2.13 Order statistics.

Write a function to use the beta distribution and the inverse CDF method to generate the \( j \) largest order statistics from a sample of size \( n \) from an exponential distribution with parameter \( \theta \). (The values of \( j, n, \) and \( \theta \) can be chosen by the user of the function. The meaning of \( \theta \) in the exponential distribution is as shown in Table D.5 in my notes on the web page.)

Now for certain values of \( k \) and \( n \) (specified below), use your program to estimate \( E(X_{(k:n)}) \), where \( X_{(k:n)} \) is the \( k \)th order statistic from a sample of size \( n \) from an exponential distribution with parameter \( \theta \). (That is, your answer will involve the parameter \( \theta \).)

Do this for \( k = 1 \) and 2, and \( n = 100 \) and 500. (There are 4 “answers”.)

As \( n \) increases without bound, of the max order statistic from an exponential distribution also grows without bound. If, however, we normalize it by subtracting \( \log(n) \) we arrive at an interesting limiting distribution. The distribution of \( Y = X_{(k:n)} - \log(n) \) is called a type 1 extreme value distribution, or a Gumbel distribution. It is easy to show that its CDF is

\[
F_Y(y) = e^{-e^{-y/\theta}} I_{(0,\infty)}(y).
\]

(This is not part of the exercise, but you might find it interesting to derive this limiting distribution.)

2.14 Random walk Markov chain.

(a) There are two simple configurations for a 1-D random walk with a finite number of states. One is linear, with two boundary points at which the particle must remain unmoved or else must reverse direction.

Suppose we have a 1-D linear random walk with 5 states and 2 boundary points. Now suppose that, at each step, the particle has an equal chance of moving or of remaining in the same state, and if there is movement from an interior (i.e., non-boundary) point, there is an equal chance of going to the right or to the left. Write the transition matrix \( P_l \) (in the old-fashioned, but standard way; i.e., the transpose of what a contemporary applied mathematician would write).

*Hint*: The first row is .5, .5, 0, 0, 0.

(b) Note that your matrix is *stochastic*; that is, each row sums to 1. Determine analytically if the random walk has a stationary, or limiting, distribution. You can use `eigen` in R. (Of course, your answer must be “yes” or “no” and the reason.)

(c) What is the stationary distribution? (Obviously, the answer to the previous question is “yes”.) Note that the distribution is a vector of nonnegative elements that sum to 1.

(d) Now, let’s try some numerical experiments. Start with the degenerate distribution \( p^{(0)} = (1, 0, 0, 0, 0) \) and consider the evolution of the distributions by

\[
p^{(t)} = p^{(t-1)} P_l.
\]

How many transitions are required to reach the stationary distribution?

(e) The previous question addressed probability distributions, not sample points.

Now suppose a particle begins at state 1, the leftmost boundary. Simulate the chain for 100000 steps and note the sample frequencies (that is, the number of times the particle was in state 1, state 2, etc.), after

(a) the first 100 steps

(b) the first 1000 steps
(c) the first 10000 steps
(d) the 100000 steps.

(f) The other simple configuration for a 1-D random walk with a finite number of states is circular, that is, there are no boundary points. The particle effectively moves around a circle. Now let’s consider a “uniform” 1-D circular random walk in which, at each step, the particle has an equal chance of moving or of remaining in the same state, and if there is movement, there is an equal chance of going in either direction. Write the transition matrix $P_c$.

Note that your matrix is stochastic; that is, each row sums to 1. Determine analytically if the random walk has a stationary, or limiting, distribution.

What is the stationary distribution?