SUPERVISOR'S REPORT

The supervisor should complete the report below and then give this cover, enclosing the final version of the extended essay, to the Diploma Programme coordinator. The supervisor must sign this report; otherwise the extended essay will not be assessed and may be returned to the school.

Name of supervisor (CAPITAL letters) _

Comments

If appropriate, please comment on the candidate's performance, the context in which the candidate undertook the research for the extended essay, any difficulties encountered and how these were overcome. These comments can help the examiner award a level for criterion H. Do not comment on any adverse personal circumstances that may have affected the candidate.

I appreciated how combined the advanced concepts of probability and calculus, along with the use of geometry to investigate a very interesting probability question. I was impressed also with the variations she came up with on the original situation. She thought of those variations on her own. I believe the quality of this paper is very high.

I have read the final version of the extended essay that will be submitted to the examiner.

To the best of my knowledge, the extended essay is the authentic work of the candidate.

I spent 2 hours with the candidate discussing the progress of the extended essay.

ASSESSMENT FORM (for examiner use only)

Candidate session number									

ACHIEVEMENT LEVEL

		First examiner	maximum	Second examiner
General assessment criteria *Refer to the general guidelines.*	**A** Research question		2	
	B Approach		3	
	C Analysis/interpretation		4	
	D Argument/evaluation		4	
	E Conclusion		2	
	F Abstract		2	
	G Formal presentation		3	
	H Holistic judgement		4	
Subject assessment criteria *Refer to the subject guidelines. Not all of the following criteria will apply to all subjects; use only the criteria which apply to the subject of the extended essay.*	**J**			
	K			
	L			
	M			
	TOTAL OUT OF 36			

Name of first examiner (CAPITAL letters): Examiner number:

Name of second examiner (CAPITAL letters): ______________________ Examiner number:

A Comparison of Two Methods of Determining the Probability of Random Meetings between Two People.

2,525 words

Table of Contents

Abstract

Probability is one of the most far-reaching subjects in mathematics, involving aspects from numerous other areas of mathematics, from elementary counting to the most complicated trigonometric equations. In particular, problems involving the probability of random meetings between two people can be solved using two separate fields of math: geometry and calculus. A problem of random meeting consists of two people who try to meet by each arriving at a certain spot at a random time within a designated time limit, staying for a certain amount of time, then leaving. This essay will examine the effectiveness of each method of solving these problems in the context of a variety of situations. The first example in this essay will be modeled after a problem selected from David Patrick's *The Art of Problem Solving: Introduction to Counting & Probability*, and the rest of the examples will be variations on this problem.

These problems can be visually modeled using geometry. Time can be represented with a square, and the successful meeting times can be shown as an area in that square. Alternately, using calculus, each person's probability of arriving at any time can be represented using a probability density function. Taking sums of selected integrals of these functions will lead to the total probability of a successful meeting.

When the circumstances of the meeting change, however, one method becomes preferable to the other. The aim of this essay is to compare the two methods in each of the three situations, and then reach a conclusion about the general effectiveness of each method.

Introduction

There are many ways to solve probability problems, such as using combinatorics, tables, charts or formulas. These methods are useful when the problem deals with discrete variables, such as the probability of heads or tails when flipping a coin or of obtaining a certain number when rolling dice. However, when variables are continuous, such as time or distance, these common methods will not include all possibilities. For example, problems of random meetings between two people deal with time, which is one of the most common continuous variables in probability. There are two common methods to solve these problems: geometry and calculus. This essay will compare the effectiveness of these methods in solving problems of random meetings between two people.

Example 1

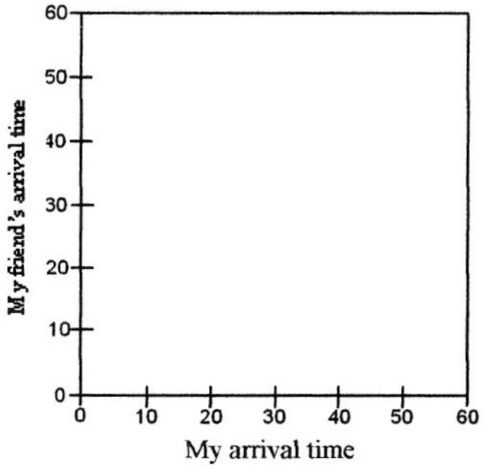
My friend and I plan to meet for lunch at our favorite restaurant. We will each arrive at a random time between noon and 1:00 PM, stay for 10 minutes, then leave. What is the probability that we will successfully meet each other at the restaurant?[1]

This problem cannot be solved by simply counting all the successful meetings and dividing that number by the number of all possible situations. Because time is continuous, there are infinitely many possibilities of each person's arrival time, so the probability of any one particular arrival is $\frac{1}{\infty}$, essentially 0. One way to overcome this obstacle is by using geometry.

---

[1] David Patrick, *The Art of Problem Solving: Introduction to Counting & Probability* (Alpine, CA: AoPS Inc., 2005) 160.

Let this square represent time between 12:00 and 1:00 PM. The horizontal axis represents my arrival time, and the vertical axis represents my friend's arrival time, in minutes after 12:00 PM.

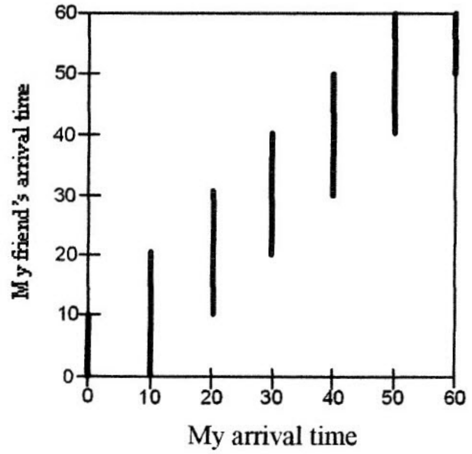


My friend and I will meet successfully if she arrives within 10 minutes before or after my arrival time. For example, if I arrive at 30 minutes after 12:00 PM, she can arrive at any time between 20 and 40 minutes after 12:00 PM, and we will successfully meet. In order to get an idea of where all the successful meetings will fall, we can look at a table:

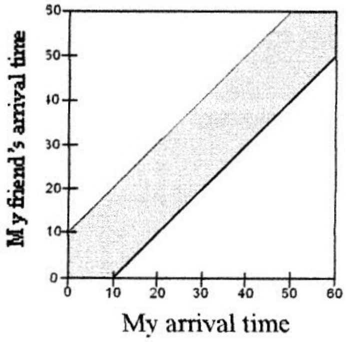
Successful Meeting Times

My arrival time	My friend's arrival time
0	0–10
10	0–20
20	10–30
30	20–40
40	30–50
50	40–60
60	50–60

Plotting these points, the diagram looks like this:



Now we can draw a trend line and shade in the area that represents the successful meetings:



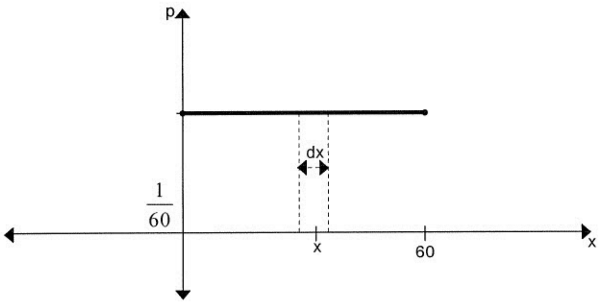
The probability that my friend and I meet successfully can be found by dividing the shaded area by the total area: $P = \frac{\text{shaded area}}{\text{total area}}$. However, the shaded area is not a regular shape, and it is easier to find the unshaded area, which consists of two right triangles. Therefore, it is easier to express the probability as $P = 1 - \frac{\text{unshaded area}}{\text{total area}}$. The total area is 60(60) = 3600, and each of the two unshaded triangles has side length 50, so the combined area of the triangles is 2(½)(50)(50) = 2500. By substituting these values, we find that $P = 1 - \frac{2500}{3600} = 1 - \frac{25}{36} = \frac{11}{36}$. The probability that my friend and I meet is $\frac{11}{36}$.

This problem can also be solved using calculus. The situation can be represented using a probability density function, which is defined as:

$$f(x) \geq 0$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

For this problem, let my probability density function be *f(x)* and my friend's function be *g(y)*.



Both graphs will be a step function where the domain is from 0 to 60, and the probability of arriving during any minute x in the domain is always $\frac{1}{60}$.

The probability that I arrive at a given time x is $f(x)dx$, the height of that interval times the width of that interval. My friend must arrive within 10 minutes of time x, that is, at a time between x-10 and x+10. This gives the equation: $P = \int_0^{60} \left[\left(\int_{x-10}^{x+10} g(y)dy \right) f(x) \right] dx$

In this case, $f(x) = g(y) = \frac{1}{60}$, so $P = \int_0^{60}\left[\left(\int_{x-10}^{x+10}\frac{1}{60}dy\right)\frac{1}{60}\right]dx$. However, between 0 and 10, x-10 does not exist, because my friend cannot arrive before 12:00 PM. Similarly, between 50 and 60, x+10 does not exist, because she cannot arrive after 1:00 PM. Because of this, we have to divide the equation into three intervals: 0-10, 10-50 and 50-60. For the first interval, the inside integral will be evaluated from 0 to x+10 because if I arrive before 10 minutes, the earliest my friend could have arrived for us to meet is at 0 minutes. The inside integral of the middle interval will still be evaluated from x-10 to x+10 because both exist. Finally, the inside integral of the third interval will be evaluated from x-10 to 60 because if I arrive after 50 minutes, the latest my friend can arrive for us to meet is at 60 minutes. With this information in mind, the equation looks like this:

$$P = \int_0^{10}\left[\left(\int_0^{x+10}\frac{1}{60}dy\right)\frac{1}{60}\right]dx + \int_{10}^{50}\left[\left(\int_{x-10}^{x+10}\frac{1}{60}dy\right)\frac{1}{60}\right]dx + \int_{50}^{60}\left[\left(\int_{x-10}^{60}\frac{1}{60}dy\right)\frac{1}{60}\right]dx$$

First, evaluate the inside integrals:

$$P = \int_0^{10}\left[\left(\frac{1}{60}(x+10) - \frac{1}{60}(0)\right)\frac{1}{60}\right]dx + \int_{10}^{50}\left[\left(\frac{1}{60}(x+10) - \frac{1}{60}(x-10)\right)\frac{1}{60}\right]dx$$
$$+ \int_{50}^{60}\left[\left(\frac{1}{60}(60) - \frac{1}{60}(x-10)\right)\frac{1}{60}\right]dx$$

For convenience, factor $\frac{1}{3600}$ out of each integral and simplify:

$$P = \frac{1}{3600}\left[\int_0^{10}(x+10)dx + \int_{10}^{50}(20)dx + \int_{50}^{60}(70-x)dx\right]$$

Finally, evaluate each integral:

$$P = \frac{1}{3600}\left[(150-0) + (1000-200) + (2400-2250)\right] = \frac{1100}{3600} = \frac{11}{36}$$

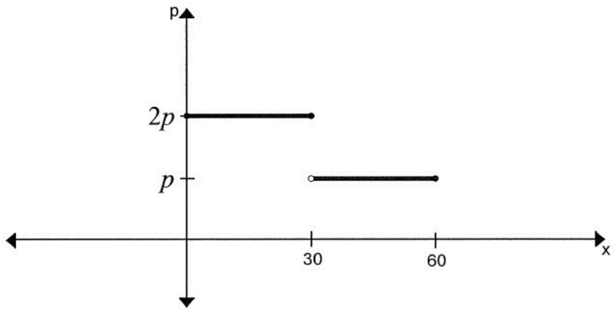
This answer agrees with the answer obtained from the geometric method.

This method seems impractical because it requires much more work than the geometric method. This method also requires the knowledge of calculus, where the geometric method requires only the ability to find areas of squares and triangles. However, with this method, there can be any functions for *f(x)* and *g(y)*.

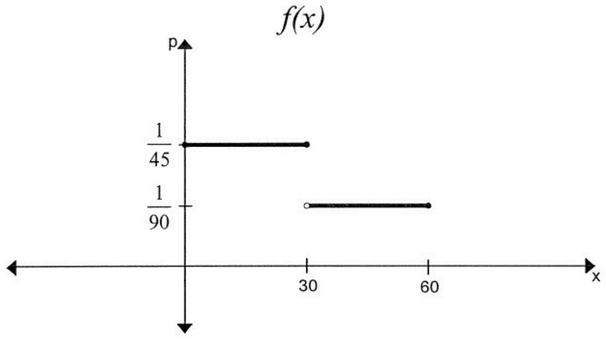
Example 2

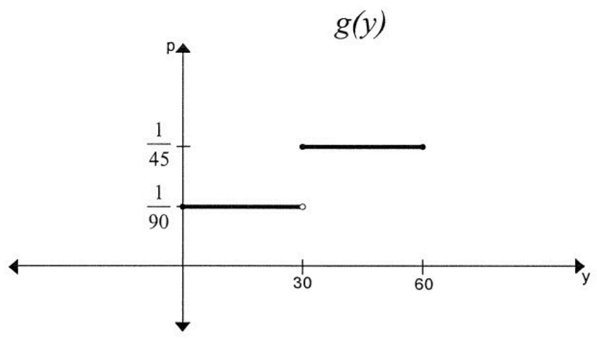
My friend and I plan to meet for lunch at our favorite restaurant. We will each arrive at a random time between noon and 1:00 PM, stay for 10 minutes, then leave. If I am twice as likely to arrive in the first half hour than in the second, and my friend is twice as likely to arrive in the second half hour than in the first, what is the probability that we successfully meet?

First, we know that the graph of *f(x)* will look like this, with the probability of arriving in any minute in the first half hour twice the probability of arriving in any minute in the second half hour:



Next, we need to know the probabilities that I will arrive in both the first and second half hours. If the probability that I arrive in the second half hour is p, then the probability that I arrive in the first half hour is $2p$. Recall that $\int_0^{60} f(x)dx = 1$, so the area under the entire graph equals one. If we divide the graph into two rectangles, one from 0 to 30 with height $2p$ and the other from 30 to 60 with height p, then $30(2p)+30p=1$, so $p=\frac{1}{90}$. Now we can accurately graph *f(x)* and *g(y)*:





Recall that $P = \int_0^{60}\left[\left(\int_{x-10}^{x+10} g(y)\,dy\right) f(x)\right]dx$. Before, we divided this into three integrals, but in this case, since *f(x)* and *g(y)* are not continuous from 0 to 60, we will have to divide the integral into six parts, each interval spanning 10 minutes. The integrals evaluated from 0-10 and from 50-60 will look like the first and last integrals in the previous problem, except the values for *f(x)* and *g(y)* will be different. From 10-20 and from 40-50, there will be no restrictions because the functions are continuous over those intervals, so the inner integrals will be evaluated from x-10 to x+10. However, from 20-30 and from 30-40, the inner integrals will also have to be divided into parts because *f(x)* and *g(y)* are not continuous over those intervals. If I arrive at a time x between 20 and 30 minutes with a probability of $\frac{1}{45}$, my friend can arrive from x-10 to 30 minutes with a probability of $\frac{1}{90}$, or she can arrive from 30 to x+10 minutes with a probability of $\frac{1}{45}$. This gives the expression $\int_{20}^{30}\left[\left(\int_{x-10}^{30}\frac{1}{90}dy + \int_{30}^{x+10}\frac{1}{45}dy\right)\frac{1}{45}\right]dx$. If I arrive at a time x between 30 and 40 minutes with a probability of $\frac{1}{90}$, my friend can arrive from x-10 to 30 minutes with a probability of $\frac{1}{90}$, or she can arrive from 30 to x+10 minutes with a probability of $\frac{1}{45}$. This gives the expression $\int_{30}^{40}\left[\left(\int_{x-10}^{30}\frac{1}{90}dy + \int_{30}^{x+10}\frac{1}{45}dy\right)\frac{1}{90}\right]dx$. After we put all of these integrals together, we get:

$$P=\int_{0}^{10}\left[\left(\int_{0}^{x+10}\frac{1}{90}dy\right)\frac{1}{45}\right]dx+\int_{10}^{20}\left[\left(\int_{x-10}^{x+10}\frac{1}{90}dy\right)\frac{1}{45}\right]dx+\int_{20}^{30}\left[\left(\int_{x-10}^{30}\frac{1}{90}dy+\int_{30}^{x+10}\frac{1}{45}dy\right)\frac{1}{45}\right]dx$$
$$+\int_{30}^{40}\left[\left(\int_{x-10}^{30}\frac{1}{90}dy+\int_{30}^{x+10}\frac{1}{45}dy\right)\frac{1}{90}\right]dx+\int_{40}^{50}\left[\left(\int_{x-10}^{x+10}\frac{1}{45}dy\right)\frac{1}{90}\right]dx+\int_{50}^{60}\left[\left(\int_{x-10}^{60}\frac{1}{45}dy\right)\frac{1}{90}\right]dx$$

Evaluate the inside integrals:

$$P=\int_{0}^{10}\left[\left(\frac{x+10}{90}\right)\frac{1}{45}\right]dx+\int_{10}^{20}\left[\left(\frac{(x+10)-(x-10)}{90}\right)\frac{1}{45}\right]dx+\int_{20}^{30}\left[\left(\frac{30-(x-10)}{90}+\frac{(x+10)-30}{45}\right)\frac{1}{45}\right]dx$$
$$+\int_{30}^{40}\left[\left(\frac{30-(x-10)}{90}+\frac{(x+10)-30}{45}\right)\frac{1}{90}\right]dx+\int_{40}^{50}\left[\left(\frac{(x+10)-(x-10)}{45}\right)\frac{1}{90}\right]dx+\int_{50}^{60}\left[\left(\frac{60-(x-10)}{45}\right)\frac{1}{90}\right]dx$$

For convenience, factor $\frac{1}{45\cdot 90}$ out of each integral and simplify:

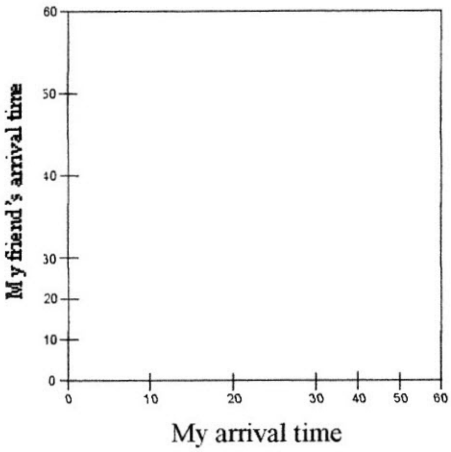
$$P=\frac{1}{4050}\left[\int_{0}^{10}(x-10)dx+\int_{10}^{20}(20)dx+\int_{20}^{30}(x)dx+\int_{30}^{40}\left(\frac{x}{2}\right)dx+\int_{40}^{50}(20)dx+\int_{50}^{60}(70-x)dx\right]$$

Finally, evaluate each integral:

$$P=\frac{1}{4050}[150+200+250+175+200+150]=\frac{1125}{4050}=\frac{5}{18}$$

In this case, the probability that my friend and I meet is $\frac{5}{18}$. This makes sense because it is close to but less than $\frac{11}{36}$, the probability of meeting when the probabilities are constant.

Geometrically, this problem is similar to the previous problem in that time can be represented as a square. However, on my axis, the increments for the first half hour must be twice as long as the increments for the second half hour because the probability that I arrive in the first half hour is twice the probability that I arrive in the second half hour. For the same reason, on my friend's axis, the increments for the second half hour must be twice as long as the increments for the first half hour:

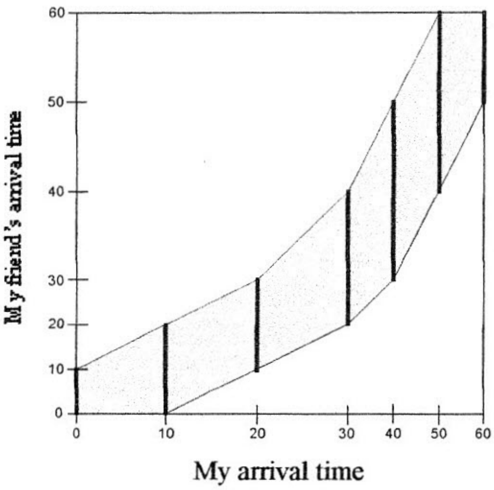


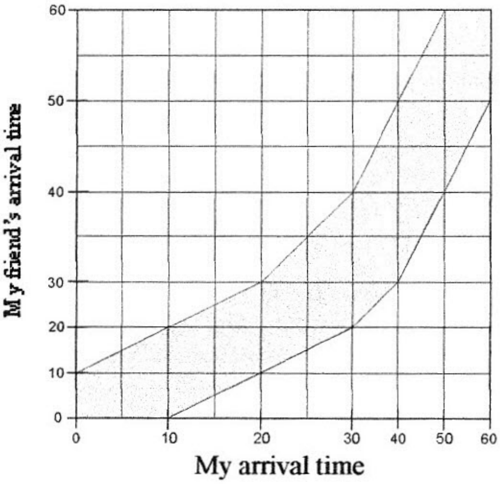
Now, take the table that we used in the first geometric representation and plot the information on this diagram:

Successful Meeting Times

My arrival time	My friend's arrival time
0	0–10
10	0–20
20	10–30
30	20–40
40	30–50
50	40–60
60	50–60

Plotting these points and drawing a trend line, the shaded area representing the successful meetings looks like this:





It is easiest to find the shaded area by drawing a grid over the diagram and simply counting squares. Each square must be weighted equally because finding the area by the length of each segment according to the scale would offset the effect of the scaling. There are 22.5 shaded squares out of a total of 9x9 squares, so the probability of a successful meeting is:

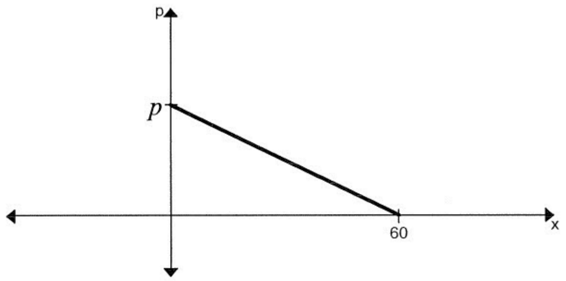
$$P = \frac{22.5}{81} = \frac{5}{18}$$

This answer agrees with the answer obtained using calculus. Again, the geometric method of solving the problem is much simpler and requires less work.

Example 3

My friend and I plan to meet for lunch at our favorite restaurant. We will each arrive at a random time between noon and 1:00 PM, stay for 10 minutes, then leave. If the probability that my friend and I arrive decreases linearly as time goes on until the probability that we each arrive at the end of the hour is zero, what is the probability that we will meet each other at the restaurant?

In this case, f(x) and g(y) both look like this:



The initial probability that I arrive is p. Because the area under the line must equal one, the area of the triangle formed by the line and the axes with base 60 and height p must equal one. If $\frac{1}{2}(60)p = 1$ then p must be $\frac{1}{30}$.

The slope of *f(x)* can be found using the two points $(0, \frac{1}{30})$ and $(60,0)$: $m = \frac{\frac{1}{30} - 0}{0 - 60} = \frac{-1}{1800}$

Therefore, $f(x) = \frac{1}{30} - \frac{1}{1800}x$ and $g(y) = \frac{1}{30} - \frac{1}{1800}y$. These functions are continuous, so the integrals only need to be broken into three parts:

$$P = \int_0^{10}\left[\left(\int_0^{x+10}\left[\frac{1}{30} - \frac{1}{1800}y\right]dy\right)\left(\frac{1}{30} - \frac{1}{1800}x\right)\right]dx + \int_{10}^{50}\left[\left(\int_{x-10}^{x+10}\left[\frac{1}{30} - \frac{1}{1800}y\right]dy\right)\left(\frac{1}{30} - \frac{1}{1800}x\right)\right]dx$$
$$+ \int_{50}^{60}\left[\left(\int_{x-10}^{60}\left[\frac{1}{30} - \frac{1}{1800}y\right]dy\right)\left(\frac{1}{30} - \frac{1}{1800}x\right)\right]dx$$

These fractions will make integration difficult, so it is easiest to factor $\frac{1}{1800^2}$ from each integral, then evaluate the inside integrals:

$$P=\frac{1}{1800^2}\left(\int_0^{10}\left[\left(\left[60(x+10)-\frac{(x+10)^2}{2}\right]-\left[60(0)-\frac{0^2}{2}\right]\right)(60-x)\right]dx\right.$$
$$+\int_{10}^{50}\left[\left(\left[60(x+10)-\frac{(x+10)^2}{2}\right]-\left[60(x-10)-\frac{(x-10)^2}{2}\right]\right)(60-x)\right]dx$$
$$\left.+\int_{50}^{60}\left[\left(\left[60(60)-\frac{60^2}{2}\right]-\left[60(x-10)-\frac{(x-10)^2}{2}\right]\right)(60-x)\right]dx\right)$$

Simplify:

$$P=\frac{1}{1800^2}\left(\int_0^{10}\left[\frac{x^3}{2}-80x^2+2450x+3300\right]dx+\int_{10}^{50}\left[20x^2-2400x+72000\right]dx\right.$$
$$\left.+\int_{50}^{60}\left[-\frac{x^3}{2}+100x^2-6650x+147000\right]dx\right)$$

Finally, evaluate each integral:

$$P=\frac{1}{1800^2}\left(\frac{1281250}{3}+\frac{2480000}{3}+\frac{21250}{3}\right)=\frac{3782500}{9720000}=\frac{1513}{3888}$$

In this case, the probability that my friend and I meet is $\frac{1513}{3888}$. This makes sense because it is close to but a little more than $\frac{11}{36}$, the probability of meeting when the probabilities are constant.

This situation is difficult to represent using pure geometry. In the last example, it was possible to rescale the axes of the square to give half of the area more weight. This worked because my friend and I both had a constant probability of arriving in each minute from 0 to 30, and although the probability changed after 30 minutes, we still had constant probabilities of arriving in any minute from 30 to 60. This cannot be said about this situation. The probability that I arrive in the first minute is not the same as the probability that I arrive in the second minute; in fact, there are not even two seconds in which I have the same probability of arriving. For this reason, there is no way to scale the axes of the square without dividing it into a grid of infinitely small squares. To better understand this concept, imagine the graph of

$y=\int_0^x\left(\frac{1}{30}-\frac{1}{1800}x\right)dx$. This is the same as writing $y=\frac{1}{30}x-\frac{1}{3600}x^2$ after solving the integral. This equation defines a parabolic curve, whereas in the other two problems, $y=\int_0^x f(x)dx$ was a line. Because of this, there is no way to represent time from 12:00 to 1:00 PM as a square. Instead, the sides must be curved to compensate for the nonlinear change in cumulative probability. At this point, the geometric method is no longer purely geometric because it involves calculus. Therefore, there is no way to solve this problem using only geometry.

Conclusion

When solving problems of random meetings, the specifics of the problem determine the preferred method. If the probability density function for each person is made up of horizontal lines, continuous or not, it is more straightforward to solve the problem using geometry. The advantage of this method is that it requires only knowledge of finding areas of geometric shapes, and no calculus is needed. There are also fewer steps in the process, so there are fewer opportunities to make mathematical errors. If the probability density functions for each person are not horizontal, it is best to use calculus to solve the problem. Although this method requires knowledge of calculus, it can be used for any type of probability density function, so it is not limited to only functions with horizontal lines.

In these three examples, my friend and I waited for ten minutes before leaving. If we had waited for a different amount of time, say m minutes, the methods of solving the problems would be the same, except instead of $x+10$ and $x-10$, we would use $x+m$ and $x-m$, and time would be measured in m-minute intervals instead of 10-minute intervals. In this way, these methods of solving random meeting problems can be generalized to include solutions to similar types of problems involving random meetings.

Bibliography

Patrick, David. *The Art of Problem Solving: Introduction to Counting & Probability*. Alpine, CA: AoPS Inc., 2005.

Tosch, Tom. Manager, Computational Mathematics Group, The Boeing Company, Seattle, WA. Interview, 7 February 2007.

Diagrams made using:
The Geometer's Sketchpad: Dynamic Geometry Software for Exploring Mathematics. Student Edition, Version 4. Designed by Nicholas Jackiw. Emeryville, CA: Key Curriculum Press, 2001.