Extra assignment 0: Who's that (simulated) pokémon?

Computational Statistics
Instructor: Luiz Max Carvalho

September 21, 2022

Hand-in date: 28/09/2022.

General guidance

- State and prove all non-trivial mathematical results necessary to substantiate your arguments;
- Do not forget to add appropriate scholarly references *at the end* of the document;
- Mathematical expressions also receive punctuation;
- Please hand in a single PDF file as your final main document.

  Code appendices are welcome, *in addition* to the main PDF document.

Background

In this (hopefully) fun little exercise I will describe a rejection sampling algorithm to sample from a mysterious distribution. Suppose we have the following procedure:

1. Generate $U_1, U_2 \sim \text{Uniform}(0, 1)$, independently;
2. Compute $Y_1 = -\log(U1)$ and $Y_2 = -\log(U2)$. If $Y_2 > \frac{(1-Y_1)^2}{2}$, accept $Y = (Y_1, Y_2)$. Else, reject and return to step 1;
3. Generate $U_3 \sim \text{Uniform}(0, 1)$; if $U_3 < 1/2$, set $X = Y_1$, otherwise set $X = -Y_1$.

Your job is to analyse this algorithm mathematically, find out its target distribution and work out its acceptance rate.

Questions

1. What is the distribution of Y_1 and Y_2?
2. What is the distribution of the "mystery" random variate X?
3. How can one take the output of the algorithm (X) and generate $W \sim \text{Normal}(\mu, \sigma^2)$, with $\mu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}_+$?
4. Can you work out what the acceptance rate of this algorithm is?
5. (bonus) Can you generalise this algorithm to sample from other distributions? For example, how would you modify the algorithm in order to sample from a Gamma distribution with parameters $a, b \in \mathbb{R}_+$?

   Hint: consider modifying step 2) to accept when $Y_2 > f(Y_1)$ and choose f carefully.