LV 11.4500 – UBUNG 5

U5.1 Consider the following of data on relative frequency of number of teeth in unicorns:

$$\pi_T = [0.0001,\ 0.0001,\ 0.0001,\ 0.0015,\ 0.0078,\ 0.0282,\ 0.0844,\ 0.1655,\ 0.2188,\ 0.2371,\ 0.1719,\ 0.0715,\ 0.0130].$$

That is, the ratio $\pi_T(0) = 0.0001$ of all unicorns have 0 teeth, the ratio $\pi_T(1) = 0.0001$ of all unicorns have 1 tooth, etc, for $k = 0, 1, \dots, 12$. The purpose of this exercise is to computationally find the distribution in the set

$$\mathcal{A} = \{\pi = Binom(12, p) \mid p \in (0, 1)\},$$

that best fits π_T in the sense of minimizing the K-L divergence from the said distribution to π_T. That is, to find

$$\pi = \arg\min_{\tilde{\pi} \in \mathcal{A}} d_{KL}(\pi_T || \tilde{\pi}).$$

Implement a computer program for computing

$$f(p) = d_{KL}(\pi_T || \pi_p)$$

on a mesh of values $p \in (0, 1)$, where $\pi_p = Binom(12, p)$. Plot f to approximately find the best value $p \in (0, 1)$.

Hint: First derive the values of $\pi_p(k)$ for $k = 0, 1, \dots, 12$ for given p.

U5.2 Compute the Kullback Leibler divergence from $\pi_Y(x) = \mathbb{1}_{[0,\infty)}(x)\beta \exp(-\beta x)$ to $\pi_X(x) = \mathbb{1}_{[0,\infty)}(x)\lambda \exp(-\lambda x)$ for some $\lambda, \beta > 0$, using the convention $\log(0) \cdot 0 = 0$. That is, i.e., compute $d_{KL}(\pi_X || \pi_Y)$.

U5.3 The accept reject sampling algorithm from lecture 12 is summarized here:

Problem setting: **Target pdf** π that we are unable to sample directly from.

Accept reject algorithm: Assume that we a **proposal density** $\hat{\pi}$ which we can draw samples from, and that for some $N \geq 1$, it holds that $N\hat{\pi} \geq \pi$.

Sample $X \sim \pi$ as follows:

1. sample $Y \sim \hat{\pi}$ and $U \sim U[0, 1]$ with $U \perp Y$.
2. accept $X = Y$ with **acceptance probability** $U \leq \pi(Y)/(N\hat{\pi}(Y))$; otherwise return to step 1.

a) verify that $X \sim \pi$.

Hint:

$$\pi_X(x) = \frac{\mathbb{P}(Y \in dx \mid U \leq \pi(Y)/(N\hat{\pi}(Y)))}{dx}$$

b) Determine which of the following candidates for proposals that can be used to sample the target

$$\pi(x) = \mathbb{1}_{(-\infty,0]}(x)\frac{\sqrt{2}\exp(-x^2/2)}{\sqrt{\pi}}$$

with the accept reject algorithm:

i $\hat{\pi}_1(x) = \exp(-|x|/2)/4$

ii $\hat{\pi}_2(x) = \frac{\exp(-x^2)}{\sqrt{\pi}}$

iii $\hat{\pi}_3(x) = \mathbb{1}_{(0,1)}(x)$

and provide N.

c) The pdf of a Weibull(λ, k) distribution is defined by

$$\pi(x) = \mathbb{1}_{[0,\infty)}\frac{k}{\lambda}\Big(\frac{x}{\lambda}\Big)^{k-1} e^{-(x/\lambda)^k}, \quad k, \lambda > 0.$$

Use the Monte Carlo and the accept reject sampling algorithm to estimate $\mathbb{E}\left[X^2\right]$ for $X \sim$ Weibull$(2, 1.2)$.

U5.4 Consider the Metropolis Hastings algorithm presented in Lecture 12 with target pdf π, conditional proposal $q(y|x)$ and acceptance probability

$$\rho(x,y) = \min\left(\frac{\pi(y)}{\pi(x)}\frac{q(x|y)}{q(y|x)}, 1\right)$$

a) Verify that for any $A \in \mathcal{B}^d$,

$$K(x,A) = \underbrace{\int_A \rho(x,y)q(y|x)dy}_{r(x,A)} + \Big(1 - r(x,\mathbb{R}^d)\Big)\delta_x(A)$$

Hint:

$$\mathbb{P}(X_1 \in A \mid X_0 = x) = \mathbb{P}(Y_0 \in A, X_1 = Y_0 \mid X_0 = x) + \mathbb{P}(x \in A, X_1 = x \mid X_0 = x) = \dots$$

b) Assuming $q(\cdot|x)$ dominates π for all $x \in \mathbb{R}^d$, prove that M-H kernel satisfies detailed balance wrt π:

$$\int_A K(x,B)\pi(x)dx = \int_B K(x,A)\pi(x)dx \qquad \forall A, B \in \mathcal{B}^d, \tag{1}$$

c) Verify that under the assumption in b), π is an invariant pdf of the M-H Markov chain.

d) **Updated question.** If $\pi \propto \exp(-x^2/2)$ and $q(y|x) = \mathbb{1}_{(0,1)}(y)$ for all $x \in \mathbb{R}$, then it turns out that (1) **still does** hold, even when the constraint that $q(\cdot|x)$ dominates π is not fulfilled (can be verified using the division-by-zero convention together with 5.4 c)). However, lack of domination may lead to loss of weak convergence of the chain: assuming the initial condition of an MCMC chain is given by $X_0 \sim \mathbb{P}_0$ with

$$\mathbb{P}_0((-\infty, 0]) \neq \int_{-\infty}^0 \pi(x)dx = 1/2,$$

explain why the above proposal q is not will not yield convergence $\mathbb{P}_n \Rightarrow \mathbb{P}$, with $\mathbb{P}$ denoting the measure associated to π.

U5.5 Let $A = \{x \in \mathbb{R}^2 \mid 2x_1^2 + 5x_2^2 \in (1, 1.2)\}$ and let $\pi(x) \propto \mathbb{1}_A(x)\exp(-|x|^{1.9})$. Construct an MCMC method for sampling π and estimate

$$\mathbb{E}^{\pi}[\exp(-2|x_1| - |x_2|)]$$

using 10000 samples in your chain.