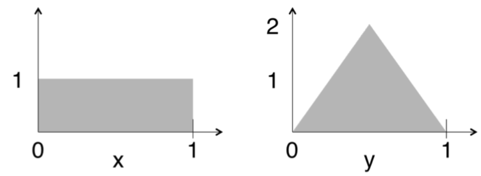
SOLUTIONS

1. **Random number generators.** I provide you with a uniform random number generator (i.e. an RNG that will give you a number *X* uniformly distributed between 0 and 1; left graph). I am interested in generating another number *Y* with the triangular distribution shown (right graph). How can you use the uniform RNG for *X* to generate values of *Y*? Hint: you are allowed to call the RNG for *X* more than once for each number *Y* you need to generate. **[5]**



Soln: Draw two values from the RNG, and set $y = \frac{x_1+x_2}{2}$. Alternatively, compute $q(y) = \int_0^y p(y')dy'$ and output $y = q^{-1}(x)$ where *x* is the output of the RNG.

2. **Markov processes.** Examine the following matrices.

$$A = \begin{bmatrix} 0.25 & 0.5 \\ 0.75 & 0.5 \end{bmatrix} \quad B = \begin{bmatrix} 0.25 & 0.1 \\ 0.4 & 0.25 \end{bmatrix} \quad C = \begin{bmatrix} 0.2 & 0.8 \\ 0.9 & 0.1 \end{bmatrix}$$

a. Which of these matrices (*A*, *B*, or *C*) can be interpreted as a Markov transition matrix for a single time step Δt of a stochastic process? Remember, we use the convention that probability distributions are column vectors. **[3]**

Soln: A is the only one whose columns are normalized, so it is the only possible Markov matrix.

b. For your choice from part (a), what is the corresponding Markov matrix for a time step $3\Delta t$? **[2]**

Soln: $A^3 = \frac{1}{64}\begin{bmatrix} 25 & 26 \\ 39 & 38 \end{bmatrix}$

3. **Master equation.** Radioactive decay is governed by the following Master equation

$$\frac{d}{dt}p_n = \gamma(-(n)p_n + (n+1)p_{n+1})$$

where *n* represents the number of radioactive nuclei in a lump of uranium, and p_n is the probability that a lump has precisely *n* radioactive nuclei ($\sum p_n = 1$).

a. We derived in class that the equation for the first moment $\langle n \rangle$ is the standard radioactive decay equation: $d\langle n \rangle/dt = -\gamma\langle n \rangle$. Now use the Master equation to derive the equation obeyed by the second moment $\langle n^2 \rangle$. **[10]**

Soln: $\frac{d}{dt} < n^2 > = \gamma\sum(-n^3 p_n + n^2(n+1)p_{n+1}) = \gamma\sum(-n^3 + n(n-1)^2)p_n = \gamma\sum(n - 2n^2)p_n$

So

$$\frac{d}{dt}\begin{bmatrix} < n > \\ < n^2 > \end{bmatrix} = \gamma \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} < n > \\ < n^2 > \end{bmatrix}$$

b. A nucleus has a probability γdt of decaying in any time interval dt. What is the probability $q(T)$ that a single nucleus that was present at $t = 0$ is *still present* at some time $t = T$? **[5]**

Soln: $q(T) = \lim_{dt\to 0}(1-\gamma dt)^{\frac{T}{dt}} = e^{-\gamma T}$

c. We know that each nucleus in a lump behaves independently. Assume you started with N nuclei in a lump at $t = 0$. Based on your answer from (b), what is the probability $p_n(T)$ that n out of N nuclei are *still present* in that lump at time $t = T$? **[5]**

Soln: The number of nuclei obeys a binomial distribution with probability $q(T) = e^{-\gamma T}$. I.e.

$$p_n(T) = \binom{N}{n} q(T)^n \big(1 - q(T)\big)^{N-n}$$

4. Coding. In a two-horse race, the probabilities that horse A or B wins are $[p_A, p_B] = [\frac{2}{3}, \frac{1}{3}]$. Answer the following questions.

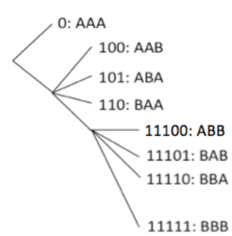
a. What is the entropy of this distribution? We know this will be the smallest number of bits per race possible in any coding scheme. You can use $\log_2(3) = 1.585$. **[5]**

Soln: $H\left(\frac{1}{3}\right) = -\left(\frac{1}{3}\right) log2\left(\frac{1}{3}\right) - \left(\frac{2}{3}\right) log2\left(\frac{2}{3}\right) = log2(3) - \frac{2}{3} = 1.585 - 0.666 \approx 0.918$

b. If you use the code $A \to 0, B \to 1$ then this is a one bit per race instantaneous code. Suppose instead you are satisfied with transmitting the results every *three* races. Build a code using the alphabet $\{0,1\}$ that has an average length of less than one bit *per race*. Please show your reasoning. **[15]**

Soln:

Outcome	Probability
AAA	8/27
AAB	4/27
ABA	4/27
BAA	4/27
ABB	2/27
BAB	2/27
BBA	2/27
BBB	1/27



We need to push more unlikely events to longer codes. The prefix-free code shown here has an average per-race length of 79/81 = 0.975 < 1 as required. It’s still not as compressed as possible if we’re willing to wait it out for more races.