In the name of GOD.
Sharif University of Technology
Stochastic Processes CE 695 Fall 2024 Hamid R. Rabiee

Homework 6 Solution (Sampling Methods)

1. (a) Inverse Transform Sampling requires computing the inverse CDF (F^{-1}), which often lacks analytical form and is computationally expensive, especially for multivariate distributions.

   (b) A good proposal distribution in Importance Sampling is essential since efficiency depends on its similarity to the target distribution.

   poor choices cause high variance in weights and inadequate sampling of important regions.

   (c) Rejection sampling becomes inefficient when acceptance rates are low, particularly in high dimensions, when target and proposal distributions differ significantly.

   (d) The burn-in period in MCMC involves discarding initial samples to allow chain convergence to the stationary distribution, with its length depending on the starting point and mixing rate.

2. (a)

$$F_X(x) = \int_0^x f_X(t)dt = \int_0^x \frac{e^t}{e-1}dt = \left[\frac{e^t}{e-1}\right]_0^x = \frac{e^x - 1}{e-1}$$

$$\Rightarrow F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{e^x-1}{e-1} & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

   (b) To find the inverse CDF, solve $F_X(x) = y$ for x:

$$y = \frac{e^x - 1}{e-1}$$
$$\Rightarrow \ln[y(e-1)+1] = x$$

$$\Rightarrow F_X^{-1}(y) = \begin{cases} 0 & y < 0 \\ \ln[y(e-1)+1] & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

(c)

$$\begin{aligned}
X &= F_X^{-1}(U) \\
&= \ln\left[\frac{1}{\sqrt{e}+1}(e-1)+1\right] \\
&= \ln\left[\frac{(e-1)+(\sqrt{e}+1)}{\sqrt{e}+1}\right] \\
&= \ln\left[\frac{e+\sqrt{e}}{\sqrt{e}+1}\right] \\
&= \ln(\sqrt{e}) \\
&= \frac{1}{2}
\end{aligned}$$

3. (a) Deriving the importance sampling estimator:
   The expected value can be written as:

$$E_X[X] = \int xf(x)dx = \int x\frac{f(x)}{q(x)}q(x)dx = E_q[Xw(x)]$$

where

$$w(x) = \frac{f(x)}{q(x)}$$

is the importance weight
Given:

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x} = \frac{1}{\Gamma(2)}x^1e^{-x}$$

(since $\alpha = 2, \lambda = 1$)

$$q(x) = \lambda_0 e^{-\lambda_0 x} = 0.5e^{-0.5x}$$

Therefore:

$$w(x) = \frac{f(x)}{q(x)} = \frac{xe^{-x}}{0.5e^{-0.5x}\Gamma(2)} = \frac{2x}{e^{0.5x}}$$

The importance sampling estimator is:

$$\hat{I} = \frac{1}{n}\sum_{i=1}^{n} x_i w(x_i) = \frac{1}{n}\sum_{i=1}^{n}\frac{2x_i^2}{e^{0.5x_i}}$$

   (b) For $x_1 = 2$: $w(2) = \frac{2(2)}{e^{0.5(2)}} = \frac{4}{e^1} \approx 1.472$
   For $x_2 = 2$: $w(2) = \frac{2(2)}{e^{0.5(2)}} = \frac{4}{e^1} \approx 1.472$
   For $x_3 = 3$: $w(3) = \frac{2(3)}{e^{0.5(3)}} = \frac{6}{e^{1.5}} \approx 1.337$
   (b) Estimate of I:

$$\begin{aligned}\hat{I} &= \frac{1}{3}(2 \cdot 1.472 + 2 \cdot 1.472 + 3 \cdot 1.337)\\ &= \frac{1}{3}(2.944 + 2.944 + 4.011)\\ &= \frac{9.899}{3} \approx 3.3\end{aligned}$$

The selected samples (2, 2, 3) do not well represent our proposal distribution and we need more samples to have a better estimate of that.

4. (a) The standard normal PDF is: $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$
   Therefore, the truncated PDF is (due to symmetry):

$$f_X(x) = \begin{cases} \frac{2}{\sqrt{2\pi}}e^{-x^2/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

   (b)

$$\frac{f_X(x)}{g(x)} = \frac{2}{\sqrt{2\pi}}e^{-x^2/2} \cdot \frac{1}{2e^{-2x}} = \frac{1}{\sqrt{2\pi}}e^{-x^2/2+2x}$$

   Differentiate with respect to x and set to 0:

$$\frac{d}{dx}(-\frac{x^2}{2} + 2x) = -x + 2 = 0$$

   The maximum value occurs at $x = 2$:

$$M = \frac{1}{\sqrt{2\pi}}e^{-2^2/2+2\cdot 2} = \frac{1}{\sqrt{2\pi}}e^{-2+4} = \frac{1}{\sqrt{2\pi}}e^2$$

   Therefore, the smallest M that satisfies the inequality for all $x \geq 0$ is $\frac{e^2}{\sqrt{2\pi}}$

   (c) Generate proposal X from exponential($\lambda = 2$):

$$\begin{aligned} u_1 &= 1 - e^{-\lambda x}\\ \Rightarrow x &= -\frac{1}{\lambda}\ln(1 - u_1)\\ \Rightarrow X &= -\frac{1}{2}\ln(U_1) \text{ where } U_1 \sim \text{Uniform}(0,1)\end{aligned}$$

   Generate U_2 ~Uniform$(0, 1)$
   Accept X if:

$$U_2 \leq \frac{f_X(X)}{Mg(X)} = \frac{\frac{2}{\sqrt{2\pi}}e^{-X^2/2}}{(\frac{e^2}{\sqrt{2\pi}})(2e^{-2X})}$$

$$= \frac{e^{-X^2/2}}{e^2 \cdot e^{-2X}}$$
$$= e^{-X^2/2+2X-2}$$

$$\Rightarrow U_2 \leq e^{(\frac{1}{2}\ln(U_1))^2/2-\ln(U_1)-2} = e^{\ln(U_1)^2/8-\ln(U_1)-2} = \frac{e^{\ln(U_1)^2/8-2}}{U_1}$$

5. (a) If 95% of the weight is on only 3 particles, the particle filter is nearly degenerate. Most particles do not contribute meaningfully, reducing diversity and increasing the risk of tracking failure.

   (b) The Selection (resampling) step duplicates high-weight particles and discards low-weight ones. This redistributes the weight more evenly, restoring diversity and mitigating degeneracy.

   (c) Use the Effective Sample Size (ESS):

   $$\text{ESS} = \frac{1}{\sum(w_i^2)}$$

   If ESS falls below a threshold (e.g., half of the number of particles), perform resampling.

6. (a) The likelihood: $L(\beta_0, \beta_1|\{X_i, Y_i\}) \propto \exp(-\frac{1}{2\sigma^2}\sum_{i=1}^n(Y_i-\beta_0-\beta_1 X_i)^2)$
   Prior for β_0: $p(\beta_0) \propto \exp(-\frac{\beta_0^2}{2\tau^2})$

   $$\Rightarrow p(\beta_0|\beta_1, \{X_i, Y_i\}) \propto \exp(-\frac{1}{2\sigma^2}\sum_{i=1}^n(Y_i - \beta_0 - \beta_1 X_i)^2)\exp(-\frac{\beta_0^2}{2\tau^2})$$

   $$\sum_{i=1}^n(Y_i - \beta_0 - \beta_1 X_i)^2 = \sum_{i=1}^n(Y_i - \beta_1 X_i)^2 - 2\beta_0\sum_{i=1}^n(Y_i - \beta_1 X_i) + n\beta_0^2$$

   $$\Rightarrow p(\beta_0|\beta_1, \{X_i, Y_i\}) \propto \exp(-\frac{1}{2}[(\frac{n}{\sigma^2} + \frac{1}{\tau^2})\beta_0^2 - \frac{2}{\sigma^2}\sum_{i=1}^n(Y_i - \beta_1 X_i)\beta_0])$$

   $$\beta_0|\beta_1, \{X_i, Y_i\} \sim \mathcal{N}(\frac{\sum_{i=1}^n(Y_i - \beta_1 X_i)}{n + \sigma^2/\tau^2}, \frac{\sigma^2}{n + \sigma^2/\tau^2})$$

   (b) The likelihood: $L(\beta_0, \beta_1|\{X_i, Y_i\}) \propto \exp(-\frac{1}{2\sigma^2}\sum_{i=1}^n(Y_i-\beta_0-\beta_1 X_i)^2)$
   Prior for β_1: $p(\beta_1) \propto \exp(-\frac{\beta_1^2}{2\tau^2})$

$$\Rightarrow p(\beta_1|\beta_0,\{X_i,Y_i\}) \propto \exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1X_i)^2)\exp(-\frac{\beta_1^2}{2\tau^2})$$

$$\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1X_i)^2 = \sum_{i=1}^{n}(Y_i-\beta_0)^2 - 2\beta_1\sum_{i=1}^{n}(Y_i-\beta_0)X_i + \beta_1^2\sum_{i=1}^{n}X_i^2$$

$$\Rightarrow p(\beta_1|\beta_0,\{X_i,Y_i\}) \propto \exp(-\frac{1}{2}[(\frac{\sum_{i=1}^{n}X_i^2}{\sigma^2}+\frac{1}{\tau^2})\beta_1^2-\frac{2}{\sigma^2}\sum_{i=1}^{n}(Y_i-\beta_0)X_i\beta_1])$$

$$\beta_1|\beta_0,\{X_i,Y_i\} \sim \mathcal{N}(\frac{\sum_{i=1}^{n}(Y_i-\beta_0)X_i}{\sum_{i=1}^{n}X_i^2+\sigma^2/\tau^2}, \frac{\sigma^2}{\sum_{i=1}^{n}X_i^2+\sigma^2/\tau^2})$$