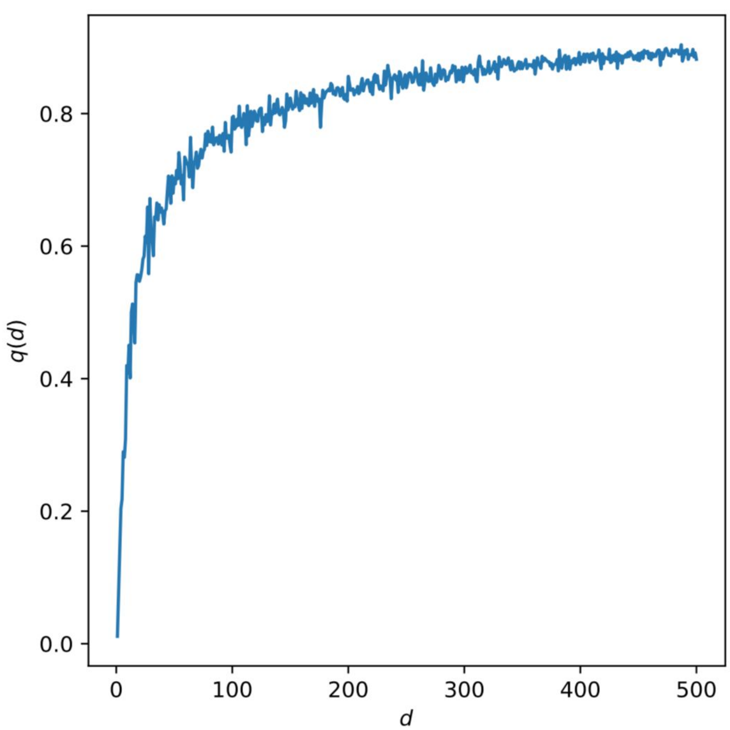
Brandt Dudziak 10/3

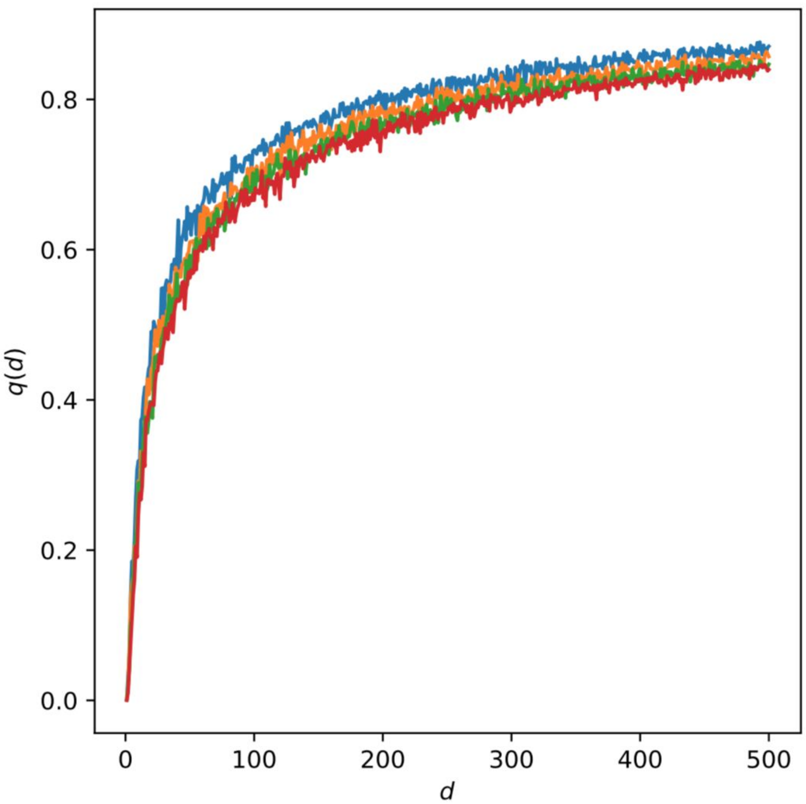
Q1.A:

As d increases, representing an increase in the dimensions, there is greater potential for big differences between points (the max R goes up), as the difference between a point (0, 0, …, 0) and (1, 1, …, 1) increases with dimensionality, according to the Euclidean distance equation. However, as d increases, the probability that things are far away, somewhat paradoxically, goes down. In other words, the expected values for both r and R go down. This can be seen concretely in expanding from one dimension to two — in one dimension, if there are two points that create the distance R, there has to be a point in between them to create the distance r, meaning $q(d)$ must be less than 1. However, in two dimensions, three points can create an equilateral triangle, meaning $q(d)$ can actually equal 1. I believe this means that a plot of d vs $q(d)$ would start at some point close to 0 when $d = 1$ and then quickly approach 1 as d increases.



This plot demonstrates what I've outlined above — $q(d)$ begins near 0 in the first dimension, then rapidly approaches 1 (or, more closely, 0.9) as d increases, flattening out once d starts getting really large, creating a logarithmic shape.

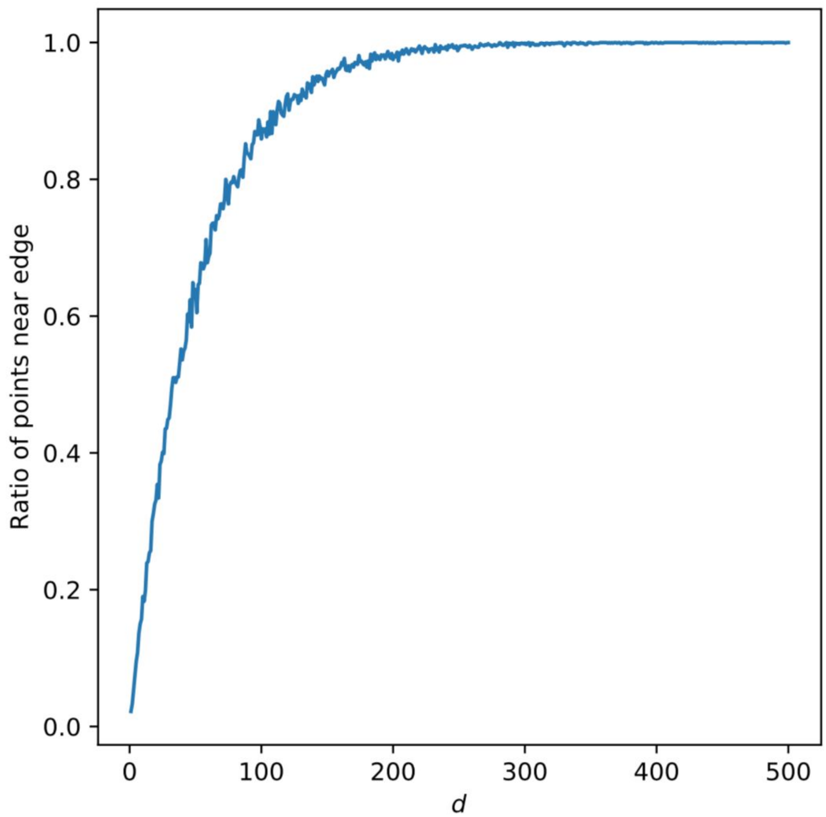
Q1.B:



In this plot, n = 20 is the blue line and n = 50 is the red. As n increases, the plots they create of *d* vs *q(d)* start becoming less logarithmic and very slightly more linear. I imagine that as n gets really large, like at 500, the plots will begin to look more like square root functions, as the estimate of the expected value of *q(d)* gets closer to the actual value, which, for the value of *R*, increases the sum under the root by one in the Euclidean distance equation. *r* stays the same while *R* increases.

Q2.A:

In lower dimensions, most points are away from the edges, whereas in higher dimensions, it’s much harder for a point to be away from one of the many, many edges. As the number of dimensions increases, the number of edges increases; thus, when sampling from 0 to 1 along these edges, there’s a greater likelihood that one of the coordinates of a given point is near an edge. This function follows the shape of *d* vs *q(d).*



Q2.B:

The ratio shoots incredibly quickly from 1 to practically 0. This aligns with the above finding about the ratio of points near edges, since as the ratio of points near an edge gets larger, the ratio of points contained within the 0.5 radius hypersphere, points near the center, naturally decreases.