**168A MIDTERM SOLUTIONS**

In problem 1, write pseudocode for your algorithms; you can also describe in words what they are 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. a. [15 points] Construct an algorithm to select a point uniformly at random from the disk of radius 1 centered at the origin. 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.

   One approach is to sample points uniformly from the square $[-1, 1] \times [-1, 1]$ which circumscribes the disk, until we find one that lies in the disk:

   $\mathit{found} \leftarrow \mathsf{False}$
   while not $\mathit{found}$:
       $x \leftarrow 2 * \text{random}() - 1$
       $y \leftarrow 2 * \text{random}() - 1$
       if $x^2 + y^2 \leq 1$, $\mathit{found} \leftarrow \mathsf{True}$
   return $(x, y)$

   b. [20 points] Now construct an algorithm to approximate the area of a disk of radius 1, *i.e.*, to approximate $\pi$. Design your algorithm so that it has probability at least 2/3 of giving an estimate for $\pi$ that is correct to one decimal place, *i.e.*, it is within the interval $[\pi - 0.05, \pi + 0.05]$.
   Hint: For large $n$ and fixed $p$, you can approximate a binomial distribution by a normal distribution with the same mean and variance, *i.e.*, $\text{Binomial}(n, p) \approx \mathcal{N}\big(np, np(1 - p)\big)$.

   If we sample points uniformly at random, independently, in the square region $[-1, 1] \times [-1, 1]$ the probability that each lies in the disk is $\pi/4$. So if we sample $N$ points, the number $C$ that lie in the disk is a binomial random variable $C \sim \text{Binomial}(N, \pi/4)$. The expectation value of $C$ is $\mathsf{E}[C] = N\pi/4$, whence $\mathsf{E}[4C/N] = \pi$. The variance in this estimator is

$$\mathsf{Var}\left[\frac{4C}{N}\right] = \frac{16}{N^2}\mathsf{Var}[C] = \frac{16}{N}\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right) = \frac{\pi(4 - \pi)}{N}$$

   so its standard deviation is $\sqrt{\pi(4 - \pi)/N}$. 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

$$\sqrt{\pi(4 - \pi)/N} < 0.05 = \epsilon, \quad \text{which implies} \quad \sqrt{N} > 20\sqrt{\pi(4 - \pi)}.$$

   Even if we did not know $\pi$, since the disk is contained in the square we know it is less than 4; since the largest $\pi(4 - \pi)$ could be is 4 (if $\pi$ were 2), if we choose

$N = 400 \cdot 4 = 1600$, the estimate will be sufficiently accurate. Here is pseudocode for this algorithm:

input: $0 < \epsilon$.
output: $\pi$.

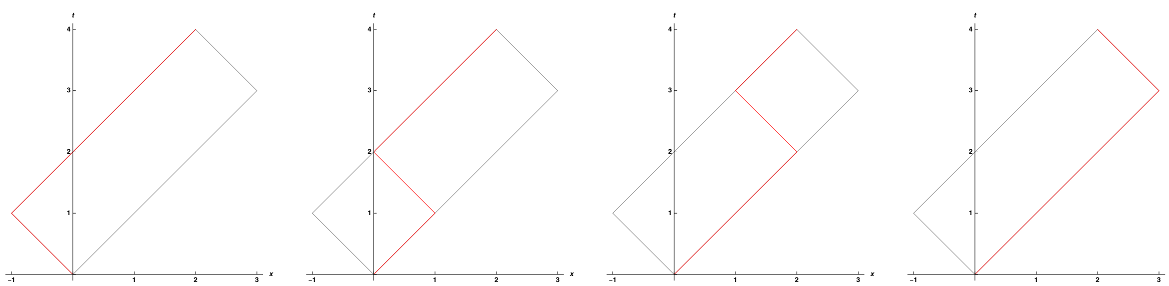
$count \leftarrow 0$
repeat $N = \lceil 4/\epsilon^2 \rceil$ times:
    $x \leftarrow 2 * \mathrm{random}() - 1$
    $y \leftarrow 2 * \mathrm{random}() - 1$
    if $x^2 + y^2 \le 1$, $count \leftarrow count + 1$
return $4 * count/N$

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. [15 points] Suppose $X_4 = 2$. Draw all the paths in $(x, t)$ space that the walk could follow.
Hint: Each path is like the path that you plotted in the first homework assignment.



b. [20 points] Compute the probability distributions for $X_1$, $X_2$, and $X_3$ given that $X_0 = 0$ and $X_4 = 2$, *i.e.*, compute $\Pr(X_t = x \mid X_0 = 0, X_4 = 2)$ for $t \in \{1, 2, 3\}$.

Each of the paths shown above has the same probability, $1/4$, conditional on $X_0 = 0$ and $X_4 = 2$. Restricting them to time $t$ gives the requested probability distributions:

$$\Pr(X_1 = x \mid X_0 = 0, X_4 = 2) = \begin{cases} 1/4 & \text{if } x = -1; \\ 3/4 & \text{if } x = 1; \\ 0 & \text{otherwise.} \end{cases}$$

$$\Pr(X_2 = x \mid X_0 = 0, X_4 = 2) = \begin{cases} 1/2 & \text{if } x = 0; \\ 1/2 & \text{if } x = 2; \\ 0 & \text{otherwise.} \end{cases}$$

$$\Pr(X_3 = x \mid X_0 = 0, X_4 = 2) = \begin{cases} 3/4 & \text{if } x = 1; \\ 1/4 & \text{if } x = 3; \\ 0 & \text{otherwise.} \end{cases}$$

3. a. [15 points] Write the transition probability matrix $P$ for a random walk on $S = \{1,2,3,4,5,6\} \subset \mathbb{N}$ with transition probabilities:

$$\Pr(X_{t+1} = x \mid X_t = y) = \begin{cases} 1/2 & \text{if } x = y; \\ 1/2 & \text{if } y \in \{1,6\} \text{ and } |x-y| = 1; \\ 1/4 & \text{if } y \notin \{1,6\} \text{ and } |x-y| = 1; \\ 0 & \text{otherwise.} \end{cases}$$

$$P = \frac{1}{4}\begin{pmatrix} 2 & 1 & & & & \\ 2 & 2 & 1 & & & \\ & 1 & 2 & 1 & & \\ & & 1 & 2 & 1 & \\ & & & 1 & 2 & 2 \\ & & & & 1 & 2 \end{pmatrix}.$$

b. [15 points] Find a probability distribution on $S$ that is unchanged by the evolution.
We want $v \in \mathbb{R}^6$ such that $Pv = v$, *i.e.*, $(P - I)v = 0$. Notice that $P$ is symmetric under the exchange of states $s$ and $7 - s$, so $v = (a,b,c,c,b,a)$. Inserting this into the equation gives $-2a + b = 0$, so $b = 2a$. Also, $2a - 2b + c = 0$, so $c = 2a$. Thus $v = (a, 2a, 2a, 2a, 2a, a)$. Since $v$ is a probability distribution, its elements must sum to 1, which means $a = 1/10$, and $v = (1,2,2,2,2,1)/10$.