1. (3 + 3) The 1-d percolation problem:

Consider an infinite 1-d lattice forest with a tree present at any site with probability $p$.

- Find the distribution of forest sizes as a function of $p$. Do this by moving along the 1-d world and figuring out the probability that any forest you enter will extend for a total length $\ell$.
- Find $p_c$, the critical probability for which a giant component exists.
  Hint: One way to find critical points is to determine when certain average quantities explode. Compute $\langle l \rangle$ and find $p$ such that this expression goes boom (if it does).

2. (3 + 3 + 3)

A courageous coding festival:

Code up the discrete HOT model in 1-d. Let’s see if we find any of these super-duper power laws everyone keeps talking about. We’ll follow the same approach as the 2-d forest discussed in lectures.

Main goal: extract yield curves as a function of the design $D$ parameter as described below.

Suggested simulations elements:
• \( N = 10^4 \) as a start. Then see if \( N = 10^5 \) or \( N = 10^6 \) is possible.
• Start with no trees.
• Probability of a spark at the \( i \)th site: \( P(i) \propto e^{-i/\ell} \) where \( i \) is tree position \((i = 1 \text{ to } N)\). (You will need to normalize this properly.) The quantity \( \ell \) is the characteristic scale for this distribution; try \( \ell = 2 \times 10^5 \).
• Consider a design problem of \( D = 1, 2, N^{1/2}, \text{ and } N \). (If \( N^{1/2} \) and \( N \) are too much, you can drop them. Perhaps sneak out to \( D = 3 \).) Recall that the design problem is to test \( D \) randomly chosen placements of the next tree against the spark distribution.
• For each test tree, measure the average yield (number of trees left) with \( n = 100 \) randomly selected sparks. Select the tree location with the highest average yield and plant a tree there.
• Add trees until the linear forest is full, measuring average yield as a function of trees added.
• Only trees and adjacent trees burn. In effect, you will be burning un-treed intervals of the line (much less complicated than 2-d).

(a) Plot the yield curves for each value of \( D \).
(b) Identify peak yield for each value of \( D \).
(c) Plot distributions of connected tree interval sizes at peak yield (you will have to rebuild forests and stop at the peak yield value of \( D \) to find these distributions.

Hint: keeping a list of un-treed locations will make choosing the next location easier. Hopefully.

3. The discrete version of HOT theory:

From lectures, we had the following.

Cost: Expected size of ‘fire’ in a \( d \)-dimensional lattice:

\[
C_{\text{fire}} \propto \sum_{i=1}^{N_{\text{sites}}} (p_i a_i) a_i = \sum_{i=1}^{N_{\text{sites}}} p_i a_i^2,
\]

where \( a_i \) = area of \( i \)th site’s region, and \( p_i \) = avg. prob. of fire at site in \( i \)th site’s region.

From lectures, the constraint for building and maintaining \((d - 1)\)-dimensional firewalls in \( d \)-dimensions is

\[
C_{\text{firewalls}} \propto \sum_{i=1}^{N_{\text{sites}}} a_i^{(d-1)/d} a_i^{-1},
\]
where we are assuming isometry.

Using Lagrange Multipliers, safety goggles, rubber gloves, a pair of tongs, and a maniacal laugh, determine that:

\[ p_i \propto a_i^{-\gamma} = a_i^{-(2+1/d)}. \]

4. **(3 + 3 + 3) Highly Optimized Tolerance:**

This question is based on Carlson and Doyle’s 1999 paper “Highly optimized tolerance: A mechanism for power laws in design systems” [11]. In class, we made our way through a discrete version of a toy HOT model of forest fires. This paper revolves around the equivalent continuous model’s derivation.

Our interest is in Table I on p. 1415:

<table>
<thead>
<tr>
<th>( x^{-(q+1)} )</th>
<th>( x^{-q} )</th>
<th>( A^{-\gamma(1-1/q)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e^{-x} )</td>
<td>( e^{-x} )</td>
<td>( A^{-\gamma} )</td>
</tr>
<tr>
<td>( e^{-x^2} )</td>
<td>( x^{-1} e^{-x^2} )</td>
<td>( A^{-\gamma[\log(A)]^{-1/2}} )</td>
</tr>
</tbody>
</table>

and Equation 8 on the same page:

\[ P_{\geq}(A) = \int_{p^{-1}(A^{-\gamma})}^{\infty} p(x)dx = p_{\geq} \left( p^{-1} \left( A^{-\gamma} \right) \right), \]

where \( \gamma = \alpha + 1/\beta \) and we’ll write \( P_{\geq} \) for \( P_{\text{cum}} \).

Please note that \( P_{\geq}(A) \) for \( x^{-(q+1)} \) is not correct. Find the right one!

Here, \( A(x) \) is the area connected to the point \( x \) (think connected patch of trees for forest fires). The cost of a ‘failure’ (e.g., lightning) beginning at \( x \) scales as \( A(x)^{\alpha} \) which in turn occurs with probability \( p(x) \). The function \( p^{-1} \) is the inverse function of \( p \).

Resources associated with point \( x \) are denoted as \( R(x) \) and area is assumed to scale with resource as \( A(x) \sim R^{-\beta}(x) \).

Finally, \( P_{\geq} \) is the complementary cumulative distribution function for \( p \).

As per the table, determine \( P_{\geq}(x) \) and \( P_{\geq}(A) \) for the following (3 pts each):

(a) \( p(x) = cx^{-(q+1)} \),
(b) \( p(x) = ce^{-x} \), and
(c) \( p(x) = ce^{-x^2} \).

Note that these forms are for the tails of \( p \) only, and you should incorporate a constant of proportionality \( c \), which is not shown in the paper.
References