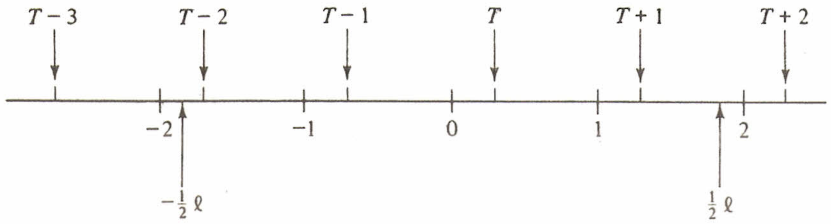


FIGURE 3.29 Illustrative positioning of random lattice.

or

$$E[N(T)] = \ell \tag{3.76}$$

as we might have expected intuitively. This is a result that generalizes readily to two and more dimensions (see [KEND 63, pp. 102–104]).

If $\ell = p + q$, where p is integral and $0 \le q < 1$, the variance is

$$\sigma^2_{N(T)} = q(1 - q) \tag{3.77}$$

How did we obtain this result?

Hint: Recognize (3.77) as the variance of a Bernoulli random variable.

Question: Can you compare the mean coverage of the random position model (Example 10) to the mean coverage of the lattice-position model (Example 11)?

Further work: Problem 3.24.

The coverage examples discussed above are illustrative of the types of problems that can be tackled using coverage concepts. However, apparently simple variations in coverage model assumptions seem to yield an intractable model much more readily than do models employing more conventional (non-area-based) random variables.

3.7 EXPECTED TRAVEL DISTANCES AND TIMES: SOME PRACTICAL RESULTS

Much of the material so far in this chapter has direct applications to the important problem of estimating the expected value of the distance covered (or travel time needed) by urban response units in traveling to the location of requests for assistance. In this section we will discuss this problem further and develop some simple approximate results which are often very useful in practical applications.

3.7.1 Simple Model

Example 12: Design of a Response District

Suppose that we have once more the situation described in Exercise 3.1, where requests for assistance are medical emergencies and the urban response unit is an ambulette. Under the assumptions that (1) locations of a medical emergency (X_1, Y_1) and of the ambulette (X_2, Y_2) are independent and uniformly distributed over the response district, and (2) travel is parallel to the sides of the rectangular response area, the travel distance [from (3.11)] is given by

$$D = |X_1 - X_2| + |Y_1 - Y_2|$$

From Exercise 3.1, we then have that

$$E[D] = \tfrac{1}{3}[X_0 + Y_0] \tag{3.12a}$$

where X_0 and Y_0 are the sides of the rectangle (see Figure 3.3). In this example we wish to formulate and solve the problem of optimal district design and to investigate the sensitivity of our results to suboptimal designs.

Solution

To find the district dimensions which lead to the minimum expected travel distance, we must keep the area of the response district $A_0 = X_0 Y_0$ constant and minimize (3.12a) subject to the condition $Y_0 = A_0/X_0$. Without this constant, a zero area (point) district would be optimal, an obviously infeasible result considering that the collection of districts in a city must usually cover the entire city (which has fixed positive area). Not surprisingly, (3.12a) is minimized when the rectangle becomes a square,

$$X_0 = Y_0 = \sqrt{A_0} \tag{3.78}$$

In that case we have

$$E[D] = \tfrac{2}{3}\sqrt{A_0} \tag{3.79}$$

More generally, if the effective travel speeds in the x-direction and the y-direction, v_x and v_y, are independent of travel distance, the *expected travel time*,

$$E[T] = \frac{1}{3}\left(\frac{X_0}{v_x} + \frac{Y_0}{v_y}\right)$$

is minimized when

$$\frac{Y_0}{v_y} = \frac{X_0}{v_x} = \sqrt{\frac{A_0}{v_x v_y}} \tag{3.80}$$

in which case

$$E[T] = \frac{2}{3}\sqrt{\frac{A_0}{v_x v_y}} \tag{3.81}$$

Intuitively speaking, the optimal shape of the district, as given by (3.80), is the one for which it takes as much time to traverse the district from "east to west" as from "north to south."

The expressions for $E[D]$ and $E[T]$ turn out to be "robust" (i.e., rather insensitive to the exact values of X_0 and Y_0). To see this, let us examine the case where

$$X_0 = \alpha Y_0 \tag{3.82}$$

where α is a positive constant. Without loss of generality we assume $\alpha > 1$ and, as before, we set $A_0 = X_0 Y_0$. Then (3.12) can be written as

$$E[D] = \frac{(\alpha + 1)\sqrt{A_0}}{3\sqrt{\alpha}} = \frac{2}{3}\sqrt{A_0} + \frac{(\sqrt{\alpha} - 1)^2}{3\sqrt{\alpha}}\sqrt{A_0} \tag{3.83}$$

The second term in (3.83) is the amount by which $E[D]$ deviates from its minimum value in (3.79). For $\alpha = 1.5$ that term becomes equal to $0.014\sqrt{A_0}$ (i.e., $E[D]$ is only about 2 percent greater than its minimum value). Even for $\alpha = 4$, $E[D]$ is only 25 percent more than its minimum value. An entirely similar analysis can demonstrate the robustness of (3.80).

Results such as those of (3.79) and (3.80) can be derived for various district shapes. The first three columns of Table 3-1 summarize the equivalents of (3.79) for a square district, a square district rotated by 45° with respect to the right-angle directions of travel, and a circular district. The following four cases are included:[8]

1. Euclidean (straight-line) travel when the response unit is randomly and uniformly positioned in the district.
2. Case 1 with right-angle travel.
3. Euclidean travel with the response unit located at the center of the district.
4. Case 3 with right-angle travel.

In all cases it is assumed that the locations of requests for service are uniformly distributed in the district and independent of the location of the service unit. When the constants in Table 3-1 are multiplied by $\sqrt{A_0}$, the square root of the area of the district in question, $E[D]$ is obtained. In some instances (e.g., a square district with a randomly positioned response unit and Euclidean travel) the constant of interest is not known exactly and the best known approximation, to two-decimal-place accuracy, is shown. Some of these constants have already been derived in this chapter or will be derived in the Problems.

[8]A few results for metrics other than Euclidean or right-angle are derived in the Problems.

TABLE 3-1 Proportionality constants for determining mean travel distances.

	Metric in Use \ *Shape of District*	*Square*	*Perfect, Four-Sided Diamond*[1]	*Circle*	*Approximation for "Fairly Compact and Fairly Convex" Areas*
Response unit is randomly positioned in the district	Euclidean travel	0.52	0.52	$\frac{128}{45\pi\sqrt{\pi}} \doteq 0.511$	0.52
	Right-angle travel	$\frac{2}{3} \doteq 0.667$	$\frac{14\sqrt{2}}{30} \doteq 0.660$	$\frac{4\cdot 128}{\pi\cdot 45\pi\sqrt{\pi}} \doteq 0.650$	0.67
Response unit is located at the center of the district	Euclidean travel	$\frac{\sqrt{2} + \ln(1+\sqrt{2})}{6} \doteq 0.383$	$\frac{\sqrt{2} + \ln(1+\sqrt{2})}{6} \doteq 0.383$	$\frac{2}{3\sqrt{\pi}} \doteq 0.376$	0.38
	Right-angle travel	$\frac{1}{2} = 0.5$	$\frac{\sqrt{2}}{3} \doteq 0.471$	$\frac{4\cdot 2}{\pi\cdot 3\sqrt{\pi}} \doteq 0.479$	0.50

[1]Square rotated by 45° with respect to the directions of travel in the case of right-angle travel.

The three district geometries included in Table 3-1 are "special cases" of rectangular, diamond-shaped, and elliptic districts. If one varies the district dimensions of each type while constraining district area to equal a constant A_0, $E[D]$ is minimized by the symmetric geometries represented in Table 3-1.

It can be seen from Table 3-1 that, for any given district area A_0, $E[D]$ is very insensitive to the exact geometry of the district. This can be confirmed by deriving $E[D]$ for other possible district geometries, such as equilateral triangles or piece-of-pie-like sectors of circles. Moreover, for any given district geometry, the value of $E[D]$ is insensitive to changes of the dimensions of the district that might make it appear to deviate appreciably from its optimum shape. This, too, can be confirmed by performing a sensitivity analysis similar to the one for the rectangular district in Example 12.

From these observations it can be concluded that we can use the first three columns of Table 3-1 to infer similar approximate expressions for $E[D]$ that apply to districts of *any shape* as long as (1) one of the dimensions (e.g., "length") is not much greater than the other dimension (e.g., width), and (2) major barriers or boundary indentations do not exist in the district. Districts that satisfy both of the conditions above will be called here, informally, "fairly compact and fairly convex districts." We can now state the following:

For fairly compact and fairly convex districts and for independently and spatially uniformly distributed requests for service,

$$E[D] \doteq c \cdot \sqrt{A_0} \tag{3.84}$$

where A_0 is the area of the district and c is a constant that depends *only* on the metric in use and on the assumption regarding the location of the response unit in the district.

The last column of Table 3-1 lists values that can be used for c in (3.84) for the four combinations of response unit locations and metrics that we have examined here. In all cases, we have selected the largest value of c listed in each row of the three leftmost columns of Table 3-1.

When the effective travel speed is independent of the distance covered, one can use the constants in the fourth column of Table 3-1 to approximate the expected travel time, $E[T]$, as well. In that case we have

$$E[T] \doteq \frac{c}{v} \sqrt{A_0} \tag{3.85}$$

in the case of Euclidean travel (assuming that the effective travel speed v is independent of the direction of travel) and

$$E[T] \doteq c\sqrt{\frac{A_0}{v_x v_y}} \tag{3.86}$$

for right-angle travel. In this latter case, the district "compactness" statement requires that

$$E[T_{\text{east-west}}] \approx E[T_{\text{north-south}}]$$

That is, it takes *on the average* about as much time to traverse the district from east to west as from north to south.

Another observation that can be made on the basis of the foregoing discussion is that both $E[D]$ and $E[T]$ are proportional to the square root of the district area, A_0, irrespective of the specific distance metric in use. This is hardly surprising since this relationship is basically a dimensional one: distance is the square root of area. More formally, if the coordinates of each point (x, y) in the district of interest are multiplied by $\sqrt{m}$ $(m > 1)$ [i.e., point (x, y) now becomes point $(\sqrt{m}\,x, \sqrt{m}\,y)$], then the area of the district increases m-fold but the length, L, of *any given* route between the pair of points (x_1, y_1) and (x_2, y_2)—in the original district—becomes equal to $\sqrt{m}\,L$ in the expanded district.

Equivalently, we can state that $E[D]$ and $E[T]$ must be proportional to the inverse of the square root of the *density* of response units in a district, for districts with more than one response unit. That is, if a district of area A is divided into n approximately equal fairly convex and fairly compact subdistricts of responsibility (whose shapes may vary), then

$$E[D] \doteq c\sqrt{\frac{A_0}{N}} = \frac{c}{\sqrt{\gamma}} \tag{3.87}$$

where γ denotes the spatial density of service units. We shall derive the same functional type of relationship in a somewhat different context later in this chapter [cf. (3.101a) and (3.104a)].

3.7.2 More Realistic Travel-Time Model

In most practical situations, the effective travel speed of urban response units depends on travel distance: longer trips, in general, are taken at a higher average speed than are shorter trips. It is therefore desirable to develop expressions for $E[T]$ that take into consideration some types of functional relationships between travel time and travel distance [unlike expressions (3.81) and (3.86), which assumed that effective travel speed remains constant with distance].

One plausible model is the following. Let us assume that urban service vehicles responding to a call, first go through an acceleration stage (perhaps while maneuvering their way through side streets, turns, etc.) until they reach a cruising speed that they maintain through the middle stage of the trip (while, perhaps, traveling on highways, thoroughfares, etc.) up to the final stage of it, during which they decelerate to a stop. Let us further assume that during the initial and final stages, vehicles accelerate (or decelerate) at a constant rate of a miles/min² and that during the middle stage, travel is at a constant cruising speed of v_c miles/min.

For trips of length less than $2d_c$ (where $d_c = v_c^2/2a$ is the distance needed to reach cruising speed) the cruising speed will never be reached; this is not the case when the travel distance D is greater than $2d_c$. Using the well-known physical relationships for accelerated and constant speed travel ($D = at^2/2$ and $D = vt$), it is then easy to conclude that the *conditional* expected travel time $E[T \mid D = d]$ for any given travel distance is

$$E[T \mid D = d] = 2\sqrt{\frac{d}{a}} \qquad \text{for } d \leq 2d_c \qquad (3.88)$$

$$E[T \mid D = d] = \frac{d - 2d_c}{v_c} + \frac{2v_c}{a} = \frac{d}{v_c} + \frac{v_c}{a} \qquad \text{for } d > 2d_c \qquad (3.89)$$

One can obviously think of many other physical scenarios that would lead to different expressions for $E[T \mid D = d]$. A considerable amount of field data, however, suggests that (3.88) and (3.89) often provide truly excellent approximations for many urban services—see, for instance, [KOLE 75, JARV 75, HAUS 75].

An expression for the unconditional expected travel time, $E[T]$, can now be written:

$$E[T] = \int_0^{\infty} E[T \mid D = x] f_D(x)\, dx = \int_0^{2d_c} 2\sqrt{\frac{x}{a}} f_D(x)\, dx + \int_{2d_c}^{\infty} \left(\frac{v_c}{a} + \frac{x}{v_c}\right) f_D(x)\, dx \qquad (3.90)$$

In order to evaluate the two integrals in (3.90) it is necessary to know the pdf for the travel distance, $f_D(x)$.

Example 13: Expected Travel Time in a Square District

Consider a 1- by 1-mile-square fire district with a firehouse located at its center. Fire alarms are distributed independently and uniformly within the district and the travel metric is right-angle with directions of travel parallel to the sides of the square. Measurements have shown that the "cruising speed" for fire engines is $v_x = v_y = v_c = 30$ miles/hr and vehicle acceleration and deceleration (as described above) is $a = 0.5$ miles/min²; that is, it takes, on the average, about 1 minute of travel for the fire engines to accelerate up to (or decelerate down from) cruising speed. (These values are rather typical, as field data show—see below.) It follows that $d_c = v_c^2/2a = 0.25$ mile.

From (3.88) and (3.89) we then have (in minutes)

$$E[T \mid D = d] = \begin{cases} 2^{3/2}\sqrt{d} & \text{for } d \leq 0.5 \text{ mile} \\ 2d + 1 & \text{for } d > 0.5 \text{ mile} \end{cases}$$

From the fact that D_x and D_y, the distances traveled in the east-west and north-south directions, respectively, are independent random variables distributed uniformly between 0 and $\frac{1}{2}$ mile, it is easy to show, using the techniques of this chapter (see Problem 3.9), that for $D = D_x + D_y$ we have

$$f_D(x) = \begin{cases} 4x & \text{for } 0 \leq x \leq \frac{1}{2} \\ 4(1 - x) & \text{for } \frac{1}{2} \leq x \leq 1 \end{cases} \qquad (3.91)$$

Substituting for $E[T \mid D = x]$ and $f_D(x)$ in (3.90), we then obtain

$$E[T] = \int_0^{1/2} 2^{3/2}\sqrt{x} \cdot 4x \cdot dx + \int_{1/2}^{1} (2x + 1)4(1 - x)\, dx = 1.97 \text{ minutes}$$

for the average travel time in responding to a fire alarm in this district.

Unfortunately, the pdf for the travel distance $f_D(x)$ is often difficult to obtain, either theoretically or from field data. The following approximate expression for $E[T \mid D = d]$ is then often used in order to overcome this problem:

$$E[T \mid D = d] = \frac{d}{v_c} + \frac{v_c}{a} \qquad \text{for } d \geq 0 \qquad (3.92)$$

This expression is compared with (3.88) and (3.89) in Figure 3.30. Note that (3.92) is a "conservative" model for $E[T \mid D = d]$ in the sense that it provides an upper bound for (3.88) and (3.89) and is also a good approximation to it for *all* values of D, when $2d_c$ is relatively small by comparison to the distances that a response unit usually travels. The physical interpretation of (3.92) is also simple: a fixed amount of time ($= v_c/a$) is spent getting ready for each trip and then the trip takes place at a constant travel speed,[9] v_c.

Obviously, the advantage of (3.92) is that $f_D(x)$ is no longer necessary to develop a simple expression for the unconditional travel time $E[T]$. For we now have

$$T = \frac{D}{v_c} + \frac{v_c}{a}$$

[9] It is also possible to have this effective travel speed depend on the direction of travel (e.g., as in the case of right-angle travel).

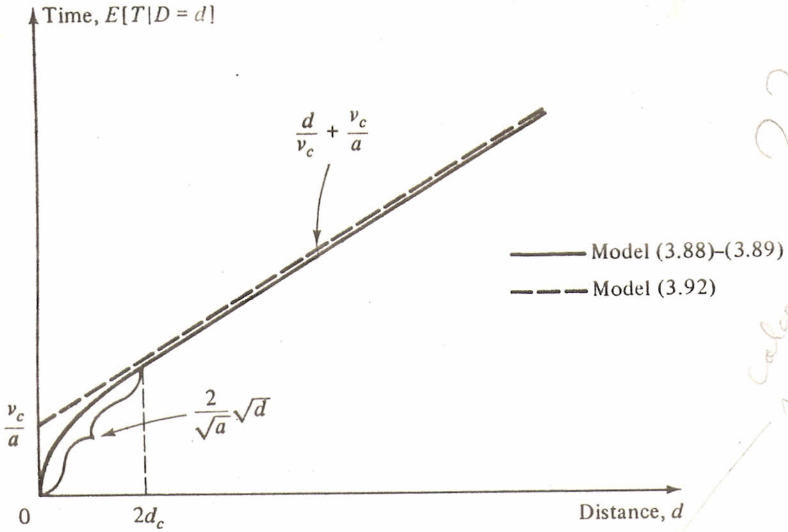


FIGURE 3.30 Comparison of two models of the expected travel time versus travel distance relationship.

and so

$$E[T] = \frac{E[D]}{v_c} + \frac{v_c}{a} \tag{3.93}$$

Example 13: (continued)

For our 1- by 1-mile-square-district example, it is obvious by inspection that $E[D] = 0.5$ mile (cf. Table 3-1). It follows that, with $v_c = 0.5$ mile/min and $a = 0.5$ mile/min^2, $E[T] = 2$ minutes, or only about 2 seconds more than our earlier exact estimate! The very close agreement between the two estimates may seem surprising in view of the fact that $2d_c = 0.5$ mile in this case, or 50 percent of the full range of values of the travel distance D.

The example above is not atypical. Estimates of $E[T]$ obtained through (3.92) and (3.93) are usually very close to estimates obtained through the time-consuming approach summarized by (3.88)–(3.90) for the values of v_c and a one encounters in urban service applications.

Table 3-2 lists four sets of measured values of v_c and a for four different cities in the United States. The data were collected by the fire departments in these cities [HAUS 75]. Note the similarity of the values for the cruising speeds. From Table 3-2 it can also be inferred that the constant v_c/a typically adds about 0.5 to 1.0 minute to $E[T]$ in (3.93).

In this section we have thus concluded that:

TABLE 3-2 Cruising speeds and accelerations.[1]

	v_c (miles/hr)	a (miles/min^2)
New York, N.Y.	39.1	0.48
Trenton, N.J.	34.8	1.09
Yonkers, N.Y.	33.9	0.86
Denver, Colo.	39.2	1.05

[1]Values of the cruising speed, v_c, and acceleration, a for four fire departments in the United States [HAUS 75].

The travel time/travel distance relationship (3.92) greatly simplifies the calculation of $E[T]$ and provides excellent approximations to results produced through more complicated analyses.

It should be noted that our earlier results regarding variation of $E[T]$ with district geometries apply unaltered in the context of (3.93). This is because of the continued linear relationship between D and T, which is augmented in (3.93) only by an additive constant, a term that does not affect variational analyses of district geometries. Even when cruising travel speed depends on direction of travel, i.e., when $v_x \neq v_y$, the addition of a constant to the travel time will not affect the variations of travel times with district designs. Most important, optimal designs remain unchanged.

3.7.3 Expected Travel Distances: The General Case

Our discussion, so far, has focused primarily on exact and approximate expressions for expected travel distances and times to and from incidents in districts with relatively regular ("fairly compact and fairly convex") geometries and uniform distribution of incidents over the districts. Although this focus may appear, at first, to cover only a limited subset of the cases that one may encounter, it turns out that, in practice, our results can be used as "building blocks" to obtain good approximations in a large number of cases where incidents are not uniformly distributed and the district itself does not have a nice rectangular (or circular, triangular, etc.) shape.

Before illustrating this, let us first discuss, in the abstract, the most general possible cases. Let (X, Y) and (X_1, Y_1) indicate, respectively, the location of calls for service and of the response unit in a district R of area A. Denote by $f_{X,Y,X_1,Y_1}(x, y, x_1, y_1)$ the joint pdf for random variables X, Y, X_1, and Y_1, and by $D = d[(X_1, Y_1), (X, Y)]$ the mathematical relationship for the distance between (X_1, Y_1) and (X, Y) [e.g., $D = \sqrt{(X_1 - X)^2 + (Y_1 - Y)^2}$ for

Euclidean distances]. Then, for the expected travel distance in the district, we have

$$E[D] = \iiiint_{\text{over } R} d[(x_1, y_1), (x, y)] f_{X,Y,X_1,Y_1}(x, y, x_1, y_1)\, dx_1\, dy_1\, dx\, dy \quad (3.94)$$

Note that the joint pdf for the coordinates of the incident and of the service unit can be made to reflect not only nonuniformities in the distribution over R but also possible dependencies between the locations of incidents and of the service unit.

Expression (3.94) can be extended to the case where N response units are located in district R. Now let (X_i, Y_i) indicate the location of the ith response unit $(i = 1, 2, \ldots, N)$ and (X, Y) the location of an incident. Then the distance between the incident and the closest response unit can be written

$$D_N = \text{Min}\,\{d[(X_1, Y_1), (X, Y)], \ldots, d[(X_N, Y_N), (X, Y)]\}$$

Since D_N is then a function of the random variables $X, Y, X_1, Y_1, X_2, \ldots, X_N, Y_N$, we can write

$$E[D_N] = \int \cdots \int_{\text{over } R} \text{Min}\,\{d[(x_1, y_1), (x, y)], \ldots, d[(x_N, y_N), (x, y)]\} \cdot f_{X,Y,X_1,Y_1,\ldots,X_N,Y_N}(x, y, x_1, y_1, \ldots, x_N, y_N)\, dx\, dy \ldots dy_N \quad (3.95)$$

where $f_{X,Y,X_1,Y_1,\ldots,X_N,Y_N}(x, y, x_1, \ldots, y_N)$ is obviously the joint pdf for the coordinates of the incident and the N response units. Thus, in both (3.94) and (3.95) we have expressed expected travel distance as the expected value of a function of random variables whose joint pdf is known. The problem of computing the expected travel distance in the general case is, therefore, no more (or less) difficult than working with any other function of these random variables (cf. Section 3.1).[10]

Obviously, in practice, there are severe limitations on how far one can go in deriving such exact expressions for $E[D]$. Problems become mathematically intractable as the number of random variables increases or as the shape of R and/or the joint pdf for the random variables becomes more complex. In many cases, however, all is not lost as long as one is willing to settle for good approximations rather than exact results. This is true any time the response units are stationary at known locations, no matter what the number, N, of these units is (and for practically any pdf for the spatial distribution of incidents/demands as well as for any shape of the district of interest). It is also true, for any value of N, in the case of mobile response units as long as

[10] This approach can also be generalized to expected distances to other than the closest unit (e.g., to the kth closest unit).

subdistricts of responsibility have been defined in such a way that each subdistrict of R is served exclusively by a very small number of mobile units (preferably 1!). In such instances, the following three-step approach will always work:

STEP 1: Divide the district R into several (possibly many) nonoverlapping parts, which we shall call "zones." Each zone must have the following two properties:

a. Its shape must be approximately rectangular, triangular, circular, or any other easy-to-work-with configuration.

b. The pdf for the spatial distribution of incidents/demands within each zone must be approximately uniform (or that pdf can be approximated by some other sufficiently simple expression as to permit easy mathematical manipulation).

STEP 2: Using the techniques of this chapter, compute all intrazone and zone-to-response unit expected distances, as required by the problem at hand.

STEP 3: Multiply the expected distances computed in Step 2 by appropriate probabilities to obtain overall expected travel distances for district R.

Note that each zone in Step 1 can have an individual shape with its "own" pdf for the distribution of incidents. Note also that the greater the degree of accuracy desired, the larger the number of district zones should be (to approximate better the shape of the district R and the pdf for the spatial distribution of incidents). In fact, the three-step approach outlined above is very similar to the approach that a computer would follow in order to compute numerically the integrals in expressions (3.94) and (3.95).

Rather than attempt a more formal statement of the above three-step approach, we now illustrate it through the following example.

Example 14: Commuter Travel in a Suburban Town

Consider the suburban town shown in Figure 3.31. Its only access to the central business district (CBD) of the metropolitan area of which this town is a part is through the single bridge shown in Figure 3.31. The CBD is 6 miles from the bridge's end, as shown. Travel in the town is right-angle, as shown.

We are interested here in the total number of person-miles traveled by the town's working residents (not including schoolchildren) each morning on their way to work. (This information might be useful in transportation planning or in estimating transportation-related fuel consumption by commuters.)

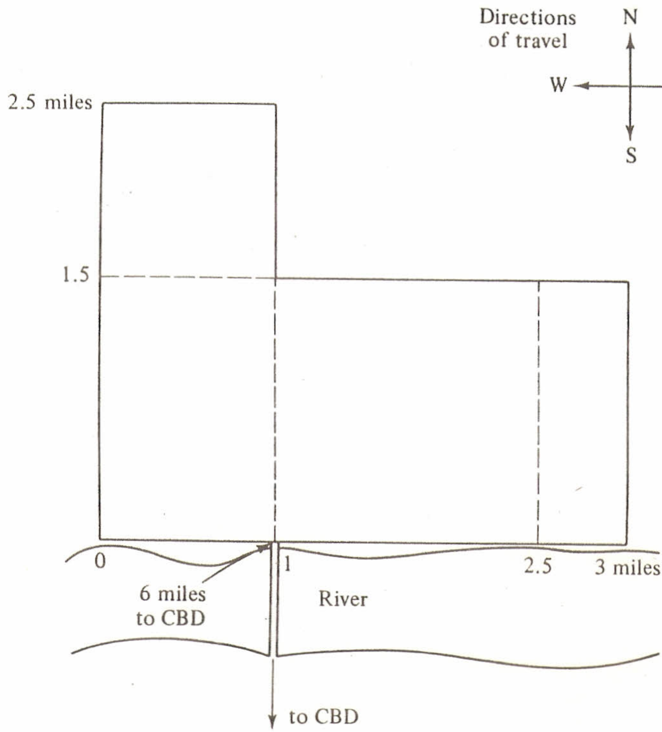


FIGURE 3.31 Configuration of a suburban city.

About 80 percent of the working residents work in the central city at the CBD. The other 20 percent work in town (and do not have to cross the bridge every morning). Trips are generated uniformly over the town at the rate of about 2,000 trips per square mile. The only exception is the rectangular area to the east of the 2.5-mile mark along the river (as shown in Figure 3.31), where the density of trips generated per square mile decreases linearly according to the function $g(d) = 4{,}000(3 - d)$, where d is the east-west coordinate $(2.5 \leq d \leq 3)$ of each point as measured from the southwesternmost point of the town (see Figure 3.31). There is no difference between the spatial distributions of trip origins to the CBD and to in-town jobs. That is, of every 100 trips generated at each part of the town, no matter where that part is located, 80, on the average, are to the CBD and 20 to in-town jobs.

The spatial distribution (and density per square mile) of in-town jobs is assumed identical to the distribution (and density) of trip-generating points for in-town jobs. (This may be the case, at least approximately, when there are no concentrations of places of employment in a city and when no major employers, such as factories, etc., are located there.) For the purposes of this example, we shall also make the more questionable assumption that the job and residence locations for in-town workers are statistically independent

(i.e., that knowledge of where an in-town worker's home is does not affect our a priori knowledge of where in town he or she works, and vice versa).

Solution

In working on this problem, we shall first compute the expected travel distance for CBD workers, then the expected travel distance for in-town workers, and finally the total passenger miles covered per day.

To start with, we need a coordinate system. Although our choice of origin does not really make much difference in this case, the edge of the bridge on the town's side is a particularly convenient one. We thus relabel the various points of interest according to this choice of origin, as shown in Figure 3.32a. We can also, using the information given, construct the pdf for the spatial distribution of trip-generation points.

Exercise 3.9 Show that this pdf is given by

$$f_{X,Y}(x, y) = \begin{cases} \frac{8}{41} & \text{for } -1 \leq x \leq 0,\ 0 \leq y \leq 2.5 \\ \frac{8}{41} & \text{for } 0 \leq x \leq 1.5,\ 0 \leq y \leq 1.5 \\ \frac{16}{41}(2 - x) & \text{for } 1.5 \leq x \leq 2,\ 0 \leq y \leq 1.5 \\ 0 & \text{otherwise} \end{cases}$$

Note that $f_{X,Y}(x, y)$ also represents the pdf for the spatial distribution of in-town jobs according to the problem statement.

With these preliminaries we can now compute:

1. **$E[D]$ for CBD workers.** Since the coordinates of the edge of the bridge that travelers to the CBD must reach are (0, 0), the distance from any point with coordinates (X, Y) to the bridge is given by $D = |X| + |Y|$.

Exercise 3.10 Show that if we define $Z = |X|$ and $W = |Y|$, then $f_Z(z)$ and $f_W(w)$ are as shown in Figure 3.32b and c. Note that both pdf's can be derived almost by inspection by first obtaining $f_X(x)$ and $f_Y(y)$ from $f_{X,Y}(x, y)$. In doing so we use the geometrical probability interpretation of pdf's (cf. Section 3.4.1).

It is now easy to obtain

$$E[D] = E[|X|] + E[|Y|] = E[Z] + E[W] = \frac{99}{164} + \frac{163}{164} \simeq 1.60$$

miles for the expected distance to point (0, 0).

2. **$E[D]$ for in-town workers.** We now partition the town into four non-overlapping zones, as shown in Figure 3.32d. We wish, in effect, to compute $E[D]$ between two random points in the town with the locations of each point determined independently, each according to the pdf $f_{X,Y}(x, y)$. To do this we consider all possible intrazone and interzone expected distances and then multiply each expected distance by the appropriate probability.

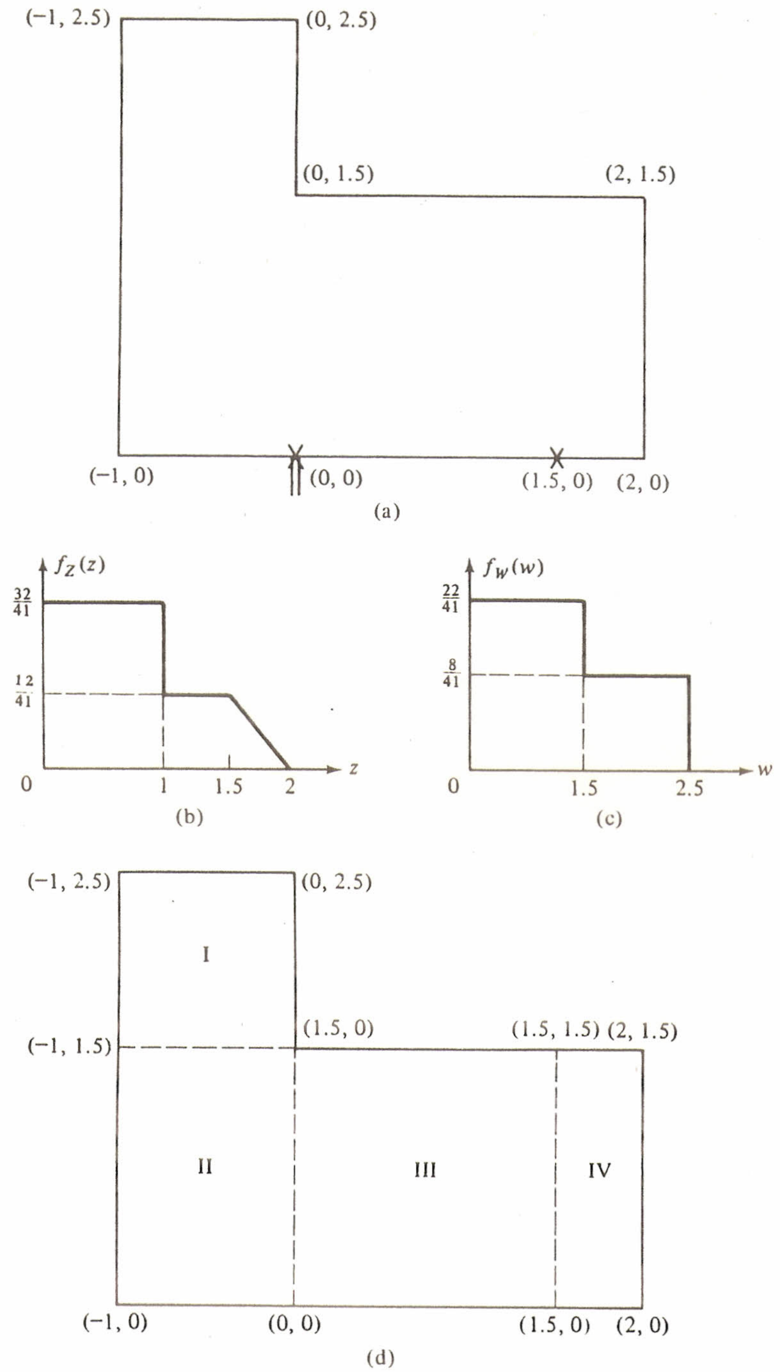


FIGURE 3.32 (a) Coordinates of corner points of suburban city; (b) The pdf for Z; (c) The pdf for W; (d) partitioning of the suburban city.

For instance, it can be seen that, given an in-town worker:

$$P\{\text{both residence and place of work are in zone I}\} = \left(\frac{8}{41}\right)^2 = \frac{64}{1{,}681}$$

$$E[D \mid \text{both residence and place of work are in zone I}] = \tfrac{1}{3} \text{ mile}$$

Exercise 3.11 Show that if both the residence and the place of work of an in-town worker are in zone IV, his or her conditional expected travel distance is equal to $\frac{19}{30}$ mile.

Exercise 3.12 By carefully considering all residence and place-of-work combinations, show that for in-town workers, $E[D] \simeq 1.655$ miles.

3. **Total expected distance.** A total of about 10,250 trips take place every morning. Of those 80 percent (= 8,200) are to the CBD and 20 percent (= 2,050) are in-town. The expected travel distance to a CBD trip is 7.60 miles [remember that point (0, 0) is 6 miles from the CBD] while an in-town trip is 1.655 miles long on the average. Therefore, the total expected distance traveled by workers each morning is 65,755 person-miles.

It should be clear that the problem of determining $E[D]$ for CBD workers was equivalent to computing $E[D]$ between an incident distributed as $f_{X,Y}(x, y)$ in the city and a fixed service unit located at the CBD. Similarly, $E[D]$ for in-town workers is equivalent to the expected travel distance between an incident spatially distributed as $f_{X,Y}(x, y)$ in the town and a mobile response unit with that same distribution for its location in the town.

Finally, we might, out of curiosity, wish to compare the result of Exercise 3.12 for the expected travel distance for *in-town* workers with the result that we would have obtained had we used the approximate expression (3.84) with $c = 0.67$ (Table 3-1), disregarding the fact that the shape of the town of interest is not quite "fairly compact and fairly convex" and that in a part of the town the distributions of demand and of the "service unit" (i.e., of the job locations) are not uniform. Since the area of the town is 5.5 square miles, we have $E[D] \approx 0.67\sqrt{5.5} \approx 1.57$ miles, for an error of about 5 percent! The reader who worked through Exercises 3.10–3.12 to obtain the exact result of 1.655 will definitely appreciate now the value of approximate expression (3.84).

3.8 SPATIAL POISSON PROCESSES

3.8.1 Description and Postulates

Suppose that we have entities distributed around the city in a completely random manner. These entities could be employees of a particular service system, recipients of a certain social service, emergency response units, crimes, and so on. We require a way of describing probabilistically the numbers of entities in given subareas and spatial interrelationships among entities. To do this, we generalize the idea of a Poisson process in time to a Poisson process in space.

We recall from Chapter 2 that for a homogeneous Poisson process in time, the probability that exactly k Poisson events occur in a fixed time interval $[0, t]$ is

$$P\{X(t) = k\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \qquad \text{for } t \geq 0, k = 0, 1, 2, \ldots$$

where λ is a positive constant interpreted as the average rate at which events are happening per unit time. The process is called "homogeneous," because λ does not vary with time.

Applying the same ideas in a spatial setting, first consider a homogeneous highway segment of length l miles. From past accident records we may know that each year an average of λ highway accidents occur per mile on this type of highway. Then the number of highway accidents that occur in the segment of length l miles can be modeled as a Poisson random variable with mean λl. Here the parameter distance (l) plays a role directly analogous to time (t). For the Poisson model to be a reasonable one, the locations of accidents must occur consistent with Poisson-type assumptions: (1) only nonnegative integer numbers of accidents can occur in any length of highway; (2) the probability distribution of the number of accidents depends only on the length of highway considered, and as this length goes to zero so does the probability of an accident occurring there; (3) the numbers of accidents occurring in nonoverlapping segments of highway are mutually independent random variables; and (4) given that an accident occurs at a particular location, the chance of a second accident occurring at the identically same location is zero. Assumptions (1) and (3) appear fairly reasonable for most highways. Since different parts of a highway (e.g., curves versus straight-aways) can be associated with different risks of accident, the first part of assumption (2) may have to be modified in practice to allow for a *spatially varying* (nonhomogeneous) Poisson process, with accompanying $\lambda(x)$ defined so that $\lambda(x)\,dx$ = probability of an accident occurring (during a year) in the road interval x to $x + dx$. A highway with overpasses, bridges, and other discernible high-risk points may yield a positive probability of at least one accident during a year at these points (e.g., at the base of an overpass), thereby negating the second part of assumption (2). Such high-risk points would also tend to negate assumption (4), which also might be invalidated by chain-reaction multiple-car accidents (such as those that occasionally occur in fog) if those are counted in accident data as more than one accident.

In practice, almost any real system will demonstrate a nonperfect degree of conformity with the postulates of the Poisson process. In assessing the

applicability of the Poisson model, the modeler must weigh the benefits of applying the Poisson model (together with the insights it provides) against the cost of inaccuracies introduced by such a simple model and the cost of constructing a more complex model. Sometimes a forced degree of ignorance involving details of a system (e.g., the locations of overpasses) will facilitate application of the model.

The ideas illustrated above for a Poisson process on a line (a highway) extend directly to the plane (to describe entities distributed over two dimensions). Let the parameter S denote a bounded region of the plane (or higher-dimensional space, for that matter). Let $X(S)$ be the number of entities contained in S. Then $X(S)$ is a homogeneous spatial Poisson process if it obeys the Poisson postulates, yielding a probability distribution

$$P\{X(S) = k\} = \frac{[\lambda A(S)]^k e^{-\lambda A(S)}}{k!} \qquad \text{for } A(S) \geq 0, \quad k = 0, 1, 2, \ldots \tag{3.96}$$

In this case λ is a positive constant called the *intensity parameter* of the process and $A(S)$ represents the area or volume of S, depending on whether S is a region in the plane or higher-dimensional space.

The underlying mathematical postulates of the model follow directly those of the time Poisson process:

1. Only nonnegative integer values are assumed by $X(S)$ and $0 < P\{X(S) > 0\} < 1$ if $A(S) > 0$.
2. The probability distribution of $X(S)$ depends on S only through the value of $A(S)$ with the further property that if $A(S) \to 0$, then $P\{X(S) \geq 1\} \to 0$.
3. If $S_1, S_2, \ldots, S_n$ $(n \geq 1)$ are disjoint regions, then $X(S_1), \ldots, X(S_n)$ are mutually independent random variables and $X(S_1 \cup \ldots \cup S_n) = X(S_1) + \ldots + X(S_n)$.
4. $\displaystyle\lim_{A(S)\to 0} \frac{P\{X(S) \geq 1\}}{P\{X(S) = 1\}} = 1.$

Generalization: As with the time Poisson process, it is not difficult to extend these ideas to a spatially varying (nonhomogeneous) Poisson process. For instance, in the plane if

$$\lambda(x, y)\,dx\,dy = P\{\text{a Poisson entity is located in the interval } x \text{ to } x + dx, y \text{ to } y + dy\}$$

then in (3.96), $\lambda A(S)$ is replaced by

$$\iint_S \lambda(x, y)\,dy\,dx \equiv \lambda(S) \tag{3.97}$$

In this case postulate (2) is changed to read: "The probability distribution of $X(S)$ depends on S only through the value of $\lambda(S)$ with the further property that as $\lambda(S) \longrightarrow 0$, then $P\{X(S) \geq 1\} \longrightarrow 0$."

3.8.2 Time–Space Poisson Process

Suppose that in some region of space S with area $A(S)$ events occur in time as a Poisson process with rate $\lambda A(S)$ per unit time. Then utilizing the foregoing ideas about multidimensional Poisson processes, the probability that k events occur in S in time t is

$$P\{X(S, t) = k\} = \frac{[\lambda t A(S)]^k e^{-\lambda t A(S)}}{k!} \qquad \text{for } k = 0, 1, 2, \ldots \qquad (3.98)$$

Problem 3.26 applies this concept.

Example 15: Distribution of Travel Distance ("Nearest Neighbor")

Suppose that emergency response units are distributed throughout a large region as a two-dimensional Poisson process with intensity parameter γ units per square mile. We wish to know the pdf of the travel distance D between an incident, whose position is selected independently of response unit positions, and the *nearest* response unit. Assume Euclidean travel distance. (This is sometimes known as a "nearest-neighbor" problem; in three-dimensional space this problem has been used to determine the distribution of distance between stars in a galaxy.)

Solution

We use the never-fail cumulative distribution method in conjunction with our new knowledge of spatial Poisson processes.

1. Assume the incident occurs at some arbitrary point (x, y).
2. Construct a circle of radius r about (x, y).
3. The probability that there are exactly k response units within the circle is

$$P\{X(\text{circle}) = k\} = \frac{(\gamma \pi r^2)^k e^{-\gamma \pi r^2}}{k!} \qquad k = 0, 1, 2, \ldots$$

4. Therefore, we obtain the cdf by the following reasoning:

$$F_D(r) \equiv P\{D \leq r\} = 1 - P\{D > r\} = 1 - P\{X(\text{circle}) = 0\}$$

or

$$F_D(r) = 1 - e^{-\gamma \pi r^2} \qquad r \geq 0 \qquad (3.99)$$

5. The pdf is

$$f_D(r) = \frac{d}{dr} F_D(r) = 2r\gamma\pi e^{-\gamma\pi r^2} \qquad r \geq 0 \qquad (3.100)$$

This is a Rayleigh pdf with parameter $\sqrt{2\gamma\pi}$. Thus, the mean and variance are

$$E[D] = \tfrac{1}{2}\sqrt{\gamma^{-1}} \qquad (3.101a)$$

$$\sigma_D^2 = \left(2 - \frac{\pi}{2}\right)\frac{1}{2\pi\gamma} \qquad (3.101b)$$

Question: How could you extend these ideas to obtain other interesting properties of the system?

Example 16: Nearest Neighbor with Right-Angle Travel Distance

If travel distance is right-angle, rather than Euclidean, the analysis in Example 15 follows straight through, except instead of a circle of radius r we have a square rotated at 45°, centered at (x, y), with area equal to $2r^2$. Following the same steps in the solution,

$$P\{X(\text{square}) = k\} = \frac{(2\gamma r^2)^k e^{-2\gamma r^2}}{k!} \qquad k = 0, 1, 2, \ldots$$

$$F_D(r) = 1 - e^{-2\gamma r^2} \qquad r \geq 0 \qquad (3.102)$$

$$f_L(r) = \frac{d}{dr} F_D(r) = 4\gamma r e^{-2\gamma r^2} \qquad r \geq 0 \qquad (3.103)$$

This is a Rayleigh pdf with parameter $\sqrt{4\gamma}$. The mean and variance are

$$E[D] = \frac{1}{4}\sqrt{\frac{2\pi}{\gamma}} \qquad (3.104a)$$

$$\sigma_D^2 = \left(2 - \frac{\pi}{2}\right)\frac{1}{4\gamma} \qquad (3.104b)$$

Question: In Example 4 in this chapter we derived that in an isotropic environment a response unit traveling according to the right-angle distance metric travels $4/\pi = 1.273$ times farther (on the average) than a unit traveling "as the crow flies." Thus, one might be tempted to think that the ratio of the mean right-angle to Euclidean distances computed in Examples 15 and 16 would be 1.273. In fact, the ratio is $\sqrt{\pi/2} < 4/\pi$. Why?

Hint: See Problems 3.9 and 3.10.

Further work: Problems 3.25 and 3.26.

3.8.3 Application to Facility Location and Districting

One could apply the ideas of spatial Poisson processes to a problem of facility location and districting of a city. Suppose that demands are distributed uniformly throughout the plane and suppose that travel distance is right-angle. We can consider two applications: (1) for each service request, a response unit is dispatched from the *nearest* facility (the service system need not be an emergency service; for instance, it could be a social service agency whose personnel make home visits); and (2) the individual requiring service travels to the nearest facility (e.g., hospital, library, "little city hall," police district station house). Each *district* about a facility would consist of all points closer to that facility than to any other.

As an agency administrator, you want to get some idea of the potential benefits (in terms of mean travel distance reduction) of a study to optimally locate service facilities.

Use of Poisson model to generate "upper bound." At one extreme, you could ask: What are the response distance characteristics of the system if facilities are distributed *at random*? We can answer this question by assuming that *facilities* are distributed as a homogeneous spatial Poisson process. This corresponds to a totally unplanned system (in terms of districting) in which the facility locations could be viewed as occurring from "throwing darts blindfolded" at a map of the city. That is, given n facilities in any particular region, their locations would be independently, uniformly distributed over the region (following the "unordered arrival times" argument of Chapter 2 for a time Poisson process).

For a right-angle distance metric, a random distribution of facilities may yield a city-wide districting as shown in Figure 3.33. Using the result of Example 16, the mean travel distance is

$$E[D] = \frac{1}{4}\sqrt{\frac{2\pi}{\gamma}} \simeq 0.627\gamma^{-1/2} \qquad (3.105)$$

where γ is the average density of facilities.

Lower bound. To achieve minimal mean travel distance, the facilities should be positioned in a regular lattice, as shown in Figure 3.34. This makes intuitive sense since a diamond gives the set of points within a given distance of its center, when right-angle distance is used (analogous to a circle for Euclidean distance), so diamonds can be used to partition a city into districts of equal coverage, where coverage of a district is measured by maximum possible distance from its facility.

We prove the desired result regarding $E[D]$ in two steps:

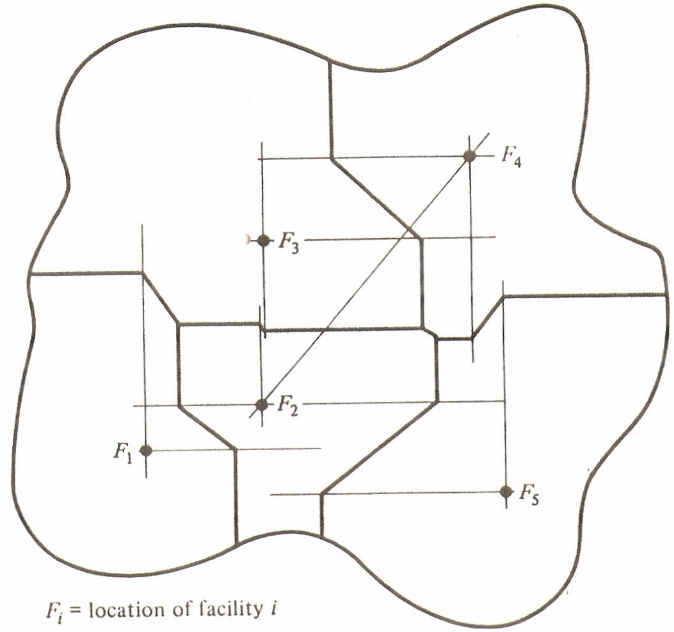


FIGURE 3.33 Illustrative districting with a random positioning of facilities.

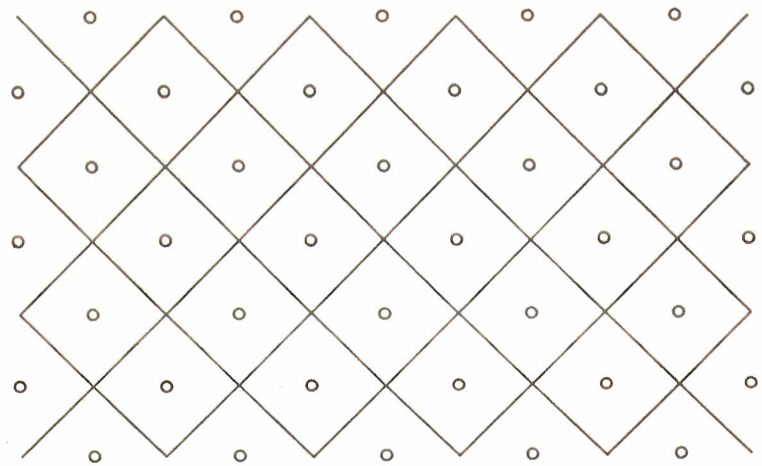


FIGURE 3.34 Regular lattice of optimal facility locations and districts.

STEP 1: Given that a facility's district must contain an area A, a square district rotated at 45°, centered at the facility's position, results in minimum mean travel distance ($E[D]$).

Proof: (Contradiction, using perturbation method.) Suppose there is some redesign of the rotated square that results in lower mean travel distance.

Then the new district can be constructed by taking a set of points B_1 of area ϵ out of the rotated square and adding a set of points B_2 of area ϵ from outside the square. The situation is shown in Figure 3.35. Then, in the redesigned district, the new mean travel distance is

$$E[D'] = E[D] - \frac{\epsilon}{A}E[D \mid B_1] + \frac{\epsilon}{A}E[D \mid B_2]$$

or

$$E[D'] - E[D] = \frac{\epsilon}{A}(E[D \mid B_2] - E[D \mid B_1])$$

But

$$E[D \mid B_1] \leq \sqrt{\frac{A}{2}}$$

$$E[D \mid B_2] \geq \sqrt{\frac{A}{2}}$$

so $E[D'] - E[D] \geq 0$, which is a contradiction.

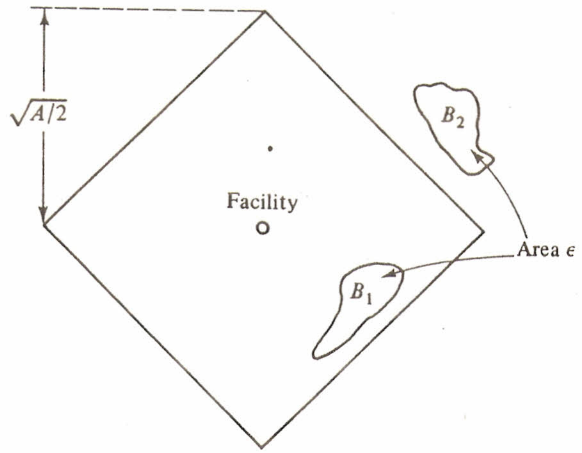


FIGURE 3.35 Exchange of subregions of area ϵ.

STEP 2: Given that we have N square districts, each rotated at 45° and centered at the respective facility's position, and given that the total area of the N districts must equal NA, minimal mean travel time is obtained by setting the area of each district equal to A.

Proof: For a random service request located in one of the N districts, assuming that district i has area A_i, the mean travel distance is

$$E[D] = \sum_{i=1}^{N} \frac{A_i}{NA} \frac{2}{3}\sqrt{\frac{A_i}{2}} \qquad (3.106)$$

(Why?) The idea is to minimize (3.106) subject to the total area constraint, $\sum_{i=1}^{N} A_i = NA$. This is a straightforward problem of constrained optimization. Using Lagrange multipliers, one finds that the minimal $E[D]$ is found by setting $A_i = A$ (all i). Finally (and fortunately!), equal-sized square districts rotated at 45° fit into the lattice shown in Figure 3.34. For this lattice

$$E[D] = \frac{2}{3}\sqrt{\frac{1}{2\gamma}} \simeq 0.472\gamma^{-1/2} \qquad (3.107)$$

Comparing this to the mean travel distance with *randomly* positioned facilities, we have the somewhat surprising result that optimal positioning (and districting) reduces mean travel distance over that obtained by random positioning by only about *25 percent*.

What are the policy implications of this result?

3.9 ALTERNATIVE SPATIAL PROCESSES

The spatial Poisson process has a "no memory" property similar to that of the time Poisson process. In this case, the existence or nonexistence of a Poisson entity in any region of space does not influence the likelihood of other Poisson entities existing in nearby disjoint regions. Moreover if we know that there are n points distributed in a fixed region of area A and that these points were generated from a spatial Poisson process, then the n points are independently uniformly distributed over the region. This is simply the two-dimensional generalization of the "unordered arrival times" argument made in Chapter 2 for Poisson processes in time.

However, many naturally occurring processes do not adhere to the Poisson assumptions. For instance, one can imagine certain processes for which the existence of one point would increase the likelihood of other points occurring nearby. Such "clustering" processes could include hospitals (which cluster due to economies of scale), certain industries, police cars, crimes, households having certain demographic characteristics, and so on. For these processes the spatial Poisson process is an inadequate model. Similarly, one can imagine other processes for which the existence of one point would decrease the likelihood of other points occurring nearby. Such "spread" processes could include certain retail establishments (e.g., supermarkets, hamburger havens), urban service facilities (e.g., libraries, outpatient clinics, "little city halls"), and street intersections.

The question is, how do we model such processes? The answer, at present, is that models for such complicated spatially dependent processes are quite inadequate. While we can quite successfully generalize from Poisson processes to renewal processes in time, there does not seem to be an analogous generalization for spatial processes. Still, urban geographers have devised various techniques for tackling this problem, and here we illustrate a popular one based on partitioning the city into a regular lattice of equal size small cells.

Suppose that we are interested in the number of points N in a particular cell of unit area. If the points are distributed according to a spatial Poisson process with parameter γ (points/unit area), then

$$P\{N = k\} = \frac{\gamma^k e^{-\gamma}}{k!} \qquad k = 0, 1, 2, \ldots$$

$$E[N] = \gamma$$

$$\sigma_N^2 = \gamma$$

In particular, we focus on the ratio of the variance to the mean,

$$r = \frac{\sigma_N^2}{E[N]} \tag{3.108}$$

For a spatial Poisson process, $r = 1$. Now, for a "spread" process in which the existence of one point reduces the chance of another point being located nearby, assuming that means are kept constant, one would expect that the Poisson distribution could be modified in a way that would take probabilities away from the tails of the Poisson distribution and add these to probabilities near the mean (center) of the distribution. Such a modification would have the effect of reducing the variance of the distribution. Thus, any spatial process for which

$$r = \frac{\sigma_N^2}{E[N]} < 1$$

is called a *spread process* (meaning points tend to be "spread out" over the plane). An extreme case occurs with a perfectly regular lattice of points which provides each geographical cell with exactly the same number of points; then $\sigma_N^2 = 0$, so that $r = 0$. A spread process, sometimes also called *regular process*, because of its closeness to a regular lattice, includes the class of processes ranging from a perfect lattice of points to (but not including) the Poisson process.

A clustered process, on the other hand, would probably have many cells with zero points and others with more than predicted by the Poisson model. Thus, again keeping means constant, one could obtain a pmf for a clustered process by taking probability away from the integer values near the mean of the Poisson process and adding this to values at or near zero and to values in the positive tail of the distribution. This would have the effect of increasing the variance of the distribution. Thus, any spatial process for which

$$r = \frac{\sigma_N^2}{E[N]} > 1$$

is called a *clustered process*. As one extreme, a perfectly clustered process would have all cells but one empty and that cell would contain a number of points totaling $E[N]\cdot$(the number of cells).

As a simple example of such spatial processes, consider a process for which a cell has probability $(1 - p)$ of containing zero points and a probability p of containing M points. Then,

$$E[N] = p \cdot M \tag{3.109a}$$

$$\sigma_N^2 = M^2 p(1 - p) \tag{3.109b}$$

so that

$$r = M(1 - p) \tag{3.110}$$

Here $r > 1$ if M is "sufficiently large" or if p is "sufficiently small." Formally, $r > 1$ and thus we have a clustered process if $M > 1/(1 - p)$ or, equivalently, if $p < (1 - 1/M)$. Interestingly, if these inequalities are reversed, we have a spread process rather than a clustered process. This makes sense since in the extreme if $p = 1$, each cell contains exactly $M = E[N]$ points. If the inequalities were strict equalities, we would have a two-valued process whose spatial randomness (as measured by r) is identical to that of the Poisson process.

3.9.1 Spread Process Yielding the Binomial PMF

Rogers has studied two particular processes—one spread and one clustered—that have appealing time-Poisson process interpretations and that have been found useful in analyzing the locations of retail trade [ROGE 74]. We consider first Rogers's spread process, the binomial process. Imagine that entities enter the cell of interest over some time interval $[0, t]$, initially with 0 entities in the cell. We are interested in the number of entities at time t, $N(t)$. Being a spread process, each time that another entity enters the cell the rate at which new entities enter the cell diminishes. Thus, suppose initially that entities enter the cell as a time-Poisson process at rate c per unit time. Then, after the first enters the cell, the cell becomes "less attractive," so the new Poisson arrival rate is $c - b$. In general, after k arrivals, the Poisson arrival rate is reduced to $c - kb$. Thus, the cell becomes less attractive in a linear manner with the number of entities already in the cell. We assume that c/b is integer, so that there exists some maximum k, $k_{max} = c/b$, at which the Poisson arrival rate is reduced to $c - k_{max}b = 0$. Thus, the maximum number of entities in a cell is $k_{\max} = c/b$. This pure birth process is characterized by the state-transition diagram shown in Figure 3.36.

Proceeding as in Chapter 2 for the Poisson process, this process is governed by the following set of coupled differential equations:

$$\frac{dP_0(t)}{dt} = -cP_0(t) \tag{3.111a}$$

$$\frac{dP_{m+1}(t)}{dt} = (c - mb)P_m(t) - [c - (m+1)b]P_{m+1}(t)$$

$$m = 0, 1, 2, \ldots, \frac{c}{b} - 1 \tag{3.111b}$$

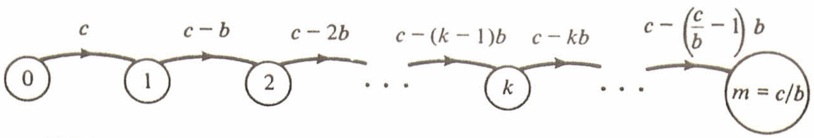


FIGURE 3.36 State-transition diagram for the binomial spread process.

Proceeding as with the Poisson process, we find that

$$P_0(t) = e^{-ct} \qquad t \geq 0$$

Successive substitutions into (3.111) for increasing values of m leads us to prove by induction that

$$P_m(t) = \binom{c/b}{m}(1 - e^{-bt})^m (e^{-bt})^{(c/b)-m} \qquad m = 0, 1, 2, \ldots, \frac{c}{b}, \quad t \geq 0 \tag{3.112}$$

This is the binomial pmf with probability of "success" equal to $(1 - e^{-bt})$. The mean and variance are

$$E[N(t)] = \frac{c}{b}(1 - e^{-bt}) \tag{3.113a}$$

$$\sigma^2_{N(t)} = \frac{c}{b}(1 - e^{-bt})e^{-bt} \tag{3.113d}$$

The ratio of the variance to the mean is

$$r = e^{-bt} \tag{3.114}$$

which is always less than unity (which is what we want with a spread process).

While the "diminishing attractiveness" interpretation of this birth process is perfectly valid, and quite appealing as a description of the dynamics of a spread process, it is not the only interpretation of the process. Alternatively,

one might imagine a population fixed with $n = c/b$ individuals. Each one will eventually locate within the cell, but the time until such location is an exponentially distributed random variable with mean $1/b$. All n such random variables are *mutually independent*. Thus, at time $t = 0$, n "Poisson generators" are turned on, yielding a rate of transition nb from state 0 to 1; after the first transition, $(n - 1)$ Poisson generators remain turned on with a net rate of occurrence equal to $(n - 1)b$. This "fixed population" interpretation of cell occupancy also yields the binomial distribution, and it could imply markedly different policy decisions in practice than the "diminishing attractiveness" interpretation. The two equally plausible interpretations provide a good example that any particular probability law may have two, three, or even a greater number of plausible underlying explanations. Thus, just because a probability law assumes a particular form does not assure us that one underlying causal model is *the* model explaining the process dynamics.

3.9.2 Clustered Process Yielding the Negative Binomial PMF

Rogers's clustered spatial process gives rise to the negative binomial distribution. In this model we assume that a cell becomes more attractive with each additional entity that locates there. In particular, if there are m entities there at time t, new entities arrive in a (time) Poisson manner at rate $c + bm$ $(c > 0, b > 0)$. The state-transition diagram for this infinite-state pure birth process is shown in Figure 3.37. Proceeding as usual, the set of coupled differential equations governing this process are

$$\frac{dP_0(t)}{dt} = -cP_0(t) \tag{3.115a}$$

$$\frac{dP_{m+1}(t)}{dt} = (c + mb)P_m(t) - [c + (m+1)b]P_{m+1}(t) \qquad m = 0, 1, 2, \ldots \tag{3.115b}$$

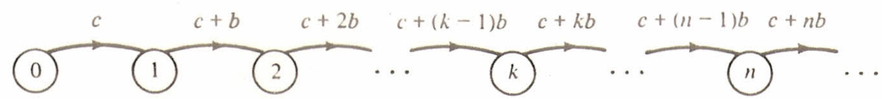


FIGURE 3.37 State-transition diagram for the negative binomial cluster process.

Again we find that

$$P_0(t) = e^{-ct} \qquad t \geq 0$$

Successive substitutions into (3.115) lead us to prove by induction that

$$P_m(t) = \binom{\frac{c}{b} + m - 1}{m}(1 - e^{-bt})^m (e^{-bt})^{c/b} \qquad m = 0, 1, 2, \ldots \quad t \geq 0 \tag{3.116}$$

This is the *negative binomial* pmf[11] with mean

$$E[N(t)] = \frac{c}{b}(e^{bt} - 1) \tag{3.117}$$

and variance

$$\sigma_N^2(t) = \frac{c}{b}(e^{bt} - 1)e^{bt} \tag{3.118}$$

The ratio of the mean to the variance is

$$r = e^{bt} \tag{3.119}$$

which is always greater than unity (which is what we want with a clustered process).

Although the "increasing attractiveness" interpretation of this process is appealing for a clustering process, there are other plausible system dynamics yielding the same negative binomial pmf.

For comparative purposes, we have sketched the mean value $E[N(t)]$ for each of the three cell occupancy laws—Poisson, binomial, negative binomial—in Figure 3.38. Note that the binomial (spread) process reaches a

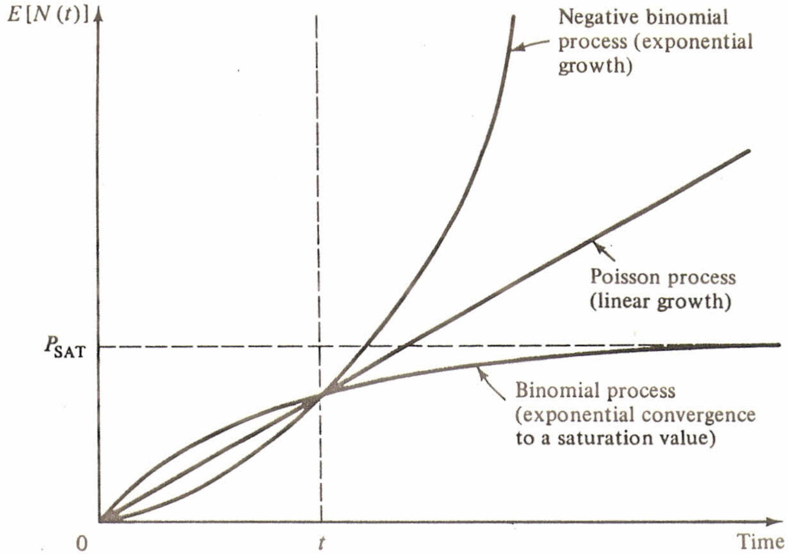


FIGURE 3.38 Mean value $E[N(t)]$ for each of three processes. Model parameters have been adjusted so that all three means are equal at time t.

[11] See, for example, A. W. Drake, *Fundamentals of Applied Probability*, McGraw-Hill, Inc., New York, 1967, pp. 128–130, 153.

"saturation" population, P_{sat}, whereas the Poisson process grows linearly in time and the negative binomial (clustered) process "explodes" at an exponential rate. Figure 3.39 illustrates each process over a 10- by 10-kilometer city.

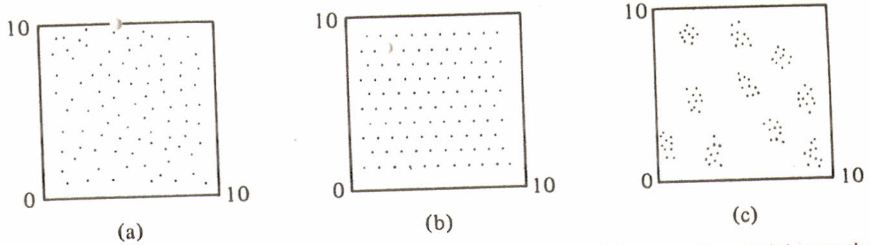


FIGURE 3.39 Illustrated examples of (a) Poisson, (b) binomial (spread) and (c) negative binomial (clustered) processes. Means normalized to 1 point per unit area.

3.10 CONCLUSION

We have now completed our tour of derived distributions, geometrical probability, and spatial processes. For those interested in further study of geometrical probability, we recommend the recent book by Solomon [SOLO 78]. Chapter 4 switches emphasis from space to time, dealing with congestion that arises in queueing systems. Later chapters rely heavily on the probabilistic modeling methods of Chapter 2, this chapter, and Chapter 4, to study congestion phenomena in a spatially oriented urban setting.

References

[BANN 72] BANNER, G. P., "Geometrical Probability: Applications to Urban Service Systems," unpublished thesis for degrees of Master of Science and Electrical Engineer, Massachusetts Institute of Technology, Cambridge, Mass., 1972.

[BERT 07] BERTRAND, J., *Calcul des probabilités*, Paris, 1907.

[BUFF 77] BUFFON, G., Essai d'arithmétique morale, Supplément à *l'Histoire naturelle*, 4, (1777).

[DRAK 67] DRAKE, A. W., *Fundamentals of Applied Probability*, McGraw-Hill, New York, 1967.

[HAUS 75] HAUSNER, J., *Determining the Travel Characteristics of Emergency Service Vehicles*, R-1687-HUD, The New York City Rand Institute, April 1975.

[JARV 75] JARVIS, J. P., K. A. STEVENSON, and T. R. WILLEMAIN, *A Simple Procedure for the Allocation of Ambulances in Semi-Rural Areas*, Technical Report TR-13-75, MIT Operations Research Center, Cambridge, Mass., March 1975.

[KEND 63] KENDALL, M. G., and P. A. P. MORAN, *Geometrical Probability*, Charles Griffen, London, 1963.

[KOLE 73] KOLESAR, P., and E. H. BLUM, "Square Root Laws for Fire Engine Response Distances," *Management Science*, *19* (12), 1368–1378 (1973).

[KOLE 75] KOLESAR, P., "A Model for Predicting Average Fire Engine Travel Times," *Operations Research*, *23*, 603–613 (1975).

[KOOP 56] KOOPMAN, B. O., "The Theory of Search II. Target Detection," *Operations Research*, *4* (5), 503–531 (1956).

[LARS 72a] LARSON, R. C., *Urban Police Patrol Analysis*, MIT Press, Cambridge, Mass., 1972.

[LARS 72b] LARSON, R. C., and K. A. STEVENSON, "On Insensitivities in Urban Redistricting and Facility Location," *Operations Research*, *20*, 596–612 (1972).

[ROGE 74] ROGERS, A., *Statistical Analysis of Spatial Dispersion: The Quadrat Method*, Pion Ltd., London, 1974.

[SOLO 78] SOLOMON, H., *Geometric Probability*, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1978.

Problems

3.1 *Review: basic concepts of probability modeling* A certain town has exactly one policeman (Jones) and exactly one burglar (Elmer). The town is divided into two police beats, each of which may be considered a straight line of length ℓ. Each night the policeman makes an equally likely choice between the two beats and then spends the whole night patrolling the selected beat. When Jones is on a beat, his position at any time is uniformly distributed over the length of that beat.

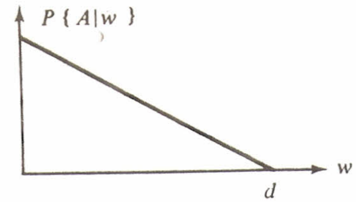
Tonight, Elmer will start committing one burglary per night until he is apprehended. On any particular night, given that he has not already been caught, Elmer is twice as likely to burglarize the beat that Jones is not patrolling than the one that he is patrolling.

Elmer's burglary position is uniformly distributed over the beat that he has selected and is independent of Jones's position, even if he and Jones happen to have selected the same beat. Assume that Jones's position remains constant throughout the duration of the burglary.

Given that Elmer and Jones are exactly w units of length apart on the *same* beat at the time of burglary, Jones will apprehend Elmer with probability $P\{A \mid w\}$, as

shown in Figure P3.1. Note that $P\{A \mid w\}$ is a *conditional probability*, not a probability density function.

$$P\{A \mid w\} = \begin{cases} 1 - \dfrac{w}{d} & 0 \le w \le d \\ 0 & \text{elsewhere} \end{cases} \quad \text{with } 0 < d < \ell$$



a. What is the probability that Elmer and Jones will both work on the same beat tonight?

b. Given that Elmer and Jones are on the same beat tonight, and also given that they are separated by a distance of more than $\ell/4$ units, what is the conditional probability that they are separated by a distance of more than $\ell/2$ units?

c. Given that Elmer and Jones are on the same beat tonight, determine the pdf $f_W(w)$ for $-\infty < w < \infty$, where W is the magnitude of the distance between them at the time of the burglary.

d. Given that Elmer has not as yet been caught and given that tonight he and Jones choose the same beat, show that P_A, the conditional probability that he will be apprehended tonight, is $(d/\ell)[1 - \frac{1}{3}(d/\ell)]$. Does this answer seem reasonable for $d = 0$ and $d = \ell$?

e. Determine the probability that Elmer is apprehended for the first time on the third night.

f. Given that Elmer has successfully completed exactly 10 burglaries, what is the probability that Jones and Elmer worked the same beats exactly three of those nights?

g. Jones is considering a new patrol strategy. He will still choose his beat randomly as before, but he will now simply stand in the center of it instead of patrolling it. If everything else remains the same (and Elmer does not change his strategy), what now is the probability of apprehension on any given night if Elmer has not previously been caught? Does your answer seem reasonable for $d = 0$ and $d = \ell$?

3.2 *Discrete random variable.* Let X_i $(i = 1, 2)$ be uniformly, independently distributed over the integers $0, 1, 2, \ldots, m$. Define the distance between X_1 and X_2 as

$$D = |X_1 - X_2|$$

a. Determine the pmf for D.

b. Show that $E[D] = \frac{1}{3}m + \frac{1}{3}\left(\frac{m}{m+1}\right)$.

3.3 *Functions of random variables.* Two emergency response units patrol uniformly and independently a 10-mile stretch of road. An emergency incident occurs on the roadway and its position is uniformly distributed, independent of the positions of the response units. The incident requires *both* response units to be dispatched to the scene. Call the two units unit a and unit b. Assume that response speed is fixed at 10 mph and that U-turns are permitted.

a. Determine the mean travel time for unit a to reach the scene.

b. Determine the mean time until the first unit (either a or b) reaches the scene.

c. Determine the probability density function for the time until the second unit reaches the scene.

3.4 *Functions of random variables.* Assume that the locations of an incident and a response unit are independently, uniformly distributed over a rectangle with dimensions X_0, Y_0 (see Exercise 3.1). The sides of the rectangle are defined parallel to directions of travel. If the incident and response unit locations are (X_1, Y_1) and (X_2, Y_2), respectively, the travel distance is

$$D = |X_1 - X_2| + |Y_1 - Y_2|$$

a. For the case $X_0 = Y_0$, find the pdf of D.

b. For the case $X_0 \geq Y_0$, identify the different regions of integration in the X, Y sample space that yield different functional forms for the pdf of D.

c. (*Optional*) For the very brave, carry out the computations for part (b) to find the pdf of D when $X_0 > Y_0$.

3.5 *Time, speed, distance.* Suppose that ambulance attendants read the following data for four random ambulance responses:

Number of Miles Driven	*Speed of Response (miles/hr)*	*Travel Time (minutes)*
1.2	10	7.2
6.0	18	20
0.2	1	12
13.0	21	37.1

a. Verify that the average distance per response $\equiv d_{av} = 5.1$ miles, the average speed per response $\equiv S_{av} = 12.5$ miles/hr, and the average time per response $\equiv t_{av} = 19.1$ minutes.

b. Intuitively, explain why $(d_{av}/S_{av})\cdot(60 \text{ minutes/hr}) = 24.48$ minutes is greater than $t_{av} = 19.1$ minutes.

c. If we compute a weighted average speed S_{wt}, where the weights sum to 1 and are proportional to the *times* driven at each speed, we find $S_{wt} = 16.03$ miles/hr. We find that $(d_{av}/S_{wt})\cdot(60 \text{ minutes/hr}) = t_{av} = 19.1$ minutes. Why is this a correct procedure?

3.6 *AVL systems.* If the travel distance metric of the vehicles being located by an AVL system is right angle, the travel distance between the vehicle's estimated and true positions is $|X_e| + |Y_e|$. Here X_e and Y_e are zero mean, variance σ^2, independent Gaussian random variables representing x- and y-location estimation error, respectively, as described in Exercise 3.2. Making the necessary isotropy assumption, argue that

$$E[|X_e| + |Y_e|] = \sigma\frac{4}{\pi}\sqrt{\frac{\pi}{2}} = 2\sigma\sqrt{\frac{2}{\pi}}$$

Thus, the expected value of a Gaussian random variable with variance σ^2 which is *truncated at zero* is $E[|X_e|] = E[|Y_e|] = \sigma\sqrt{2/\pi}$.

3.7 *AVL systems.* Suppose that the individual x and y errors of an AVL system are independently distributed according to a *Laplace* pdf,

$$f_{X_e}(x) = f_{Y_e}(x) = \frac{\lambda}{2}e^{-\lambda|x|} \qquad -\infty < x < +\infty, \quad \lambda > 0$$

Suppose that the radius of error is measured directly in terms of right-angle distance,

$$R = |X_e| + |Y_e|$$

Show that

$$f_R(r) = r\lambda^2 e^{-\lambda r} \qquad r \geq 0 \qquad (*)$$

and thus

$$E[R] = \frac{2}{\lambda}$$

[You might try obtaining (*) using both the "never-fail" method and the "infinitesimal method" outlined in Example 3.]

3.8 *Test of the "right-angle" distance metric, revisited* In Example 4 we obtained the probability law for the random variable R, the ratio of the right-angle and Euclidean distances. In particular, we found

$$E[R] = \frac{4}{\pi} \approx 1.273$$

$$\sigma_R^2 = 1 + \frac{2}{\pi} - \frac{16}{\pi^2}$$

Derive these results directly *without* first obtaining $F_R(\cdot)$ or $f_R(\cdot)$.

3.9 *Functions of random variables (derived distributions)* Consider a square service region of unit area in which travel is right-angle and directions of travel are parallel to the sides of the square. Let (X_1, Y_1) be the location of a mobile unit and (X_2, Y_2) the location of a demand for service. The travel distance is

$$D = D_x + D_y$$

where

$$D_x = |X_1 - X_2| \quad \text{and} \quad D_y = |Y_1 - Y_2|$$

We assume that the two locations are independent and uniformly distributed over the square.

a. Show that the joint pdf for D_x and D_y is

$$f_{D_x,D_y}(x,y) = \begin{cases} 4(1-x)(1-y) & 0 \le x \le 1, \quad 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

b. Define $R_{yx} = D_y/D_x$. Show that the pdf for R_{yx} is

$$f_{R_{yx}}(r) = \begin{cases} \dfrac{2}{3} - \dfrac{1}{3}r & 0 \le r \le 1 \\ \dfrac{2}{3r^2} - \dfrac{1}{3r^3} & 1 \le r < \infty \end{cases}$$

3.10 *Ratio of right-angle and Euclidean travel distances* In this problem we test the reasonableness of the isotropy assumption used in Example 4. It is appropriate to question this assumption since most service regions in a city are such that Ψ will not be uniformly distributed between 0 and $\pi/2$. We consider three cases.

a. *Case 1.* For the square service region of Problem 3.9, in which directions of travel are parallel to the sides of the square, one might expect intuitively that $E[R] > 4/\pi \approx 1.273$. Why? Show that $E[R] = \frac{1}{3}[5 \log(1 + \sqrt{2}) + \sqrt{2} - 2] \approx 1.274$.

Hint: In terms of Problem 3.9, recognize that

$$R = \frac{1 + R_{yx}}{\sqrt{1 + R_{yx}^2}}$$

b. *Case 2.* Suppose that the square-unit-area service area of part (a) is rotated at a 45° angle to the directions of travel. In such a case intuition might lead one to think that $E[R]$ would be less than $4/\pi$. Why? To investigate this conjecture it is helpful to use the relationship

$$|x_1 - x_2| + |y_1 - y_2| = \sqrt{2}\,\text{Max}\,[|x_1' - x_2'|, |y_1' - y_2'|]$$

where the primed variables are defined relative to a coordinate system rotated at 45° with respect to the original coordinate system. Show that in such a case

$$E[R] = \tfrac{1}{3}[4\sqrt{2}\log(1 + \sqrt{2}) + 2\sqrt{2} - 4] \approx 1.271$$

(Intuition is correct but the result is closer to $4/\pi$ than might otherwise have been expected.)

c. *Case 3.* Suppose that the mobile unit is located uniformly *on* the perimeter of a square rotated at 45° to the directions of travel. Suppose that the unit travels in a shortest (right-angle) distance manner to the center of the square. Again, Ψ is the angle at which the directions of travel are rotated with respect to the straight line connecting the unit's initial position to the center of the square. Show that

i. $F_\Psi(y) = \frac{1}{2}[1 + \tan(y - \pi/4)]$, $0 \le y \le \pi/2$.

ii. $F_R(r) = 1 - \dfrac{\sqrt{2 - r^2}}{r}$, $\quad 1 \le r \le \sqrt{2}$.

iii. $E[R] = \sqrt{2}\log(1 + \sqrt{2}) \approx 1.246$.

Do all the results for $E[R]$ check with your intuition?

3.11 *Quantization model (continued)* In Example 5 we described a quantization model for odometer readings. We stated that [(3.31)]

$$E[K] = E[D]$$
$$\sigma_K^2 \ge \sigma_D^2$$

Prove these results. What implications do these results have for an actual data-gathering experiment?

Hint: Do not work directly with (3.30); instead, demonstrate the validity of the desired results for any $\{D = d\}$ and then integrate over all d.

3.12 *Truncated times* Assume that an activity commences at time T_1 and terminates at time T_2. The exact duration of the activity is $T_2 - T_1 \equiv T$. Now assume that times are recorded by some mechanism that records time x as $\lfloor x + \alpha \rfloor$, for some fixed α. Using this mechanism, the recorded duration of the activity is $\lfloor T_2 + \alpha \rfloor - \lfloor T_1 + \alpha \rfloor$.

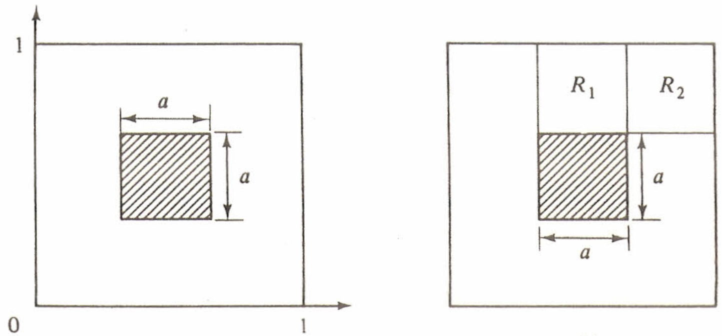
a. Argue why it is reasonable to apply the distance quantization model (Example 5) to this situation, with the following correspondences:

$$D\colon \quad T_2 - T_1 \equiv T$$
$$\Theta\colon \quad T_1 + \alpha - \lfloor T_1 + \alpha \rfloor \equiv \Phi$$
$$K\colon \quad \lfloor T_2 + \alpha \rfloor - \lfloor T_1 + \alpha \rfloor \equiv J$$

b. Show that $J = \lfloor T + \Phi \rfloor$.

3.13 *Zero-demand zone* Consider a unit-square response area, as shown in Figure P3.13(a). We assume that a response unit and incident (i.e., requests for service) are distributed uniformly, independently over that part of the unit square *not contained* within the central square having area a^2. Travel occurs according to the right-angle metric, *and* travel is allowed through the zero-demand zone. We want to use conditioning arguments to derive the expected travel distance $\bar{W}(a)$ to a random incident.

Let (X_1, Y_1) and (X_2, Y_2) denote the locations of the response unit and incident, respectively. Let S (S') denote the set of points within (outside) the central square. Let $A = \{(X_1, Y_1) \in S\}$ and $B = \{(X_2, Y_2) \in S\}$.



(a) (b)

Now focus on a unit square on which incidents and the response unit are uniformly, independently distributed over the *entire* square, yielding an expected travel distance $E[D]$.

a. Show that

$$E[D] = \tfrac{2}{3} = E[D \mid A \cap B]P\{A \cap B\} + 2E[D \mid A \cap B']P\{A \cap B'\}$$
$$+ E[D \mid A' \cap B']P\{A' \cap B'\}$$
$$= \tfrac{2}{3}a(a^2)^2 + 2E[D \mid A \cap B']a^2(1 - a^2) + E[D \mid A' \cap B'](1 - a^2)^2$$

b. We wish to derive $E[D \mid A' \cap B'] = \bar{W}(a)$. The relationship above allows us to compute this quantity by finding the easier-to-compute quantity $E[D \mid A \cap B']$. (Note the similarity of approach to Crofton's method.)

i. To find $E[D \mid A \cap B']$, argue that one need only consider the incident to be located in R_1 or R_2, as shown in Figure P3.13(b).

ii. Show that

$$P\{(X_2, Y_2) \in R_1 \mid (X_2, Y_2) \in R_1 \cup R_2\} = \frac{2a}{a+1}$$

iii. Show that

$$E[D \mid A \cap R_1] = \tfrac{1}{4} + \tfrac{7}{12}a$$
$$E[D \mid A \cap R_2] = \tfrac{1}{2} + \tfrac{1}{2}a$$

c. Finally, find $\bar{W}(a)$. As a check, $\bar{W}(0) = \tfrac{2}{3}$, $\bar{W}(1) = \tfrac{11}{12}$. (Why?)

3.14 *Square barrier* Suppose that the conditions of Problem 3.13 apply, except that in addition, no travel is allowed through the central square. We wish to derive

$$\bar{W}'(a) = \text{expected travel distance to a random incident}$$

We use perturbation arguments to write

$$\bar{W}'(a) = \bar{W}(a) + \bar{W}_E(a)$$

where $\bar{W}(a)$ is the mean travel distance from Problem 3.13 and $\bar{W}_E(a)$ is the mean extra (perturbation) distance due to the barrier.

a. Show that the probability that the perturbation term is strictly positive is

$$\left(\frac{a}{a+1}\right)^2$$

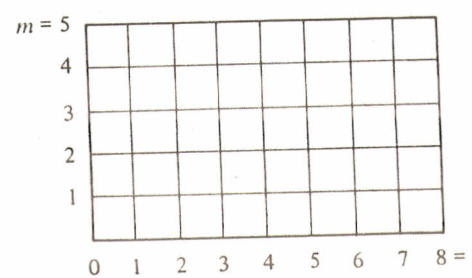
b. Show that the conditional extra travel distance, given that the perturbation term is positive, is $\tfrac{1}{3}a$. Thus,

$$\bar{W}_E(a) = \frac{1}{3}a\left(\frac{a}{a+1}\right)^2$$

As a check, verify the reasonableness of the result $\bar{W}'(1) = 1$.

3.15 *Rectangular grid of two-way streets* Consider an $n \times m$ rectangular grid of two-way streets running north-south and east-west as shown in Figure P3.15. Assume that incident positions are distributed uniformly over the grid. A response unit patrols the grid in a uniform manner. The incident location and the response unit location are independent. Let

D = travel distance between the response unit and the incident, assuming the unit follows a shortest path that remains on the streets of the grid



a. (*Optional*) By carefully and patiently conditioning on the various possible locations for the incident and response unit, show that

$$E[D] = \frac{1}{3}(n + m) + \frac{4n(m+1)m(n+1)}{3[(n+1)m + (m+1)n]^2}$$

b. Show that

$$\tfrac{1}{3}(n + m) \leq E[D] \leq \tfrac{1}{3}(n + m + 1)$$

where the left-hand inequality becomes an equality when n or m is zero and the right-hand inequality becomes an equality when $n = m$. Thus, *the continuous approximation*, (3.12a), *is never in error by more than* $\frac{1}{3}$ *block length.*

3.16 *Perturbation variables: one-way streets* Consider a very large grid of equally spaced one-way streets, with the direction of travel alternating from street to adjacent parallel street. Assume that the positions of the response unit and the incident are independent and uniformly distributed over the grid. It is assumed that the response distance from the response unit to the incident is a shortest path that remains on the streets of the grid and obeys the one-way constraints. Use perturbation variables to demonstrate that the mean *extra* distance traveled to the incident, due to the one-way travel constraints, is *two blocks*.

3.17 *Cauchy random variable* We recall from Section 3.3.3 that random variable X_i has a Cauchy pdf if

$$f_{X_i}(y) = \frac{1}{\pi(1 + y^2)} \qquad -\infty < y < +\infty$$

a. Suppose that $S_2 = X_1 + X_2$, where X_1 and X_2 are independent Cauchy random variables, each having pdf $f_{X_i}(\cdot)$. Using the integral identity

$$\frac{1}{\pi^2}\int_{-\infty}^{+\infty} \frac{dw}{(1 + w^2)\{n^2 + (y - w)^2\}} = \frac{(1/\pi)(n+1)}{n} \frac{1}{y^2 + (n+1)^2}$$

show that S_2 has a pdf $2/[\pi(4 + y^2)]$.

b. Proceeding by induction, show that

$$S_n \equiv X_1 + X_2 + \ldots + X_n \quad (\text{all } X_i \text{ independent})$$

has a pdf $n/[\pi(n^2 + y^2)]$.

c. Thus, verify that the average of n independent Cauchy samples (i.e., $V_n \equiv S_n/n$) has a Cauchy pdf $1/[\pi(1 + y^2)]$. Thus, "averaging together" a number of independent Cauchy samples yields a pdf for the average identical to that of any one of the individual samples. (This result contrasts sharply to most random variables, for which averaging of n independent samples reduces the variance by a factor of n^{-1}.)

3.18 *Crofton's method* Suppose that X_1 and X_2 are two points independently uniformly distributed over a highway segment of length a. Define

$$D^p \equiv |X_1 - X_2|^p \qquad p > 0$$

Use Crofton's method to show that

$$E[D^p] = \frac{2a^p}{(p+1)(p+2)}$$

3.19 *Crofton's method* Consider again policeman Jones and burglar Elmer of Problem 3.1. Use Crofton's method to verify that the apprehension probability P_A [Problem 3.1(d)] equals $(d/\ell)[1 - \frac{1}{3}(d/\ell)]$.

Hint: Here the homogeneous solution of the associated differential equation cannot be discarded.

3.20 *Crofton's method* Here we wish to apply Crofton's method for finding mean values to the problem of finding the mean Euclidean distance,

$$E[D] = E[\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}]$$

where (X_1, Y_1) and (X_2, Y_2) are uniformly and independently distributed over a circle of radius r. Here, for instance, (X_1, Y_1) and (X_1, Y_2) could be the locations of an emergency and a helicopter response unit, respectively, and D would be the travel distance to the emergency.

a. By arguments similar to those used in the text, show that

$$\delta\mu = \frac{4(\mu_1 - \mu)}{r}\delta r$$

where

$$\mu = E[D]$$

$$\mu_1 = E[D \mid \text{exactly one of the points is in the infinitesimal ring of width } \delta r \text{ on the circle circumference}]$$

$$\mu + \delta\mu = \text{mean travel distance in a circle of radius } r + \delta r$$

b. Show that

$$\mu_1 = \frac{1}{\pi r^2}\int_0^{2r} 2x^2 \cos^{-1}\frac{x}{2r}\,dx = \frac{32r}{9\pi}$$

Hint: From tables of integrals, you may find useful

$$\int x^n \cos^{-1} x\,dx = \frac{x^{n+1}\cos^{-1} x}{n+1} + \frac{1}{n+1}\int \frac{x^{n+1}\,dx}{\sqrt{1 - x^2}}$$

$$\int \frac{x^3\,dx}{\sqrt{a^2 - x^2}} = -\frac{x^2}{3}\sqrt{a^2 - x^2} - \frac{2}{3}a^2\sqrt{a^2 - x^2}$$

c. Use your results in parts (a) and (b) to obtain

$$\mu = E[D] = \frac{128}{45\pi} r \approx 0.906r$$
$$\approx 0.51\sqrt{A}$$

where $A = \pi r^2$ is the area of the circle.

Note: You have just derived one of the constants in Table 3-1.

3.21 *Expected values* Suppose that two points (X_1, Y_1) and (X_2, Y_2) are uniformly and independently distributed over a circle of area A. Assume that the travel distance D between the two points is the right-angle travel distance

$$D = |X_1 - X_2| + |Y_1 - Y_2|$$

Argue that

$$E[D] = \frac{4}{\pi}\frac{128}{45\pi}\sqrt{\frac{A}{\pi}} \approx 0.650\sqrt{A}$$

Hint: Consider D to be the product of the Euclidean distance and a scaling factor R, the ratio between the right-angle and Euclidean distances.

3.22 *Crofton's method* Use Crofton's method to rederive (3.12a) for the mean travel distance of a rectangular response area.

Hint: Points in the infinitesimally thick "frame" surrounding the original rectangle are not indistinguishable, as they are for the circles.

3.23 *Coverage; Robbins's theorem on random sets* Imagine a square region of a city having unit area. Suppose that there are N ambulettes whose positions are independently and uniformly distributed over a region T consisting of all points in the city whose distance from the square is not greater than a. The area of T is $1 + 4a + \pi a^2$. A point in the unit square is said to have sufficient ambulance coverage if at least one ambulance is within a (Euclidean) distance a of the point. Find the expected area within the square which is sufficiently covered.

3.24 *Coverage of a square lattice by a rectangle* A city's geographical structure is being placed on a computer. All coordinate positions are being quantized, where the unit of quantization is 500 feet. The quantization points comprise a lattice that runs east-west and north-south. The board of elections wishes to know how many quantization points will be contained in an arbitrary rectangular election district of dimension ℓ (east-west) and m (north-south).

Assume that the location of the election district on the lattice can be modeled as random (but the sides are parallel to the two directions of the lattice). Let N be the number of lattice points contained within the election district.

a. Show that

$$E[N] = \ell m$$

b. Let $\ell = p + q$, $m = P + Q$ $(0 \le q, Q < 1)$. Show that

$$\sigma_N^2 = q(1-q)Q(1-Q) + \ell^2 Q(1-Q) + m^2 q(1-q)$$

X **3.25** *Spatial Poisson process* Suppose that response units are distributed throughout the city as a homogeneous spatial Poisson process, with an average of γ response units per square mile. Assume that the travel time between (x_1, y_1) and (x_2, y_2) is

$$t = \frac{|x_1 - x_2|}{v_x} + \frac{|y_1 - y_2|}{v_y}$$

where v_x and v_y are travel speeds in the directions of the abscissa and ordinate, respectively.

Assume that an incident occurs somewhere in the city, independent of the locations of the response units.

a. Find the pdf for T_k, where

T_k = travel time to the kth nearest response unit, $k = 1, 2, \ldots$

Hint: The set of points that are a given travel time from the incident is given by a diamond centered at the incident.

b. Find $E[T_k]$ and $\sigma_{T_k}^2$ and note their functional dependence on k.

3.26 *Space–time Poisson process* Consider a highway that starts at $x = 0$ and extends infinitely eastward toward increasing values of x. Automobile accidents and breakdowns occur along the highway in a Poisson manner in time and space at a rate γ per hour per mile. Any accident or breakdown that occurs remains at the location of occurrence until serviced.

At time $t = 0$, when there are no unserviced accidents or breakdowns on the highway, a helicopter starts from $x = 0$ flying eastward above the highway at a constant speed s. As a service unit, the helicopter will land at the site of any accident or breakdown that it flies over. Moreover, given at time t the helicopter is located at $x = st$, the helicopter can be dispatched (by radio) to service any accident or breakdown that occurs *behind* it (i.e., at values of $x \le st$). We assume that any such dispatch occurs immediately after the accident or breakdown occurs.

We are interested in the time the helicopter *first* becomes busy, *either* by landing at an accident/breakdown site *or* by being dispatched to an accident/breakdown behind its current position; in the latter case, the instant of dispatch (not the time of arrival at the scene) is the time of interest.

Let

T = time that the helicopter first becomes busy

a. Show that T has a Rayleigh pdf with parameter $\sqrt{2s\gamma}$:

$$f_T(t) = 2s\gamma t e^{-s\gamma t^2} \qquad t \geq 0$$

implying that

$$E[T] = \frac{1}{2}\sqrt{\frac{\pi}{s\gamma}}$$

$$\sigma_T^2 = \left(2 - \frac{\pi}{2}\right)\frac{1}{2s\gamma}$$

b. Let

$\beta \equiv$ probability that the first accident/breakdown is a dispatch incident behind the helicopter

$1 - \beta \equiv$ probability that the first accident/breakdown occurs as a result of patrol (i.e., the helicopter discovers it)

Show that $\beta = 1 - \beta = \frac{1}{2}$.

Hint: Condition on the event that the first accident/breakdown occurs in the time interval $(t, t + dt)$.

c. Let

$X \equiv$ location of the first accident/breakdown that the helicopter services

Show that

$$E[X] = \frac{3}{8}\sqrt{\frac{s\pi}{\gamma}}$$

d. Suppose that

$L_1 =$ time of first accident/breakdown that the helicopter flies over, assuming that it is no longer dispatched by radio (i.e., all incidents are helicopter—discovered incidents)

$L_2 =$ time of first accident/breakdown that the helicopter is dispatched to, assuming that it never services accident/breakdowns that it flies over

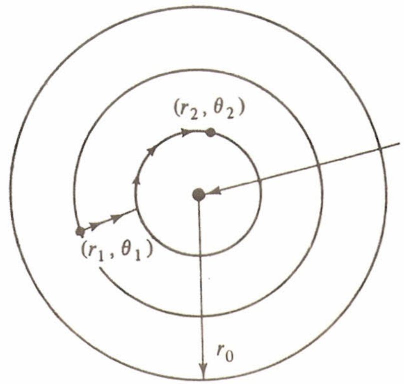
Then, for instance, $T = \text{Min}[L_1, L_2]$. Show that L_1 and L_2 are identically distributed Rayleigh random variables, each with parameter $\sqrt{s\gamma}$. Finally, argue that L_1 and L_2 are independent, thereby concluding that the minimum of two independent Rayleigh random variables, each with parameter $\sqrt{\gamma}$, is itself a Rayleigh random variable with parameter $\sqrt{2\gamma}$.

3.27 *Circular city, revisited* Suppose that two points (R_1, Θ_2) and (R_2, Θ_2) are independently, uniformly distributed over a circular city of radius r_0 and area $A = \pi r_0^2$. Suppose further that this city has a large number of radial routes and circular

ring routes so that the travel distance between (R_1, Θ_1) and (R_2, Θ_2) can be accurately approximated as

$$D = |R_1 - R_2| + \text{Min}[R_1, R_2]|\Theta_1 - \Theta_2|$$

where $0 \leq |\Theta_1 - \Theta_2| \leq \pi$ signifies the magnitude of the angular difference between Θ_1 and Θ_2. In words, travel from an outer point, say, (R_1, Θ_1) if $R_1 > R_2$, to an inner point (R_2, Θ_2) first occurs along a radial route to a ring located a distance R_2 from the city center, and then along that ring (in the direction of minimum travel distance) to (R_2, Θ_2); the same path is traveled in reverse if travel is from (R_2, Θ_2) to (R_1, Θ_1). A sample path is shown in Figure P3.27.



a. Define

$$W \equiv |R_1 - R_2|$$
$$Z \equiv \text{Min}[R_1, R_2]$$

Show that

$$f_W(w) = \begin{cases} \frac{1}{r_0}\left[\frac{8}{3} - 4\frac{w}{r_0} + \frac{4}{3}\left(\frac{w}{r_0}\right)^3\right] & 0 \leq w \leq r_0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_Z(z) = \begin{cases} \frac{4}{r_0}\left[\frac{z}{r_0} - \left(\frac{z}{r_0}\right)^3\right] & 0 \leq z \leq r_0 \\ 0 & \text{otherwise} \end{cases}$$

b. Verify that

$$E[W] = \frac{4}{15}r_0$$

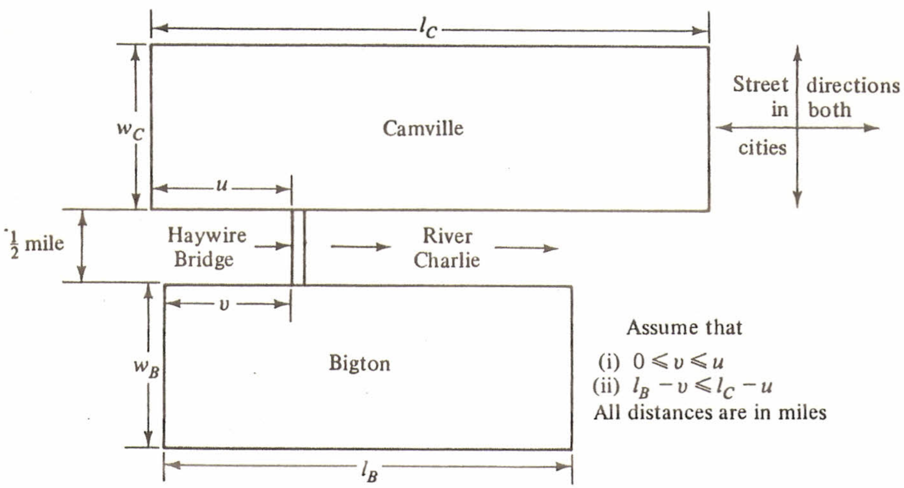
and

$$E[Z] = \frac{8}{15}r_0$$

g. Determine the mean and variance of the queue length (number of vehicles) at traffic light 1 at the instant before the light turns green. (This is a primer for Chapter 4.)

h. A traffic engineer adjusts the phases of the two traffic lights so that $\Theta_1 = \Theta_2 = 0$ (relative to 12:00 noon). Suppose that at 12:00 noon we are given conditional information that *no vehicles* have left *a* during the last 8 minutes. Carefully sketch and label the probability density function for the time of arrival at *b* of the next vehicle to arrive there.

3.30 *Expected travel distances in a two-city area* Consider the cities of Camville and Bigton, which are separated by the River Charlie. Only the $\frac{1}{2}$-mile-long Haywire Bridge connects the two towns. The layout of the cities is shown in Figure P3.30. The population of each of the two cities is uniformly distributed throughout the cities. Camville has a density of λ_C people per square mile; Bigton's density is λ_B people per square mile.



a. Find the expected travel distance between two randomly selected people in Camville; find the expected distance between two randomly selected people in Bigton.

b. Find the expected distance between a randomly selected person in Camville and a randomly selected person in Bigton.

c. Find the expected distance between two randomly selected people in the greater Camville–Bigton metropolitan area.

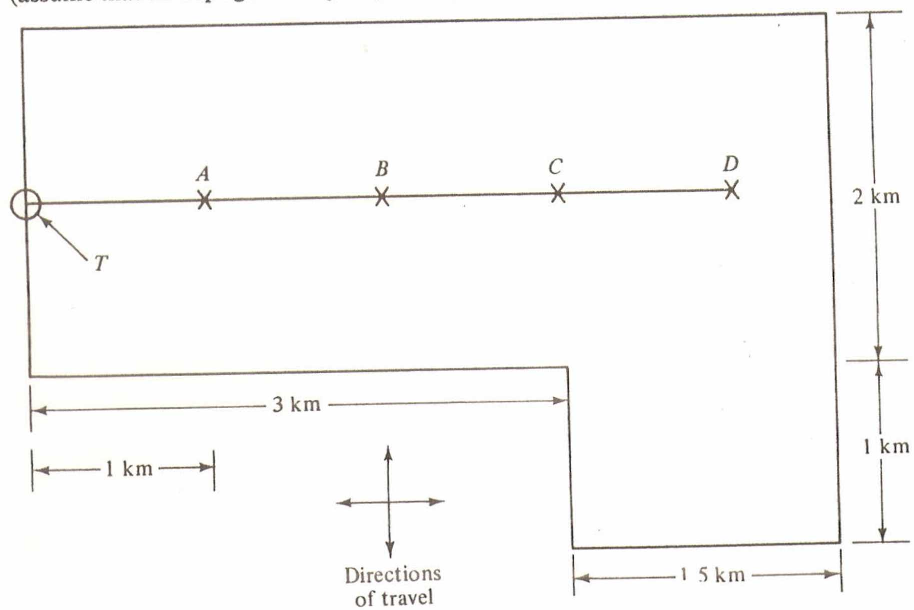
d. Find the expected distance between two randomly selected people, one from *each* town, assuming that one can cross the River Charlie at *any* point along the ℓ_B miles of Bigton that run along the river. By how much does your answer differ from your answer in part (b)?

e. Compute numerical values for all the questions above assuming that

i. $\ell_B = 4$ miles
$\ell_C = 6$ miles
$u = 2$ miles
$v = 1$ mile
$w_B = 5$ miles
$w_C = 3$ miles

ii. $\lambda_B = 2{,}000$ people/square mile
$\lambda_C = 1{,}000$ people/square mile

3.31 *Assigning commuters to subway stations* Figure P3.31 shows an urban area which is served only by a mass-transit rail line with four stops (*A*, *B*, *C*, and *D*) as shown. There is a single destination at point *T* for all trips generated from this area (assume that all trips go exactly to point *T*) during each day's morning rush hour.



In order to get to a transit line station or to walk directly to *T*, the residents of the area must walk on an "infinitely dense" grid of urban streets whose directions run parallel to the boundaries of the area.

The following information is now given:

1. During the morning rush hour the area generates 200 trips per km² with trip origins distributed uniformly.

2. Headways between trains are constant and equal to 6 minutes, and each train rider is equally likely to arrive at a station at any time between two successive departures of trains from that station. (All riders are assumed to be able to ride on the first train to leave a station after their arrival there.) Stops at each station are 1 minute long.

3. Trains travel between stations at a speed of 30 km/hr (this includes an adjustment for acceleration and deceleration periods). People walk at a constant speed of 5 km/hr.
4. The *sole* criterion that each individual uses to determine his/her route is to minimize the expected total trip time to T (including time spent waiting for and riding on trains). Each individual is assumed to know all the information given above concerning travel speeds, headways, and so on.

a. Determine the number of riders who will be using each of stations A, B, C, and D each day, as well as the number of those that will be walking directly to T.

b. Compute the expected travel time for a random resident of this area each day.

c. Draw the boundary of the region whose residents are 9 minutes or less away from T. Repeat for the 20-minute boundary. (Be careful in your work.)

d. Repeat part (a) by making the change in the initial data indicated below, while keeping everything else the same as before. (Each part below is separate.)

1. The train speed increases to 40 km/hr.
2. Train headways are increased to 10 minutes.
3. Train speed $\gg$ walking speed.

e. Repeat part (a) by assuming that headways between trains are described by a negative exponential pdf with a mean of 6 minutes.

3.32 *Optimal district design in terms of city blocks* In this problem we shall examine the question of optimal district design for cases in which the dimensions of a district can assume only integer values due to the grid structure of streets.

Consider again the case described in Problem 3.15, which involved an $n \times m$ rectangular grid district with two-way streets. In that problem you have shown that

$$E[D] = \frac{1}{3}(n+m) + \frac{4}{3}\frac{n(m+1)m(n+1)}{[(n+1)m+(m+1)n]^2}$$

and that

$$\tfrac{1}{3}(n+m) \leq E[D] \leq \tfrac{1}{3}(n+m+1)$$

Suppose now that we wish to design a district so as to minimize $E[D]$, subject to the constraint that the area of the district is greater than or equal to some value A (i.e., $mn \geq A$). (It is not reasonable to set an equality constraint since m and n can take only integer values.) For instance, if $A = 44$, we wish to find n and m such that $E[D]$ is minimized *and* the district contains at least 44 city blocks.

a. Consider the quantity

$$Q = \frac{4}{3}\frac{n(m+1)m(n+1)}{[(n+1)m+(m+1)n]^2}$$

under the constraint $n + m = C$ (where C is a positive integer constant). Argue that under this constraint Q is maximized when $n = m$ (if C is an even number) or $n = m + 1$ (if C is an odd number), and minimized when $n = 0$ or when $m = 0$.

Hint: Q is symmetric in n and m.

b. Show that the following algorithm will give the optimal district dimensions [LARS 72a]:

STEP 1: Set $i = \lceil \sqrt{A} \rceil$, where $\lceil x \rceil$ = smallest integer greater than or equal to x.

STEP 2: Set $j = i - 1$ if $i(i-1) \geq A$; otherwise, set $j = i$.

STEP 3: Set $k = i + 1$, $\ell = j - 1$.

STEP 4: If $k\ell \geq A$, set $i = k$, $j = l$; return to Step 3. Otherwise, *stop;* the optimal district dimensions are $n = i$, $m = j$.

c. Apply your algorithm to find the optimal district dimensions for the case $A = 44$.

d. Compare numerically $E[D]$ for the cases: $m = 7$, $n = 7$; $m = 6$, $n = 8$; $m = 5$, $n = 9$; $m = 4$, $n = 11$. (All these designs satisfy $mn \geq 44$.) What do you conclude from these numerical comparisons?

(23) TORRES, O. FADIGAS: "Curso de Estatística", Depto. Engª de Produção, Escola Politécncia da USP, 1967.

(24) VOLLMANN, T.E. e E.S. BUFFA: "The Facilities Layout Problem in Perspective", Management Science, Vol 12, nº 10, pp 450-468, 1966.