**Part I: The Theoretical (Calculus) Approach**

*Step 1: Geometric analysis of the problem*

The first thing to note about this problem is that the shortest distance to any side in a square will always be the perpendicular distance to the side. This can be easily proved by Pythagoras’s theorem, as any distance between a point and a side that is not perpendicular to the side (c) can always be regarded as the hypotenuse of a right triangle made from the perpendicular distance (b) and the side (a), as seen on *Figure-1.1*. Since that the Pythagoras’s theorem states that:

, a non-perpendicular distance will always be longer than a perpendicular one between a point and a side.

Given that a square has four sides in total, any point in a square will have four different perpendicular distances between it and each of the four sides, but only one of them will be the shortest distance. Instead of considering all four sides at once, we can focus on one side at a time if we split the unit square into four analogous quarter triangles as shown on *Figure-1.2A*. As on *Figure-1.2A*, if a point falls within the area of Quarter Triangle A, it can be said for definite that this point will have the shortest perpendicular distance with Side A, and the same can be said about Side-Triangle pairs B, C and D.



*Figure-1.2B* shows the dimensions of a quarter triangle. Since that each of the four quarter triangles are completely analogous, the probability that a random point in the unit square falling within the bounds of any of the four quarter triangles are equally likely. This means that the average distance between any point within a quarter triangle and the side it shares with unit square is equivalent to the average shortest distance between any point in the unit square and the closest side, hence the problem can be simplified by considering a single quarter triangle instead of the whole unit square.

A quarter triangle can be modeled as follows on a cartesian plane, as seen on *Figure-1.3*:

­

*Step 2: Finding the probability of randomly selecting a specific point in the triangle*

To find the exact average of the distances between any point and the side, all possible points within the quarter triangle will have to be considered. One way to do this is to find an expression for the probability of randomly selecting a specific point out of all possible points in the quarter triangle.

To find the probability of randomly selecting a specific point in the quarter triangle, we should first consider the probability of selecting a random point in the square formed by the domain and range of and . Let and be any continuous length of *x* and *y* values within the restricted domain and range, as represented *Figure-1.4A*. The probability of randomly selecting a point that falls within the length of or are respectively and , hence the probability of randomly selecting a point that falls within the rectangular area formed by and , “P(Area)”, can be expressed as:



*Figure-1.4B* shows that when Δ*x* and Δ*y* approach infinitesimal and become *dx* and *dy*, the rectangular area they form will approach a specific point*.* Therefore, the probability of randomly selecting a single point, “P(Point)”, in the restricted domain and range, would be:

Since the area of the quarter triangle is half of the restricted domain and range, there is half as many possible positions that can be randomly selected in the quarter triangle than the restricted domain and range, which also means that the probability of randomly selecting a specific point in the quarter triangle will be twice of the probability of random selecting a specific point within the domain and range. P(Point) in a quarter triangle is hence calculated as:

*Step 3: Finding the perpendicular distance between any random point and the side*

After finding the probability of each point being selected in a quarter triangle, we need to find an expression for the distance between any point in the quarter triangle and the side of the unit square.

Let a random point *C* on *Figure-1.5* have the coordinates (x, y). Distance *d* is the perpendicular distance between point *C* and the side, the goal here is to find an expression for distance *d* in terms of *x* and *y*.

Suppose point *C* have the horizontal distance *b* and the vertical distance *a* from the side, extending to point *B* and point *A* to make triangle *ABC*. With information derived from *Figure-1.5*, the lengths of triangle *ABC*’s three sides *a*, *b*, and *c* can be expressed in terms of *x* and *y* as follows:

There are 2 ways to calculate the area of triangle *ABC*:

We can hence equate the two ways of finding the same area and substitute in the known expressions for *a*, *b* and *c* in terms of *x* and *y* to find calculate distance *d* as follows:

Given that the area is restricted to :

*Step 4: Finding the average perpendicular distance between any random point and the side in the quarter triangle*

To find the mean or expected value of any random variable, we take the sum of the products between all possible values of the random variable and their respective probability of occurring. From previous steps, we know how to find the probability of any specific point being randomly selected and how to find distance *d* for any specific point; Hence the product between any possible value of the continuous random variable *d* and its probability of occurring can be expressed as:

To find the expected value of *d*, “E(*d*)”, we put the product expression into a double integral with respect to both *x* and *y* to calculate the sum of the product of all possible *d* values and their probability of occurrence for all possible pairs of *x* and *y* values within the area of the quarter triangle. The first integral with respect to *x* will have the same bounds as the domain of the quarter triangle model to make sure that only *x* values that fall into the quarter triangle are considered. The bounds of the second integral with respect to *y* will change with respect to *x* given the inequality that models the quarter triangle, for that the possible *y* values changes at different *x* values. This finally gives the expression:

This expression can be simplified and solved as follows:

From the result of solving this equation, it can be hence concluded for Part I that the average shortest distance between any point in a unit square and the side is units.

*Reflection on Part I*

There is a minor error in the method used in this part of the investigation. By dividing the unit square into four quarter triangles to isolate all the possible points that would be closest to a single side, Step 1 of this investigation assumes that any specific point in the unit square can only be closest to one of its four sides. However, this is not true as all points that are positioned exactly on the cleavages of the quarter triangles are closest to two or more of the unit square’s sides. An example of this is demonstrated on *Figure-1.6A*, where the labelled point positioned on the cleavage between Quarter Triangles A and B is equally close to both sides A and B.

This causes the problem of double counting all the points that are closest to two or more sides of the unit square, when Steps 2, 3 and 4 of this investigation moves on to consider a quarter triangle independent from the rest of the unit square as shown on *Figure-1.6B*, where the labeled point that belongs to both Quarter Triangle A and B is counted as two separate points by how the average is calculated in this investigation.



Despite having the aforementioned error, the method used in Part I remains valid in terms of achieving the aim. This can be said because that only points along the cleavages between the quarter squares can be double counted, and the area of the cleavage is infinitely small compared to the area of the unit square, which hence makes the relative error caused by double counting infinitely small. However, infinitely small error does not mean there is no error, hence the result obtained from Part I cannot be taken as “exactly accurate”. On the other hand, with the purpose of this investigation being to solve a real-life problem, it is only meaningful to omit the infinitely small error and assume that is the exact solution to the problem, for that we can’t take an infinitely small measurement of the length of a swimming pool in practice.