Predicting the Response Time of an Urban Ambulance System

By David W. Scott, Lynette E. Factor, and G. Anthony Gorry

Response time, i.e., the time from dispatch of an ambulance to its arrival at the scene of an emergency, is an important measure of performance in an urban ambulance system. We developed a model that predicts the entire distribution of response time, explicitly accounting for the rate and spatial distribution of demand, variable ambulance velocities, and queueing effects. We tested the model using data sampled from 3,936 ambulance runs in Houston and achieved close agreement between empirical and predicted distributions of response time. Our use of probability theory to predict response times yielded a model that complements those previously reported for planning and evaluating urban ambulance systems.

Through recent legislation and financial support, the federal government has encouraged improvement in the quality of community response to emergencies. However, this support has not been given unreservedly but has been accompanied by an increased emphasis on evaluation [1]. The Emergency Medical Services Systems (EMSS) Act of 1973 mandated the "review and evaluation of the extent and quality of the emergency health care services provided in systems service areas" [2]. However, the factors that affect the performance of emergency medical systems must be understood more fully before proper evaluation studies can be carried out.

Two problems confront the planner of an urban ambulance system —finding appropriate measures of performance for the system and predicting its performance in terms of these measures. Willemain [3] critically reviewed various performance measures for ambulance systems. The number of ambulances per capita and the number of runs per ambulance are common input measures, but they are inadequate for comparing the systems of regions with different spatial characteristics. Output measures, such as changes in mortality rates, are much

This research was supported in part by grant no. 17269 from the National Heart and Blood Vessel Research and Demonstration Center (Baylor College of Medicine), a research project of the National Heart, Lung, and Blood Institute, NIH.

Address communications and requests for reprints to David W. Scott, Assistant Professor of Mathematics, Department of Community Medicine, Baylor College of Medicine, Texas Medical Center, Houston, TX 77030. Lynette E. Factor is a Master's Candidate in the Department of Mathematical Sciences, Rice University, and G. Anthony Gorry is Professor of Health Management, Department of Community Medicine, and Director, Program for Health Management, Baylor College of Medicine.

0017-9124/78/04040414/$02.00/0

more useful since they allow an evaluation of the cost-effectiveness of an ambulance system. In practice, however, it is very difficult to estimate the extent to which an outcome is attributable to an ambulance system alone. Process measures are less useful than output measures but, in the absence of the latter, can be used to assess the performance of an urban ambulance system.

The design of an urban ambulance system can be approached through a variety of analytic techniques and models. The choice of a particular model may be dictated by any of several considerations, such as the development of the system in question, the data available for estimating parameters of the model, the computational resources at hand, and the purpose for which the model is sought. Although substantial progress has been made in the development of urban ambulance system models, no single formulation of the problem seems satisfactory for all purposes. However, it seems likely that most of the needs of urban ambulance system planners can be met through the use of several different approaches.

In a recent article, ReVelle et al. [4] reviewed much of the analytic work on emergency medical systems. They noted that although some models had been developed to address specific problems, others were more general and had been applied to emergency medical systems only after development. ReVelle et al. also distinguished between models concerned mainly with locational and spatial considerations and those focusing on the stochastic nature of ambulance systems.

Queueing theory has been the principal foundation for models of the stochastic type. Because of the analytic complexities of queueing models, simulation has been used in conjunction with them in urban ambulance system analysis. Savas [5], Fitzsimmons [6], Swoveland et al. [7], Berlin and Liebman [8], Stevenson [9], and Cretin [10] have addressed various aspects of urban ambulance systems, using queueing theory, simulation, or a combination of the two. The work of Larson [11,12] is particularly noteworthy, because his dynamic hypercube model addresses many of the important problems of urban ambulance system design.

In their review, ReVelle et al. [4] also discussed context-free location models. These are, in general, optimization models that are used to locate a given number of facilities in a network. The initial formulation of this problem, the p-median problem [13,14], emphasized the mean response time or mean travel time in a region. As ReVelle et al. [4] pointed out, solutions to this problem may leave some portions of a region with unacceptably large response times. Set-covering models [15] improved matters by setting an upper bound on the response time for an emergency in the city, but they did not account for variations in population density. The maximal covering model [4] addressed population density more directly and retained the response time standard of the p-median problem.

The purpose of this research was to complement previous efforts through the development of an analytic model that accounts for the rate and spatial distribution of demand, queueing, and variable ambulance velocities in predicting the probability distribution of response time in an urban region as a function of the number of ambulances and their locations.

Methods

Because of the need to reach an emergency quickly, time is a logical process measure for ambulance systems. For example, the number of ambulances sufficient for an urban area was defined by the EMSS Act as the number that could respond within 10 minutes to 90 percent of all requests. Although it is relatively easy to measure the performance of an ambulance system against this standard, it is difficult to predict the response times of a new system or a proposed expansion of an existing one. We developed a model to determine the probability distribution function of response time, i.e., the time it takes an ambulance once dispatched to reach an emergency. In the model, this probability distribution depends on the rate and spatial distribution of demand, the number of ambulances in the system, the distribution of ambulance velocities, and certain assumptions about the locations of the ambulances. This approach also explicitly includes queueing effects that can be important in determining overall system performance.

We assumed a roughly circular urban region with no prominent barriers to travel within it, such as a river with only one bridge. We also assumed that the spatial distribution of demand was known and that an emergency would be served by the nearest available ambulance. Finally, we assumed that demand for service is the sole determinant of ambulance locations and that the probability that an ambulance is located at any point in the city is an arbitrary but known function of the demand density at that point. It has been noted that response time may not be very sensitive to locations [16], and our latter assumption was analytically convenient because it allowed us to develop the complete probability distribution of ambulance response times. It did not permit the optimization of locations for a particular city, but by specializing our results for particular functions, we were able to investigate the effects of different allocation strategies on response-time distribution.

Given the distribution of demand for ambulance service and our assumed model of ambulance locations within the region, we estimated the probability that when an emergency occurs, the closest ambulance is within a given Euclidean distance, the response distance, of that emergency. We related distance to time through the distribution of travel velocity for ambulances in the city to obtain the probability distribution of response time and then extended our model to account for queueing effects. We did not separately model delays in summoning the ambulance, dispatching delays, time spent at the site of emergency, or time in transit to the hospital. Rather, we combined these times as part of the total ambulance run or service time, which we used in the analysis of queueing. Three allocation strategies were studied in our analysis of the response time distribution for the ambulance system in Houston: in the first strategy, the probability that an ambulance is located at a given point is proportional to the demand density at that point; in the second, proportionally more ambulances are concentrated in regions of high demand; and in the third, ambulances are more dispersed from centers of demand. The effect of queueing was considered for each strategy.

Model Development

The Basic Model of Response Distance. In what follows, random variables are represented by upper-case letters and particular values of random variables by lower-case letters. Let D be the random variable of response distance for an emergency in a region of area A, in which N ambulances are uniformly and independently distributed. Then D approximately follows the spatial Poisson process [17]. For any response distance d greater than zero

$$P(D \leqslant d) = 1 - \exp(-\pi d^2 N/A) \tag{1}$$

We divide the region into K subregions with equal areas $A_i = A/K$ and with expected numbers of emergencies, $\bar{E}_i$, for a given time period. $\bar{Z}_i = \bar{E}_i/A_i$ is the expected demand per unit area in the ith region for the time period in question, and $\bar{E} = \sum_{i=1}^{K} \bar{E}_i$ is total expected emergencies. We omit explicit reference to the time period in what follows.

To model the locations of ambulances, assume that the expected number of ambulances in the subregion $\bar{N}_i$ is given by

$$\bar{N}_i = Nh(\bar{Z}_i) \Big/ \sum_{j=1}^{K} h(\bar{Z}_j)$$

where N is the total number of ambulances and h is an arbitrary nonnegative function. From Eq. 1 and for the ith subregion, we obtain

$$P_i(D \leqslant d) = 1 - \exp(-\pi d^2 \bar{N}_i/A_i)$$

Summing over i, noting that $A_i\bar{Z}_i$ is the expected number of emergencies in the ith subregion, and dividing by the total expected emergencies $\bar{E}$, we obtain:

$$P(D \leqslant d) = 1/\bar{E} \sum_{i=1}^{K} A_i\bar{Z}_i \,[1 - \exp(-\pi d^2 \bar{N}_i/A_i)]$$

Letting $K \rightarrow \infty$ and noting $\bar{E} = A\bar{Z}_m$, where $\bar{Z}_m$ is the mean of the random variable $\bar{Z}$, we obtain:

$$P_N(D \leqslant d) = 1 - 1/\bar{Z}_m \int_0^\infty \bar{z} \;\exp\{[-\pi d^2 N/A\bar{h}(\bar{Z})]\, h(\bar{z})\}\, g(\bar{z})\, d\bar{z} \tag{2}$$

where $g(\bar{z})$ is the probability density function of $\bar{Z}$ and $\bar{h}(\bar{Z})$ is the expectation of $h(\bar{Z})$ given by $\int h(\bar{z})g(\bar{z})\, d\bar{z}$. The subscript N on the probability emphasizes that N ambulances are assumed to have been allocated according to $h(\bar{z})$.

The interpretation of the function $g(\bar{z})$ is that, for a randomly chosen point (x,y) in the region, the ratio of the probability that the density of demand for ambulance service $\bar{Z}$ has a value $\bar{z}_1$ to the probability that it has a value $\bar{z}_2$ is $g(\bar{z}_1)/g(\bar{z}_2)$. We shall show how

$g(\bar{z})$ can be obtained directly from spatial demand data or indirectly from the function chosen to represent the spatial distribution of demand for ambulance service in the region.

The following example justifies the derivation of Eq. 2: If ambulances are distributed uniformly and independently, then $h(\bar{z}) = 1$. In this case, we know that the distance between any emergency site and the nearest ambulance follows the spatial Poisson process. Substitution of $h(\bar{z}) = 1$ into Eq. 2 leads to Eq. 1, as is appropriate for this case.

Obtaining Response Time from Response Distance. The average velocity of an ambulance varies from run to run. Let $\bar{V}$ be the random variable denoting the average velocity for a randomly selected run and associate with $\bar{V}$ the probability density function $f(\bar{v})$. We shall assume that $f(\bar{v})$ is the same for all ambulances and thus solve for the random variable response time, T, by the random variable equation $T = D/\bar{V}$. The distribution function of the response time can be obtained from Eq. 2 using the laws of conditional probability:

$$\begin{aligned} P_N(T \leqslant t) &= P_N(D/\bar{V} \leqslant t) \\ &= \int P_N(D \leqslant \bar{V}t \text{ given } \bar{V} = \bar{v}) \text{ Prob } (\bar{V} = \bar{v}) d\bar{v} \\ &= \int_{\bar{v}} P_N(D \leqslant \bar{v}t) f(\bar{v}) d\bar{v} \end{aligned} \tag{3}$$

Accounting for Queueing. Equation 3 overestimates the response for a system of N ambulances because the ambulance closest to an emergency may be busy when the emergency occurs. Using queueing theory, we can estimate the probability $Q_N(n)$ that exactly n of N ambulances are busy for $n = 0, 1, \ldots, N$. Specifically, if we assume that emergencies arise in a Poisson manner and that the total time required for an ambulance to complete an emergency run follows an exponential distribution, we can calculate the steady-state probability $Q_N(n)$ that n of N ambulances are available. If a call is received when all the ambulances are busy, then we shall assume that other transportation facilities are utilized but, typically, the probability of a well-developed urban system's being fully utilized is extremely small.

Let λ be the Poisson distribution parameter (mean number of requests per hour), and let μ be the parameter of the exponential distribution (inverse of the average time in hours that an ambulance requires to complete an emergency run and return to service). For a finite steady-state solution to exist, we assume the system demand intensity $\lambda/N\mu < 1$. Then from Hillier and Lieberman [18], we have

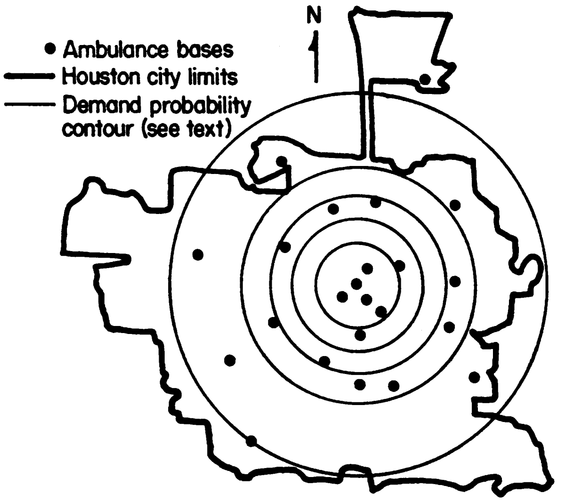
$$Q_N(n) = \rho^{N-n}/[w(N-n)!] \tag{4}$$

where $\rho = \lambda/\mu$ and

$$w = \rho^N/[N!(1-\rho/N)] + \sum_{j=0}^{N-1} \rho^j/j!$$

Using the laws of conditional probability, we can now calculate the probability distribution function of response time for the ambulance system when it is in steady state. We shall assume that, on average, the response time is a small fraction of the total service time, an assumption we shall confirm for our case study. From Eq. 3, we take the

Fig. 1. Spatial distribution of emergency demand and locations of ambulances in Houston.



response time distribution of the system when n ambulances are available, $P_n(T \leqslant t)$, with N replaced by the more general number of available ambulances n. The probability that n ambulances are available is $Q_N(n)$, and the desired probability is

$$P_N^*(T \leqslant t) = \sum_{n=1}^{N} P_n(T \leqslant t) Q_N(n) \tag{5}$$

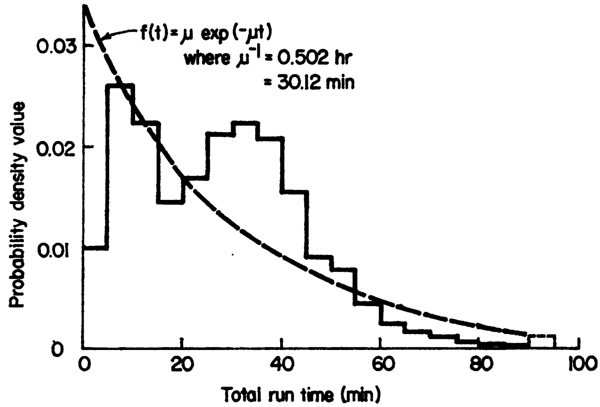
since $P_0(T \leqslant t) = 0$ and ambulance locations are assumed mutually independent. The latter assumption implies that the system responds as if it had a total of n ambulances allocated according to $h(\bar{z})$ for a fraction of time given by $Q_N(n)$. Equation 5 is our general solution to the distribution of response times in an urban ambulance system and may be solved analytically for specific system parameters or numerically if analytic solutions are not obtainable.

Application

Case Study of Houston

In 1973–74, Houston had 23 ambulances stationed at 23 firehouses around the city (Fig. 1). A 28-day stratified random sample of ambulance runs was made by Benson [19] for the period July 1973–June 1974. Location and time data (pickup address, time of dispatch, arrival at the scene, departure from the scene, arrival at the hospital, and return to service) were taken from ambulance records for every run made on those 28 days. The number of runs recorded was 3,936, and an analysis was made of daily and hourly demand variation. Although early morn-

Fig. 2. **Histogram of the total run time for a 28-day stratified random sample in Houston.**



ing demand was lower than average, we assumed a constant demand rate throughout the day to simplify the presentation. The same analysis could be performed for arbitrary daily time periods. From these data we estimated that the average demand rate $\lambda = 5.86$ calls per hour and that the average time required to complete an emergency run $\mu^{-1} = 0.502$ hours per call. The mean response time was about five minutes, only 10 percent of the average total run time. Figure 2 shows the histogram of total run time for the sample and a fitted negative exponential distribution.

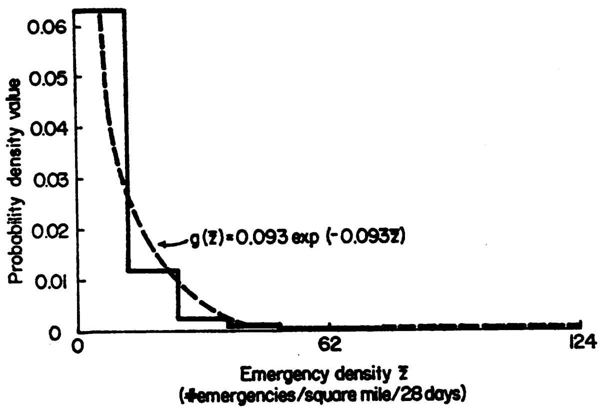
To estimate the distribution of the demand density, the histogram shown in Fig. 3 was developed by calculating demand density for census tracts for the sampled time period of 28 days. For Houston, $g(\bar{z})$ can be represented by an exponential distribution $b \cdot \exp(-b\bar{z})$. One estimate of the exponential parameter b can be obtained from the spatial demand distribution (appendix, Eq. A3). This approach yields $b = A/\bar{E} = 0.102$. Alternatively, the maximum likelihood estimate of b using data shown in Fig. 3 is 0.093. These two estimates differ only slightly. We use $b = 0.093$ in what follows. Notice that $\bar{Z}_m = 1/b$ for the negative exponential.

In Houston, the probability that an ambulance is located at a given point appears to be proportional to $\bar{Z}$ (see Fig. 1) since each of the five annuli contains 20 percent of the demand as well as approximately 20 percent of the ambulances. With this assumption, Eq. 2 becomes (recall $\bar{Z}_m = 1/b$):

$$P_n(D \leqslant d) = 1 - (1 + \pi d^2 n/A)^{-2} \tag{6}$$

The distribution of average velocity in Houston was obtained from the measured Euclidean distances between ambulance station and pickup locations for a random sample of 104 of the 3,936 runs. Dividing this distance by the response time (recorded to the nearest minute) we

Fig. 3. Histogram of the density of emergency demand levels in Houston.



obtained an estimate of the mean velocity for that run. Since ambulances generally are not required to be at their home stations, the calculated mean velocities are biased. However the overall average of the calculated mean velocities provides an unbiased estimate of the mean of $\bar{V}$. The average of the 104 calculated mean velocities was 22.5 miles per hour. We fitted the piecewise linear $f(\bar{v})$ shown in Fig. 4 to approximate a Normal fit for $f(\bar{v})$ with mean 22.5 miles per hour and standard deviation 5 miles per hour. The value for the mean of $f(\bar{v})$ is greater than that reported for some other communities [10] because

Fig. 4. Probability density function for the mean ambulance velocity en route to an emergency in Houston.

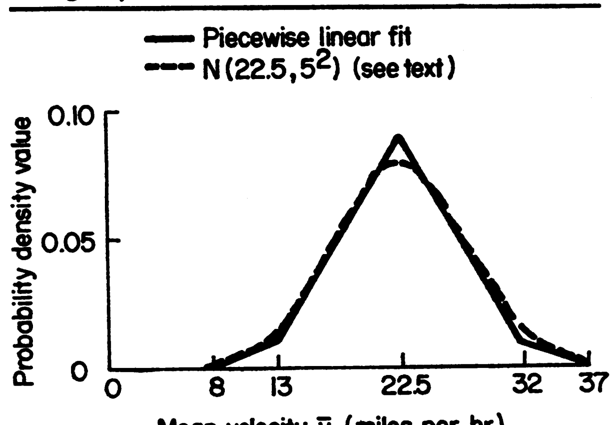
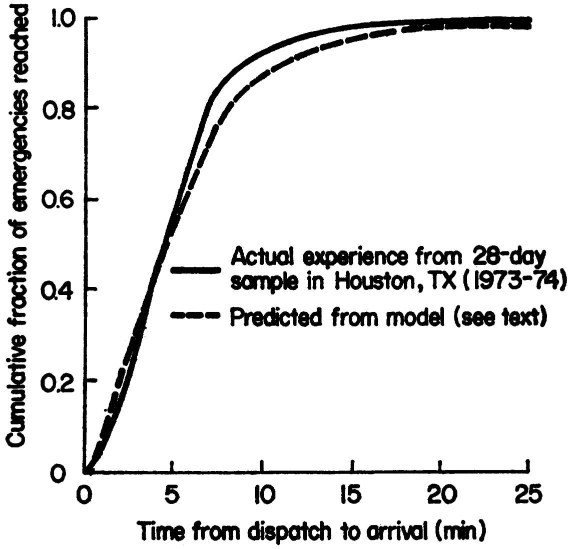


Fig. 5. Comparison of sample Houston ambulance data and the model's prediction of the fraction of emergencies reached as a function of time.



Houston is less densely populated than these other communities and has a road system that permits more rapid travel.

If the probability density function of mean velocities $f(\bar{v})$ is represented by a piecewise linear function $f(\bar{v}) = a_i\bar{v} + b_i$ for $\bar{v}_{i-1} < \bar{v} < \bar{v}_i$ for $i = 1, m$, then for the proportional allocation with response distance distribution given by Eq. 6, Eq. 3 becomes

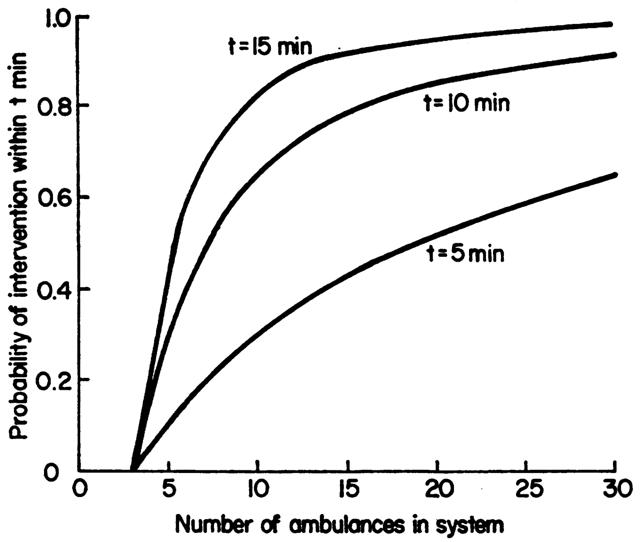
$$P_n(T \leqslant t) = 1 - \frac{1}{2}\sum_{i=1}^{m}\left[\frac{b_i\bar{v} - (a_i/f)}{1 + f\bar{v}^2} + \frac{b_i}{f^{\frac{1}{2}}} \arctan(f^{\frac{1}{2}}\bar{v})\right]_{\bar{v} = \bar{v}_{i-1}}^{\bar{v} = \bar{v}_i} \quad (7)$$

where $f = \pi t^2 n/A$. Then the desired distribution of response time is given by Eq. 5, where $P_n(T \leqslant t)$ is as in Eq. 7 and $Q_N(n)$ is given by Eq. 4.

We computed the empirical cumulative distribution of the time between the dispatch of an ambulance and its arrival at the scene of the emergency. Figure 5 shows this distribution as well as the model's predicted distribution of the response time. We believe that the close correspondence between the prediction and the empirical data justifies the approach we have taken.

In Fig. 6 we present the model's predictions of the probability that an ambulance would be within 5, 10, and 15 minutes of an emergency as a function of the total number of ambulances in Houston. For 23 ambulances the model predicts a 90-percent chance that an ambulance will reach an emergency within 10.7 minutes of its dispatch. (Recall

Fig. 6. The model's prediction of the probability of timely emergency intervention (dispatch to arrival) as a function of the number of ambulances in Houston.



the 90-percent standard of the EMSS Act.) If, however, 15 minutes is an acceptable response time, then only 14 ambulances are needed to achieve the 90-percent level. Whether the increase of 4.3 minutes in this portion of the response time can be justified by the cost savings of nine fewer ambulances is a problem for a further cost/benefit analysis.

The model can also be useful in planning for the future needs of the emergency system. If the area of the region and the shape of the emergency demand remain constant but the demand rate changes, we can evaluate the impact on ambulance system responsiveness. The accompanying table lists, for several choices of the average rate of demand, the model's predictions of the number of ambulances needed in Houston to insure that an ambulance will be within 10 minutes of an emergency with 90-percent probability. In this case, the required number of ambulances is relatively insensitive to the rate of demand. For example, if the 1973–74 demand rate of six calls per hour doubled but the geographical distribution of that demand remained unchanged, only three additional ambulances would be required to provide comparable service. On the other hand, if both the area and the demand rate doubled, then the model predicts that 52 ambulances would be required. When an ambulance system is performing well, for example reaching 90 percent of emergencies within 10 minutes, its responsiveness is much more sensitive to increases in the area of the service region than to increases in the rate of emergency calls.

Predicted Number of Ambulances for Houston Needed for a Response Time of 10 Minutes to 90 Percent of Emergencies

(Constant urban area, see text; 23 ambulances are required if queueing effects are ignored)

Average no. of calls/hr	Ambulances required*	Predicted demand intensity
3	25 (0.903)	0.06
6†	27 (0.905)	0.11
9	28 (0.902)	0.16
12	30 (0.904)	0.20
15	31 (0.901)	0.24
18	33 (0.903)	0.27

* Predicted coverage probabilities are in parentheses.
† 1973 demand level in Houston.

Other Allocation Strategies for Houston

In our analysis, we assumed that ambulances in Houston are allocated according to the function $h(\bar{z}) = \bar{z}$, the proportional strategy. Another allocation strategy for Houston is to allocate proportionally fewer ambulances to regions with high emergency demand. In that case, $h(\bar{z})$ might take the form of the square root of $\bar{z}$, allocating relatively more ambulances in low demand regions than the proportional strategy. On the other hand, high demand regions might actually require more ambulances because of queueing effects, so we also investigated allocation according to $h(\bar{z}) = \bar{z}^2$.

The distribution of response distance for these choices of $h(\bar{z})$ may be obtained analytically from Eq. 2. For $h(\bar{z}) = \bar{z}$, we obtain Eq. 6. For $h(\bar{z}) = \bar{z}^2$,

$$P_N(D \leqslant d) = 1 - f[1 - (2\pi f)^{1/2} \cdot \exp(f/2) \cdot \Phi(-f^{1/2})] \tag{8}$$

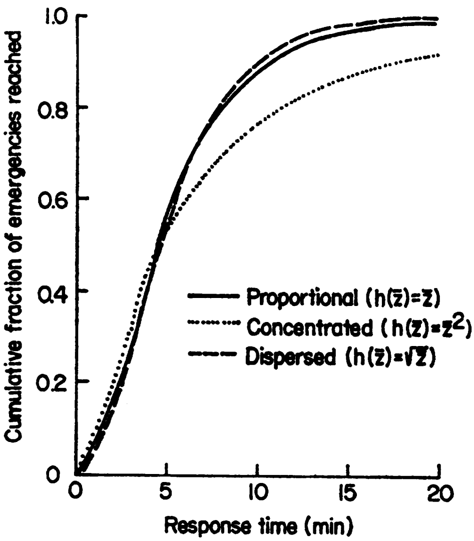
where $f = A/\pi d^2 N$, and Φ represents the cumulative distribution of a standardized Normal random variable. If $h(\bar{z})$ is the square root of $\bar{z}$, then

$$P_N(D \leqslant d) = f^2 [-1 + \pi^{1/2}(2f^2 + 3)/f \cdot \exp(f^2) \cdot \Phi(-2^{1/2} f)] \tag{9}$$

where $f = \pi^{1/2} d^2 N/A$.

To compare the potential effects of these allocation strategies on response times in Houston, we solved Eq. 3 numerically for response distance distributions given by Eqs. 8 and 9 and incorporated queueing effects using Eq. 5 with data from the Houston case study. The results are shown in Fig. 7, and it is apparent in this case that the "proportional" and "square root" strategies yield similar response time distributions, although the latter strategy seems slightly better. The "square" strategy, which concentrates ambulances in regions of high demand, is markedly inferior to the other two, unless very short response times are valued highly. We believe that the "proportional" and "square

Fig. 7. Comparison of the model's predictions of the distribution of response time in Houston for three ambulance allocation strategies.



root" strategies cover a range of plausible allocations and thus that Fig. 7 supports the notion that the response time distribution is quite insensitive to plausible choices of allocations. In turn, this suggests that our use of a probability model of locations was justified.

Discussion

Our use of probability theory to predict response times of an urban ambulance system yielded a model that complements those previously reported. From this model, the entire probability distribution of response time can be obtained from the distribution of the demand for service and from the number, locations, and velocities of ambulances. We were able to deal directly with both queueing and variable ambulance velocities, principally because of our representation of ambulance locations. Specifically, we assumed that the probability of an ambulance's being at a given point in the city is some function of the demand for emergency service at that point. This is intended as a descriptive rather than a normative formulation, and we believe that it is quite adequate for modeling the response of a broad range of ambulance systems. The optimal allocation of a given number of ambulances within a particular urban region remains an important problem, the solution of which must account for many idiosyncratic features of the region. The approaches of ReVelle et al. [4] and Larson [12] are powerful aids for dealing with this problem.

Because our model yields the entire response time distribution, it is useful for studying certain features of an urban ambulance system. With data from Houston, we showed how allocation strategies influence the response time distribution in general. Compared to Houston's allocation in 1973, which seems to have been in proportion to demand for service, a greater dispersion of ambulances away from centers of demand would have slightly improved the tail of the predicted response time distribution. A greater concentration of ambulances in areas of high demand would have resulted in much poorer performance, as measured by the tail of the cumulative distribution. The large area of Houston, the relatively high velocities of ambulances, and the relatively low demand intensity contributed to the latter result. Even with high demand intensities, however, the analysis for Houston shows that concentrating ambulances in regions of high demand increases the cumulative probabilities of service within short intervals but decreases the overall probability of service within intervals of 10 to 15 minutes.

We did not consider a number of performance measures such as delays in summoning and dispatching an ambulance. These are areas in which we plan to expand the existing model, but it may then not be possible to maintain its analytical structure, and simulation may be required [5–8].

Appendix

The two-dimensional histogram of emergency demand data is shown in Fig. 1 (p. 409). Each annulus contains 20 percent of the total emergency calls. The downtown area (the center of Houston) has a dramatic peak of demand, which decreases monotonically as the distance from the center increases. Thus the histogram has one mode, and the decrease from the peak is radially symmetric over a circle with radius a. The functional form chosen to fit the histogram is

$$H(x,y) = -\hat{E}/\pi a^2 \quad \ln[(x^2+y^2)/a^2] \qquad 0 \leqslant x^2+y^2 \leqslant a^2 \tag{A1}$$

where (x,y) are coordinates relative to the point (0,0) at the center of Houston. $\hat{E}$ is the expected number of calls in the city during the time period in question, 3,936 for our example. Houston's area is 400 square miles and so a was taken to be 11.28 miles.

To obtain $g(\hat{z})$, the distribution of demand density, from the fitted H(x,y), we proceed as follows. Consider a randomly selected point (X,Y) in the region. Let R be the distance from (0,0) to the point (X,Y), that is, $R^2 = X^2 + Y^2$. Substituting into Eq. A1 and noting $A = \pi a^2$, we obtain

$$\hat{Z} = H(R) = -2\hat{E}/A \quad \ln[R/a]$$

or alternatively

$$R = R(\hat{Z}) = a \quad \exp(-A\hat{Z}/2\hat{E}) \tag{A2}$$

The probability that a randomly selected point in the region will yield a radius less than or equal to r is r^2/a^2, so that the probability density function of R is

$$f(r) = 2r/a^2 \qquad 0 \leqslant r \leqslant a$$

Using this result with the absolute value of the Jacobian, J, of the transformation given in Eq. A2

$$J = aA/2\bar{E} \qquad \exp(-A\bar{z}/2\bar{E})$$

yields $g(\bar{z})$ by the change-of-variables technique:

$$g(\bar{z}) = f[R(\bar{z})]J = A/\bar{E} \qquad \exp(-A\bar{z}/\bar{E}) \tag{A3}$$

which is a negative exponential distribution with parameter $A/\bar{E}$.

Acknowledgments. The authors would like to thank Lanie J. Benson for assistance and the referees for their comments, which led to substantial revision of an earlier version.

REFERENCES

1. Gibson, G. Evaluation criteria for emergency ambulance systems. *Soc Sci Med* 7:425 June 1973.
2. *Emergency Medical Services Systems Act. Federal Register*, Vol. 39, No. 62, Part 3, pp. 11758–11766 Mar. 29, 1974.
3. Willemain, T.R. The status of performance measures for emergency medical services. *J Am Coll Emergency Physician* 4:143 Mar.–Apr. 1975.
4. ReVelle, C., D. Bigman, D. Schilling, J. Cohon, and R. Church. Facility location: A review of context-free and EMS models. *Health Serv Res* 12:129 Summer 1977.
5. Savas, E.S. Simulation and cost-effectiveness analysis of New York's emergency ambulance service. *Manage Sci* 15:B608 Aug. 1969.
6. Fitzsimmons, J.A. A methodology for emergency ambulance deployment. *Manage Sci* 19:627 Feb. 1973.
7. Swoveland, C., D. Uyeno, I. Vertinsky, and R. Vickson. Ambulance locations: A probabilistic enumeration approach. *Manage Sci* 19:686 Dec. 1973.
8. Berlin, G.N. and J.C. Liebman. Mathematical analysis of emergency ambulance location. *Socio-Economic Plann Sci* 8:323 Dec. 1974.
9. Stevenson, K.A. Emergency Ambulance Transportation. In A. Drake (ed.), *Analysis of Public Systems*, pp. 133–149. Cambridge, MA: M.I.T. Press, 1972.
10. Cretin, S. *A Model of the Risk of Death from Myocardial Infarction*. Technical Report No. TR-09-74. Cambridge, MA: M.I.T. Operations Research Center, Nov. 1974.
11. Larson, R.C. A hypercube queueing model for facility location and redistricting in urban emergency services. *Comput Oper Res* 1:67, 1974.
12. Larson, R.C. Approximating the performance of urban emergency service facilities. *Oper Res* 23:845 Sept.–Oct. 1975.
13. Hakimi, S. Optimum locations of switching centers and the absolute centers and medians of a graph. *Oper Res* 12:450 May–June 1964.
14. ReVelle, C. and R. Swain. Central facilities location. *Geogr Anal* 11:30 Jan. 1970.
15. Toregas, C., R. Swain, C. ReVelle, and L. Bergman. The location of emergency service facilities. *Oper Res* 19:1363 Oct. 1971.
16. Larson, R.C. and K.A. Stevenson. On insensitivities in urban redistricting and facility location. *Oper Res* 20:595 May–June 1972.
17. Kendall, M.G. and P.A.P. Moran. *Geometric Probability*. London: Charles Griffin, 1963.
18. Hillier, F.S. and G.J. Lieberman. *Introduction to Operations Research*. San Francisco: Holden-Day, 1967.
19. Benson, L.J. *Utilization Study: Houston Fire Department Emergency Ambulance Service*. M.P.H. thesis, School of Public Health, University of Texas, Houston, 1976.