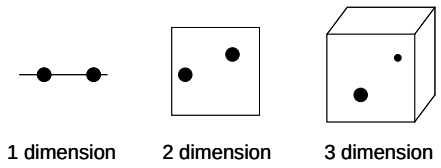
Chem 160/260 Project 1

You must log on to biopathway to do this Project. For full credit complete this Project before 12:00AM (midnight) on Sunday, October 25, 2015. You can turn this in up to one week late (by midnight Sunday, November 1, 2015) for a 25% penalty**. If you choose to turn your Project in late, email me so I know to look for it the following week.**

Project overview: Choose **one** of the following project assignments and complete it in your **`~/projects/project1`** directory. **Extra credit:** Do both of the project assignments.

Project Option 1: A long-standing mathematical question is the *Hypercube Line Picking* problem, which asks what is the average distance between two points randomly chosen in a hypercube with *n* dimensions and unit sides (i.e. each side has length=1). For example, consider the hypercubes with 1, 2, and 3 dimensions. The problem involves picking two points in the space, defined by 1, 2, or 3-coordinates for the 1-, 2-, and 3-dimensional cases, (and *n* coordinates for the *n*-dimensional case).



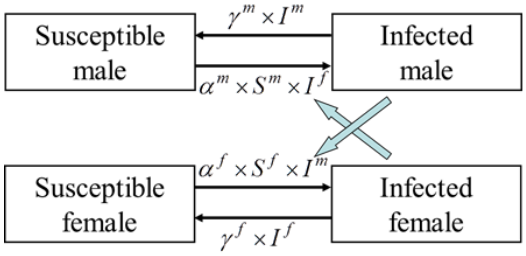
$$\frac{1}{3}\sqrt{n} \le Dist(n) \le \sqrt{\frac{1}{6}n}\sqrt{\frac{1}{3}\left[1+2\sqrt{\left(1-\frac{3}{5n}\right)}\right]}$$

Low n High n

The limiting average interpoint distances for the low-*n* and high-*n* limits are given by the equations shown at left (equations from http://mathworld.wolfram.com/HypercubeLinePicking.html)

Write a Monte Carlo R script (named **`hypercube.R`**) to calculate the average distance between random pairs of points in 1- to 30-dimensions and make a plot of this average distance vs the number of dimensions. Hint 1: Have an outer loop over the number of dimensions and an inner loop over number of trials (use at least 1000 trials). Hint 2: This problem closely follows the example in section III of the class 9 lecture notes. Hint 3: For each trial store the distance between the points in an array that grows by 1 for every new trial, and then calculate the average value in the array using the **`mean()`** command and store that result in another vector that grows by one with each new dimension added. After your script is complete this second vector will contain 30 values (the first for the 1-D case, the second for 2-D, etc. up to 30-D).

Project Option 2: In class 8 we talked about the SIR model which is appropriate for diseases that spread homogeneously regardless of gender and for which individuals become permanently resistant. Many sexually transmitted diseases do not lead to permanent resistance and are more accurately treated as multi-population Susceptible-Infected-Susceptible (SIS) models shown in the following diagram and with the following difference equations:



$$\Delta S^m = \gamma^m \times I^m - \alpha^m \times S^m \times I^f$$
$$\Delta I^m = \alpha^m \times S^m \times I^f - \gamma^m \times I^m$$
$$\Delta S^f = \gamma^f \times I^f - \alpha^f \times S^f \times I^m$$
$$\Delta I^f = \alpha^f \times S^f \times I^m - \gamma^f \times I^f$$

Write an R script (named **`SIS.R`**) that implements this multi-population SIS model. Use the following initial populations and parameters to produces a graph showing the populations of infected males (I^m) and females (I^f) plotted on a single graph, running separate simulations and making separate graphs for both values of α^m shown in the table. Hint: modify the SIR.R script we used in class (provided in the project folder), to have a separate variable for S^m, I^m, S^f, and I^f. You should run this simulation for at least 2000 time steps.

Initial S^m=14000	Initial I^m=1000	γ^m=0.05	α^m=0.000006 α^m=0.000002
Initial S^f=9000	Initial I^f=1000	γ^f=0.007	α^f=0.0000009