## Assignment 3: Generating Random Variables (Part I)

*Due on Feb. 7*

1. A pair of dice are to be continually rolled until all the possible outcomes $2, 3, \ldots, 12$ have occurred at least once. Develop a Monte-Carlo simulation experiment to estimate the expected number of dice rolls that are needed.

2. Prove that the acceptance-rejection algorithm for discrete r.v.'s (given on p.32 of notes) is correct. (Mimic the proof for the acceptance-rejection method for continuous r.v.'s on p.25–27). *Hint: Try to prove that the PMF* $\mathbb{P}(X = x_i) = \cdots\cdots = p(x_i)$ *for the r.v.* $X$ *which is the output of the algorithm on p.32.*

3. For I.I.D. uniform(0,1) r.v.'s $U_1, U_2, \ldots$. Define the following two random variables

$$N \equiv \min\left\{n : \sum_{i=1}^{n} U_i > 1\right\} \quad \text{and} \quad M \equiv \left\{n \geq 0 : \text{such that } \prod_{i=1}^{n} U_i \geq e^{-\lambda} > \prod_{i=1}^{n+1} U_i\right\}.$$

   (a) Compute $\mathbb{E}[N]$ by generating $n = 100, 1000, 10000$ values of $N$. What do you is $N$?

   (b) Compute $\mathbb{E}[M]$ by generating $n = 100, 1000, 10000$ values of $M$ with $\lambda = 1$ and 2. What do you think is $N$?

4. Let $X$ be a discrete r.v. with state space $\{1, \ldots, 6\}$ and PMFs

$$p(1) = 0.05,\ p(2) = 0.05,\ p(3) = 0.1,\ p(4) = 0.1,\ p(5) = 0.6,\ p(6) = 0.1.$$

   Define cumulative sum of the PMF: $q(i) \equiv \sum_{k=1}^{i} p(i)$, $i = 1, \ldots, 6$.

   (a) Explain why the algorithm below is exactly a *discrete inverse transform* method with a simple *left-to-right* search.

   STEP 1: Generate $U \sim$ uniform(0,1) and set $i = 1$.

   STEP 2: If $U \leq q(i)$, stop and return $X = i$. Otherwise, continue to STEP 3.

   STEP 3: Let $i = i + 1$, go to STEP 2.

   (b) Let $N$ be the number of times STEP 2 is executed (e.g., number of comparisons until a final acceptance). Note that $N$ measures the efficiency of your algorithm. What is the exact value of $\mathbb{E}[N]$?

   (c) Implement the above algorithm to generate $X$ for $n = 10000$ times and use Monte-Carlo simulation to estimate $\mathbb{E}[N]$ and compare to (b).

   (d) Alternatively, we first sort the $p(i)$'s in decreasing order and form a sorted version of $q$:

$$q'(1) = 0.6,\ q'(2) = 0.7,\ q'(3) = 0.8,\ q'(4) = 0.9,\ q'(5) = 0.95,\ q'(6) = 1,$$

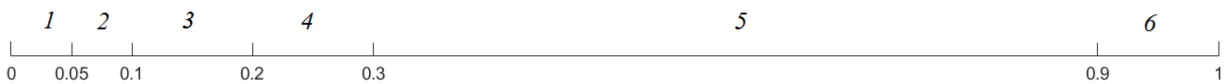
   and an identity vector

$$id(1) = 5,\ id(2) = 3),\ id(3) = 4,\ id(4) = 6,\ id(5) = 1,\ id(6) = 2.$$

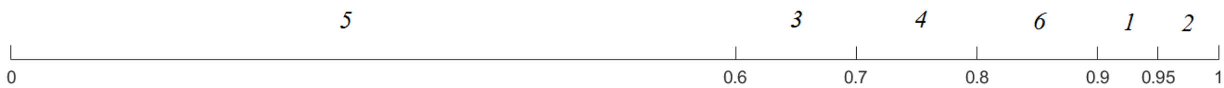
   Explain why the following algorithm is valid: (see the figure)

   STEP 1: Generate $U \sim$ uniform(0,1) and set $i = 1$.

   STEP 2: If $U \leq q'(i)$, stop and return $X = id(i)$. Otherwise, continue to STEP 3.




*(a) Discrete inverse transform with simple left-to-right search*




*(b) Discrete inverse transform with states sorted according their probability masses*

STEP 3: Let $i = i + 1$, go to STEP 2.

(e) Let $N'$ be the number of comparisons in STEP 2 in (d). What is the exact value of $\mathbb{E}[N']$? Use Monte-Carlo simulation to estimate $\mathbb{E}[N']$ and compare to the exact value.

Remarks: You should observe that $\mathbb{E}[N'] < \mathbb{E}[N]$. Note that this saving in the marginal execution time depends on the particular distribution and must be weighed against the extra setup time and storage for the identity vector $id(i)$.

5. Consider the piecewise linear majorizing function $g(x)$ (on p.31 of the notes) for the Beta distribution (p.28 of the notes), with $x_1 = 0.36$ and $x_2 = 0.84$.

   (a) Compute the constant $c$ and the probability of acceptance.

   (b) Give an algorithm to generate $Y \sim h(x)$.

   (c) Give an acceptance-rejection method to generate the Beta distributed $X \sim f(x)$ (p.26) using this new majorizing function $g(x)$.

6. Suppose we want to generate a random variable $X$ which has a PDF

$$f(x) = \frac{1}{2}x^2 e^{-x}, \quad x > 0.$$

   We hope to use the acceptance-rejection method with an exponential density having rate $\lambda$ (mean $1/\lambda$). Find the value of $\lambda$ that minimizes the expected number of iterations of the algorithm used to generate $X$.