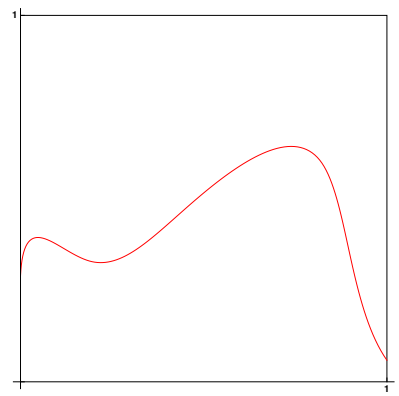
**168A PRACTICE MIDTERM SOLUTIONS**

In problem 1, write pseudocode for your algorithm; you can also describe in words what it is doing. In problems 2 and 3, showing your work in addition to the final answer might earn some partial credit in case the latter is not exactly right.

1. Construct a randomized algorithm to approximate the integral of an arbitrary function $f : [0,1] \to [0,1]$. Assume that the function is provided as a subroutine which returns $f(x)$ when passed argument $x$; also assume you have access to a random number generator random() that returns a uniformly random number in the interval $[0,1]$, *e.g.*, $x \leftarrow$ random() means that $x \in [0,1]$, sampled from the constant probability density function. Design your algorithm so that it has probability at least 2/3 of giving an estimate for the integral that is correct to one decimal place.

   The unknown function $f$ might look like this plot; in any case our task is to compute its integral, namely the area $A$ below its curve. If we sample points uniformly at random, independently, in the square region $[0,1] \times [0,1]$ the probability that each lies below the curve is $A$. So if we sample $N$ points, the number $C$ that lie below the curve is a binomial random variable $C \sim \text{Binomial}(N, A)$. The expectation value of $C$ is $\mathsf{E}[C] = NA$, whence $\mathsf{E}[C/N] = A$. The variance in this estimator is




$$\mathsf{Var}[C/N] = \mathsf{Var}[C]/N^2 = A(1-A)/N,$$

   so its standard deviation is $\sqrt{A(1-A)/N}$. Since the binomial distribution is approximated by a normal distribution, and slightly more than 2/3 of the probability is within 1 standard deviation of the mean in a normal distribution, we need to have

$$\sqrt{A(1-A)/N} < 0.05 = \epsilon, \quad \text{which implies} \quad \sqrt{N} > 20\sqrt{A(1-A)}.$$

   The right hand side of this inequality is largest when $A = 1/2$, so we choose $N > 100$. Here is pseudocode for this algorithm:

   input: $0 < \epsilon$; $f : [0,1] \to [0,1]$.
   output: $\int_0^1 f(x)\,\mathrm{d}x$.

   *count* $\leftarrow 0$
   repeat $N = \lceil 1/(4\epsilon^2) \rceil$ times:
   $\quad x \leftarrow$ random()
   $\quad y \leftarrow$ random()
   $\quad$ if $f(x) < y$, *count* $\leftarrow$ *count* $+1$
   return *count*$/N$

**168A PRACTICE MIDTERM SOLUTIONS**

2. Consider a random walk $\{X_t \in \mathbb{Z} \mid t \in \mathbb{N}\}$ with transition probabilities

$$P_{xy} = \begin{cases} 1/2 & \text{if } |x-y| = 1; \\ 0 & \text{otherwise,} \end{cases}$$

and initial state $X_0 = 0$.

a. Suppose $X_2 = 0$. What is the probability that $X_t \neq 0$ for any $0 < t < 2$, *i.e.*, that given the random walk has returned to 0 at $t = 2$, that this is the first return?

Since $X_1$ must be odd, it is definitely not 0, so the probability is 1.

b. Suppose $X_4 = 0$. What is the probability that $X_t \neq 0$ for any $0 < t < 4$, *i.e.*, that given the random walk has returned to 0 at $t = 4$, that this is the first return?

Since $X_4 = 0$ the random walk must take two positive steps and two negative steps. There are $\binom{4}{2} = 6$ such paths, exactly 2 of which ($++--$ and $--++$) have $t = 4$ as their first return time, so the probability is $2/6 = 1/3$.

c. Suppose $X_6 = 0$. What is the probability that $X_t \neq 0$ for any $0 < t < 6$, *i.e.*, that given the random walk has returned to 0 at $t = 6$, that this is the first return?

Since $X_6 = 0$ the random walk must take three positive steps and three negative steps. There are $\binom{6}{6} = 20$ such paths, exactly 4 of which ($+++---$, $++-+--$, and the reverses of each of these) have $t = 6$ as their first return time, so the probability is $4/20 = 1/5$.

3. a. Let $G$ be a graph with 4 vertices and adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

What is the transition probability matrix for the random walk on $G$ that for each vertex $v$ has equal transition probability to each vertex connected to $v$, and transition probability 0 to every other vertex?

The nonzero transition probabilities are in the same places as the 1s in $A$, but each column must sum to 1; thus

$$P = \begin{pmatrix} 0 & 1/3 & 0 & 0 \\ 1 & 0 & 1/2 & 1/2 \\ 0 & 1/3 & 0 & 1/2 \\ 0 & 1/3 & 1/2 & 0 \end{pmatrix}.$$

b. Find a probability distribution on the vertices of $G$ that is unchanged by the random walk evolution defined in problem 3.a.

Suppose the probability at vertex 1 is $p$. Then, since it all transfers to vertex 2, and 1/3 of vertex 2's probability transfers to vertex 1, there must be $3p$ at state 2. Similarly, there must be $2p$ at vertices 3 and 4. Since $p + 3p + 2p + 2p = 1$, $p = 1/8$, and the probability distribution is $(1, 3, 2, 2)/8$.