Importance Sampling

National University of Singapore
A. Thiery

1 Variance of the Importance Sampling Estimator

Consider a density $p(x)$, a proposal density $q(x)$ and a function $\varphi(x)$ of interest.

1. What is the importance sampling estimator of $\mathrm{E}[\varphi(X)]$?
2. Prove that the variance of this importance sampling estimator is given by

$$\frac{1}{N}\,\mathrm{Var}\left(\varphi(X)\,r(X)\right)$$

where $r(x) = p(x)/q(x)$ and $X \sim q$.

2 Optimal Proposal Density

Consider a density $p(x)$ and a function $\varphi(x)$ of interest.

1. Prove that the optimal proposal density $q^*(x)$ that minimizes the variance of the importance sampling estimator is given by

$$q^*(x) = \frac{1}{Z}|\varphi(x)|\,p(x)$$

where $Z = \int |\varphi(x)|\,p(x)\,dx$. *(Hint: Use the Cauchy-Schwarz inequality.)*

2. What is the variance of the importance sampling estimator when using the optimal proposal density when the function $\varphi(x)$ is non-negative?
3. Why is this proposal density not practical in most cases?

3 Ratio of normalizing constants

Consider two densities $p(x) = p_u(x)/Z_p$ and $q(x) = q_u(x)/Z_q$ where $p_u(x)$ and $q_u(x)$ are unnormalized densities and Z_p and Z_q are the normalizing constants. Prove that for $X \sim q$ we have:

$$\frac{Z_p}{Z_q} = \mathrm{E}_q\left[\frac{p_u(X)}{q_u(X)}\right].$$

4 Multivariate Gaussian

Consider a Gaussian distribution $p(x, y)$ in $\mathbb{R}^2$ with mean zero and covariance matrix C given by $C_{1,1} = C_{2,2} = 1$ and $C_{1,2} = C_{2,1} = 0.9$. Estimate as accurately as possible the probability that a sample (X, Y) from this Gaussian distribution is such that $X > 4$ and $Y > 4$.

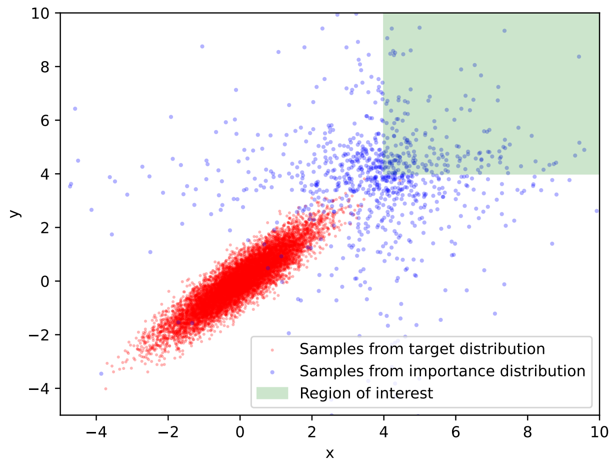

Figure 1: Importance sampling with a Cauchy proposal density

5 Effective sample size

Consider a density $p(x)$ and a proposal density $q(x)$. Recall that the effective sample size is defined as

$$\text{ESS} = \frac{1}{\sum_{i=1}^{N} w_i^2}$$

where the w_i are the normalized importance weights.

1. Recall the definition of the normalized importance weights.
2. Prove that $1 \leq \text{ESS} \leq N$.

6 Variance of a Student-t distribution

Consider $p(x)$ the density of a Student-t distribution with $d = 3$ degrees of freedom,

$$p(x) \;\propto\; \frac{1}{(1 + x^2/3)^2}.$$

1. Use importance sampling with a Gaussian proposal density with mean zero and variance 1 to estimate the variance of $p(x)$.

2. Use importance sampling with a Cauchy distribution as proposal density to estimate the variance of $p(x)$.

7 Enumerating random walks

Estimate the number of paths $(x_0, x_1, \ldots, x_{30})$ where $x_0 = 0$ and $|x_{i+1} - x_i| = 1$ and

$$\max_{0 \le i \le 30} x_i \ge 20.$$

8 Project Evaluation and Review Technique (PERT)

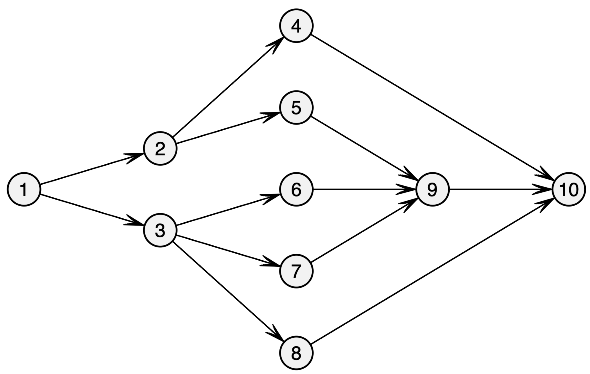

Figure 2: Project Evaluation and Review Technique (PERT) dependencies

Figure 2 shows the dependencies between the tasks of a project. For example, task 9 can only start after tasks 5 and 6 and 7 are completed. Similarly, task 4 can only start once task 2 is completed. The duration T_i of task i is exponentially distributed with mean one,

$$\mathrm{P}(a < T_i < b) = \int_a^b e^{-x}\, dx.$$

For modeling purposes, we assume that the tasks are independent. The total duration of the project is the earliest time at which all tasks are completed, respecting the dependency constraints shown in Figure 2 (i.e., the length of the critical path through the dependency graph). There is a very large penalty if the project is not completed within 17 units of time.
Estimate as accurately as possible the probability that the project runs late. For this purpose, use importance sampling with a suitable proposal density.