Exercise sheet 6
Solutions

Exercise 1. i) If the target distribution $\pi(\cdot)$ has an unknown normalising constant, that is $\pi(\cdot)$ is of the form

$$\pi(\theta) = \frac{\gamma(\theta)}{Z}$$

with Z unknown, then one can instead consider the self-normalised importance sampling algorithm which gives an estimator of $I(\varphi)$ of the form

$$\tilde{\boldsymbol{I}}(\varphi) = \frac{1}{\sum_{j=1}^{N} \tilde{w}(\boldsymbol{\theta}^{(j)})} \sum_{i=1}^{N} \varphi(\boldsymbol{\theta}^{(i)}) \tilde{w}(\boldsymbol{\theta}^{(i)})$$

with $\tilde{w}(\theta) = \gamma(\theta)/s(\theta)$

ii) An unbiased estimator $\boldsymbol{Z}$ for Z is

$$\boldsymbol{Z} = \frac{1}{N} \sum_{i=1}^{N} \tilde{w}(\boldsymbol{\theta}^{(i)}).$$

This estimator is indeed unbiased since

$$\mathbb{E}(\boldsymbol{Z}) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}(\tilde{w}(\boldsymbol{\theta}^{(i)})) = \int \frac{\gamma(\theta)}{s(\theta)} s(\theta) \, \mathrm{d}\theta = Z.$$

iii) The obtained estimator of $I(\varphi)$ is not unbiased because

$$\mathbb{E}(\tilde{\boldsymbol{I}}(\varphi)) = \mathbb{E}\left(\frac{N^{-1} \sum_{i=1}^{n} \varphi(\boldsymbol{\theta}^{(i)}) \tilde{w}(\boldsymbol{\theta}^{(i)})}{N^{-1} \sum_{i=1}^{N} \tilde{w}(\boldsymbol{\theta}^{(i)})} \right) \neq \frac{\mathbb{E}(N^{-1} \sum_{i=1}^{n} \varphi(\boldsymbol{\theta}^{(i)}) \tilde{w}(\boldsymbol{\theta}^{(i)}))}{\mathbb{E}(N^{-1} \sum_{i=1}^{N} \tilde{w}(\boldsymbol{\theta}^{(i)}))} = I(\varphi).$$

In general, estimators define as the quotient of two estimators are often biased.

Exercise 2. i) Z_n is a high-dimensional integral in general and is difficult to compute directly with quadrature or basic Monte Carlo methods.

ii) Given that it is possible to sample from $p_0(\cdot)$ and $q_k(\cdot \,|\, \theta)$ for any $\theta \in \Theta$, the simplest proposal distribution that one can consider is

$$s_n(\theta_{0:n}) = p_0(\theta_0) \prod_{k=1}^{n} q_k(\theta_k \,|\, \theta_{k-1}). \tag{1}$$

iii) To sample from $s_n(\cdot)$ one can fist sample $\boldsymbol{\theta}_0$ from $p_0(\cdot)$, then for all $k \in \{1, \ldots, n\}$, sample $\boldsymbol{\theta}_k$ from $q_k(\cdot \,|\, \boldsymbol{\theta}_{k-1})$.

iv) Because of the special form of the proposal distribution, it holds that

$$\tilde{w}(\boldsymbol{\theta}_{0:n}) = \frac{\gamma_n(\boldsymbol{\theta}_{0:n})}{s_n(\boldsymbol{\theta}_{0:n})} = \ell_0(y_0 \,|\, \boldsymbol{\theta}_0) \prod_{k=1}^{n} \ell_k(y_k \,|\, \boldsymbol{\theta}_k)$$

v) The importance sampling algorithm for this choice of proposal distribution is given in Algorithm 1.

Algorithm 1 Sequential importance sampling for the proposal (1)

for $i = 1, \dots, N$ **do**

  Sample $\boldsymbol{\theta}_0^{(i)} \sim p_0(\cdot)$

  Define the importance weight

$$\boldsymbol{w}_0^{(i)} = \ell_0(y_0 \,|\, \boldsymbol{\theta}_0^{(i)})$$

end for

for $k = 1, \dots, n$ **do**

  for $i = 1, \dots, N$ **do**

    Sample $\boldsymbol{\theta}_k^{(i)} \,|\, \boldsymbol{\theta}_{k-1}^{(i)} \sim q_k(\cdot \,|\, \boldsymbol{\theta}_{k-1}^{(i)})$

    Define the importance weight

$$\boldsymbol{w}_k^{(i)} = \boldsymbol{w}_{k-1}^{(i)} \ell_k(y_k \,|\, \boldsymbol{\theta}_k^{(i)})$$

  end for

end for

Output:

$$\hat{\boldsymbol{I}}_n(\varphi) = \frac{1}{\sum_{j=1}^N \boldsymbol{w}_n^{(j)}} \sum_{i=1}^N \boldsymbol{w}_n^{(i)} \varphi(\boldsymbol{\theta}_n^{(i)})$$

and $\hat{\boldsymbol{Z}}_n = \frac{1}{N} \sum_{i=1}^N \boldsymbol{w}_n^{(i)}$