen's inequality[1], that $\sigma^2_n(\phi)$ tends to infinity geometrically fast with n as soon as g is different from f on a set with positive probability.

2. Recall that the *coefficient of variation* of a set of importance weights $(\omega^i_n)_{i=1}^N$ is given by

$$\mathrm{CV}_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(N\frac{\omega^i_n}{\sum_{\ell=1}^N \omega^\ell_n} - 1\right)^2}.$$

Check that the value of CV_N is $\sqrt{N-1}$ in the case of maximally skew importance weights, i.e., when all weights are zero but a single one.

[1] Recall that Jensen's inequality states that if X is a random variable and φ is a convex function, then $\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$. If φ is strictly convex, the inequality is strict unless X is a constant with probability one.