# SHORTEST PATHS IN PROBABILISTIC GRAPHS

**H. Frank**

*Office of Emergency Preparedness, Executive Office of the President, Washington, D.C.*

(Received June 26, 1968)

This paper considers the problem of finding shortest-path probability distributions in graphs whose branches are weighted with random lengths, examines the consequences of various assumptions concerning the nature of the available statistical information, and gives an exact method for computing the probability distribution, as well as methods based on hypothesis testing and statistical estimation. It presents Monte Carlo results and, based on these results, it develops an efficient method of hypothesis testing. Finally, it discusses briefly the pairwise comparison of paths.

LET $G$ BE a graph whose branches $b_1, b_2, \cdots, b_m$ are weighted with nonnegative real numbers. The number $W_i$ associated with a given branch $b_i$ is called the length of $b_i$. Let the vertices of $G$ be labeled $v_1, v_2, \cdots, v_n$ and let $\Pi_i(s, t)$ be a path from $v_s$ to $v_t$. Then a classical problem in graph theory is to find the *shortest* path from $v_s$ to $v_t$ where the length $|\Pi_i(s, t)|$ of $\Pi_i(s, t)$ is defined as

$$|\Pi_i(s, t)| = \sum_{b_j \epsilon \Pi_i(s,t)} W_j, \tag{1}$$

and the shortest $s$-$t$ path is any path $\Pi_j(s, t)$ with

$$|\Pi_j(s, t)| = \min_i [|\Pi_i(s, t)|]. \tag{2}$$

An enormous amount of attention has been given to finding shortest-path algorithms for graphs with fixed, deterministic branch weights.[1,2] A number of very efficient algorithms exist for finding the shortest path between a particular pair of vertices, from a given vertex to all other vertices, or between all pairs of vertices. If all branch weights are nonnegative, these algorithms are equally effective for directed or undirected graphs. On the other hand, almost no attention has been given to the situation where branch weights are random variables with possibly known probability distributions.

In numerous instances, the length of a branch is the measure of the cost or time needed to traverse that branch or to complete a given operation. In practical situations, such times or costs are often random. Hence, the shortest path is a function of random events characterized by probability distributions. In this paper, we examine the problems of finding or estimating the probability distribution of the shortest path, testing hypotheses

concerning the shortest-path distribution, and comparing paths according to probabilistic length criteria. These problems will be considered under a variety of different assumptions about the nature of the probabilistic information available.

## EXACT PROBABILISTIC ANALYSIS

LET $G$ be an $n$-vertex $m$-undirected branch graph (*see* Note 1), and let $W_i$ be a random variable associated with branch $b_i$ for $i=1, \cdots, m$. Assume that $W_i$ is a continuous random variable and that the joint probability density of the random vector $\boldsymbol{W} \triangleq (W_1, W_2, \cdots, W_m)$ is $p_W(w_1, w_2, \cdots, w_m)$. Let $\Pi_1(s, t), \Pi_2(s, t), \cdots, \Pi_{q_{s,t}}(s, t)$ be the set of $s$-$t$ paths of $G$ that, for convenience, will be denoted by $\Pi_1, \Pi_2, \cdots, \Pi_q$. Let $M$ be the random variable

$$M = \min_i [|\Pi_i|] = \min_i [\textstyle\sum_{b_j \epsilon \Pi_i} W_j]. \tag{3}$$

$M$ obviously corresponds to the length of the shortest $s$-$t$ path. In this section, we discuss the general problem of computing the probability distribution $F(l)$ of $M$, where

$$F(l) = \text{prob}\,\{M < l\}. \tag{4}$$

Let $|\boldsymbol{\Pi}| \triangleq (|\Pi_1|, |\Pi_2|, \cdots, |\Pi_q|)$. To compute $F(l)$ we must first find the probability density $p_\pi(x_1, x_2, \cdots, x_q)$ of $|\Pi|$. Then, $F(l)$ is easily found as

$$F(l) = 1 - \int_l^\infty \cdots \int_l^\infty p_\pi(x_1, \cdots, x_q)\, dx_1, \cdots, dx_n. \tag{5}$$

The probability density $p_\Pi(x_1, \cdots, x_q)$ can be found by the following method.[3,4] Let $\boldsymbol{w} = (w_1, w_2, \cdots, w_m)$ be a particular value of $\boldsymbol{W}$. Given $\boldsymbol{W} = \boldsymbol{w}$, then $|\boldsymbol{\Pi}| = \mathbf{x} \triangleq (x_1, \cdots, x_q)$, where

$$\boldsymbol{x} = \boldsymbol{w} P_{s,t}, \tag{6}$$

and $P_{s,t} = [p_{ij}]_{m \times q}$ is a matrix of zeros and ones such that $p_{ij} = 1$ if and only if $b_j \epsilon \pi_i$.

Let $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_m)$, $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_q)$ and $\varphi(\boldsymbol{\alpha})$ and $\psi(\boldsymbol{\beta})$ be the characteristic functions of $\boldsymbol{W}$ and $|\boldsymbol{\Pi}|$, respectively. By definition,

$$\varphi(\boldsymbol{\alpha}) = E\{e^{i\boldsymbol{\alpha}\boldsymbol{W}}\} \triangleq \int_0^\infty \cdots \int_0^\infty \cdot \exp\left[ i \sum_{j=1}^m \alpha_j w_j \right] p_W(w_1, \cdots, w_m)\, dw_1, \cdots, dw_m \tag{7a}$$

and

$$\psi(\mathfrak{z}) = E\{e^{i\mathfrak{z}|\mathbf{\Pi}|}\} \triangleq \int_0^\infty \cdots \int_0^\infty \cdot \exp\left[i \sum_{j=1}^{j=q} \beta_j x_j\right] p_\Pi(x_1, \cdots, x_q)\, dx_1, \cdots, dx_q, \tag{7b}$$

where $E$ represents the expectation operator.

Then, since $|\mathbf{\Pi}| = \mathbf{W}P_{s,t}$ and since the expectation of a function of a random variable can be computed as the integral of that function and the original variable,

$$\psi(\mathfrak{z}) = \int_0^\infty \cdots \int_0^\infty \exp\left[i \sum_{j=1}^{j=q} \beta_j \sum_{k=1}^{k=m} w_k P_{kj}\right] p_{\mathbf{W}}(w_1, \cdots, w_m)\, dw_1, \cdots, dw_m. \tag{8}$$

In other words,

$$\begin{aligned} \psi(\mathfrak{z}) &= \int_0^\infty \cdots \int_0^\infty \exp\left[i \sum_{k=1}^{k=m} \left\{\sum_{j=1}^{j=q} \beta_j P_{k,j}\right\} w_k\right] \\ &\quad \cdot p_w(w_1, \cdots, w_m)\, dw_1, \cdots, dw_m \\ &= \varphi(\mathfrak{z} P_{s,t}). \end{aligned} \tag{9}$$

Now that the characteristic function $\psi(\mathfrak{z})$ of $|\mathbf{\Pi}|$ is known, we can attempt to find the probability density of $|\mathbf{\Pi}|$ by inversion.[3] However, although $\psi(\mathfrak{z})$ is well defined, the corresponding density may not be unless the rank of $P_{s,t}$ is $q$. Suppose the rank of $P_{s,t}$ is $r < q$. Instead of finding the characteristic function of $(|\Pi_1|, \cdots, |\Pi_q|)$, we must find the characteristic function of the random vector $|\hat{\mathbf{\Pi}}| = (|\Pi_{i_1}|, |\Pi_{i_2}|, \cdots, |\Pi_{i_r}|)$ where rows $i_1, i_2, \cdots, i_r$ of $P_{s,t}$ are a maximum set of linearly independent rows of $P_{s,t}$. Let $\hat{P}_{s,t}$ be the submatrix of $P_{s,t}$ corresponding to these rows. Then,

$$|\Pi_j| = \sum_{k=1}^{k=r} d_{j,i_k} |\Pi_{i_k}|, \tag{10}$$

where the $d_{j,i_k}$ are constants found by Gauss reducing $P_{s,t}$. The characteristic function $\hat{\psi}(\hat{\mathfrak{z}})$ where $\hat{\mathfrak{z}} = (\beta_{i_1}, \beta_{i_2}, \cdots, \beta_{i_r})$ is simply

$$\hat{\psi}(\hat{\mathfrak{z}}) = \varphi(\hat{\mathfrak{z}} \hat{P}_{s,t}), \tag{11}$$

and if $\hat{p}_{\hat{\Pi}}(y_1, \cdots, y_r)$ is the nonsingular density function of $|\hat{\Pi}|$, then

$$F(l) = 1 - \int \cdots \int_{\Omega} \hat{p}_{\hat{\Pi}}(y_1, \cdots, y_r)\, dy_1, \cdots, dy_r, \tag{12}$$

where $\Omega$ is the convex region defined by the set of linear inequalities

$$\sum_{k=1}^{k=r} d_{j,i_k} y_k \geqq l \quad \text{for} \quad j = 1, \cdots, q. \tag{13}$$

The computation of $\hat{p}_{\hat{n}}(y_1, \cdots, y_r)$ and $F(l)$ each require $r$ integrations where $r \leqq m$. Moreover, the region $\Omega$ over which the integration in (12) must be performed is complicated. However, the above procedure could be useful in networks with relatively few branches. Moreover, the original problem may be formulated so that it is necessary to find the distribution of the minimum length path in a *subset* of the $s$-$t$ paths of $G$. For example, it may be desirable to examine only the set of $s$-$t$ paths that contain fewer than $k<n-1$ branches. In this case the computational difficulties are greatly reduced. It should also be noted that if the random lengths come from discontinuous distributions, then one should replace all Riemann integrals by Stieltjes integrals. In the case of purely discrete distributions all integrals can be replaced by summations.

## NONPARAMETRIC STATISTICAL ANALYSIS

LET US assume that the probability distributions of $W_1, \cdots, W_m$ are unknown but that a set of $K$ time observations may be made. Let $\boldsymbol{W}(k) = [W_1(k), \cdots, W_m(k)]$ be the random vector corresponding to the value of $\boldsymbol{W}$ at time $k$. We will assume that $\boldsymbol{W}(1), \cdots, \boldsymbol{W}(k)$ are indentically and independently distributed.

Let $|\Pi_1(k)|, |\Pi_2(k)|, \cdots, |\Pi_q(k)|$ be the random values of the $s$-$t$ paths $\Pi_1, \cdots, \Pi_q$ at the time $k$, and let $m(k) = \min_i \{|\Pi_i(k)|\}$.

The problem we consider here is to determine whether prob $\{M \geqq l\}$ is at least some quantity $p_0$. Hence, we pose the following problem: Test the hypothesis $H_1: p \overset{\Delta}{=} 1-F(l) \geqq p_0$ against the alternative $H_2: p < p_0$, at level $\alpha$, where $\alpha$ is the probability of Type I error.

Consider the following statistical test. If $m_1, \cdots, m_K$ are the path lengths at times $1, \cdots, K$, reorder the $m_k$'s such that

$$m_{i_1}, m_{i_2}, \cdots, m_{i_N} < l \leqq m_{j_1}, m_{j_2}, \cdots, m_{j_{K-N}}. \tag{14}$$

Then, if $\hat{K}$ and $\gamma (0 \leqq \gamma \leqq 1)$ are predetermined constants:

Reject hypothesis $H_1$ if $N > \hat{K}$. (15a)

Reject hypothesis $H_1$ with probability $\gamma$ if $N = \hat{K}$. (15b)

Accept hypothesis $H_1$ if $N < \hat{K}$. (15c)

The constants $\hat{K}$ and $\gamma$ are determined from

$$\sum_{i=\hat{K}+1}^{i=K} \binom{K}{i} (1-p_0)^i p_0^{K-i} + \gamma \binom{K}{\hat{K}} (1-p_0)^{\hat{K}} p_0^{K-\hat{K}} = \alpha. \tag{15d}$$

The left-hand side of (15d) is readily seen to be prob $\{X > \hat{K}\} + \gamma$ prob $\{X = \hat{K}\}$, where $X$ is a binomial random variable with parameters $1-p_0$ and

$K$. Hence $\hat{K}$ and $\gamma$ are easily found from tables of the binomial distribution.

It may be shown that the test suggested above is uniformly most powerful;[5] in other words, it is optimum in the sense that if $H_1$ is true ($p \geqq p_0$), the probability of making a wrong decision (i.e., rejecting $H_1$) is $\leqq \alpha$, and if $H_1$ is false ($p < p_0$), the probability of making a wrong decision (i.e., accepting $H_1$) is *minimum*, over all possible statistical tests that utilize the same information. However, although the probability of making the incorrect decision if $p < p_0$ (Type II error) is minimum, it still may be large. The magnitude of this probability is a function of both the sample size and the true value of the parameter and will decrease as the sample size becomes larger and the true value of $p$ becomes smaller. If the 'cost' of making a Type II error is greater than the 'cost' of making a Type I error, it would be more advisable to formulate the dual statistical testing problem $H_1: p \leqq p_0$ versus $H_2: p > p_0$. However, the details of the solution are similar to the test discussed here and will not be considered further. Also, the above test may be generalized to a sequential test.[6,7] However, this generalization is straightforward and will not be discussed here.

## MONTE CARLO SIMULATION

A HIGHLY desirable goal is to compute the probability distribution of the shortest path length. Because of the great complications that arise in the exact computation procedure, it seems reasonable to expect that approximations are necessary to obtain tractable results. One such approximation can be made by using Monte Carlo programming.

We simulate the system by generating a set of random branch-length vectors according to given probability distributions. We then apply a shortest path algorithm to compute the value of $M$ for this simulation. This process is repeated $K$ times and the empirical distribution function $S_K(z)$ is defined as

$$S_K(z) \triangleq (1/K) \sum_{k=1}^{k=K} h_z(m_k), \tag{16a}$$

where

$$h_z(u) = \begin{cases} 0, & u \geqq z, \\ 1, & u < z. \end{cases} \tag{16b}$$

Thus, $KS_K(z)$ is the number of $m_k$ that are smaller than $z$. There is a very strong relation between $S_K(z)$ and the true probability distribution of $M$.[7] Hence, it is desirable to examine $S_K(z)$ when $K$ is large, since in a very strong sense $S_K(z)$ converges to $F(z)$.

A computer program was written to perform the following steps:

1. Generate a graph.

**2.** Generate random branch lengths according to specified probability distributions.

**3.** Calculate the shortest path in the graph between a specified pair of vertices.

**4.** Repeat steps 2–3 $K$ times and plot $S_K(z)$.

The above investigation was applied to both graphs generated randomly and to graphs selected from a class of regular graphs. Branch lengths were selected to be uniformly distributed, normally distributed, or exponentially distributed. The $S_K(z)$ were plotted, curves such as the one shown in Fig. 1b for the graph shown in Fig. 1a were obtained. In Fig. 1b, the circles represent observation points while the solid curve is one minus the cumulative distribution function of a normal random variable. This simulation was performed for a large number of cases. In nearly all cases the results of the simulation indicated that $F(l)$ could be accurately approximated by a normal distribution function. To investigate this conjecture, the $\chi^2$ goodness-of-fit test for normal distributions was applied.[7]

In the $\chi^2$ test, the hypothesis that $F(l)$ is normal is tested against the alternative that it is not normal. A sample $m_1, \cdots, m_K$ of $M$ is taken, reordered such that $m_1 \leqq m_2 \leqq \cdots \leqq m_K$ and then grouped into $r$ disjoint sets $T_1, T_2, \cdots, T_r$ such that $T_k = [a_k, a_{k+1}]$, where $a_k < a_{k+1}$, $a_0 = 0$, and $a_{r+1} = \infty$. Suppose there are $n_k$ observations in group $T_k$. For different samples, the $n_k$ vary and are therefore random variables. The sample mean $\bar{M}$ and sample variance $V^2$ are calculated from the equations

$$\bar{M} = (1/K) \sum_{k=1}^{k=K} m_k, \tag{17a}$$

$$V^2 = (1/K) \sum_{k=1}^{k=K} (m_k - \bar{M})^2. \tag{17b}$$

The 'theoretical frequences' $x_1, \cdots, x_r$ are then defined by

$$x_k = \Phi[(a_{k+1} - \bar{M})/V] - \Phi[(a_k - \bar{M})/V], \tag{18}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal variable. The value of the statistic

$$\chi_0^2 \overset{\Delta}{=} \sum_{k=1}^{k=r} (n_k - Kx_k)^2 / Kx_k \tag{19}$$

is calculated; $\chi_0^2$ has the $\chi^2$ distribution with $r-3$ degrees of freedom if the hypothesis is true. Finally, given the significance level $\alpha$ of the test (say $\alpha = 0.05$), a constant $X_0$ such that prob $\{\chi^2 > X_0\} = \alpha$ is computed. If $\chi_0^2 \leqq X_0$ the hypothesis of normality is accepted, while if $\chi_0^2 > X_0$ the hypothesis of normality is rejected. As a measure of acceptance level of the test, prob $\{\chi^2 > \chi_0^2\}$ may be considered. The larger this probability, the more certain the normality conjecture.

The $\chi^2$ test was performed for a wide variety of graphs and branch length probability distributions. The class of regular graphs as illustrated

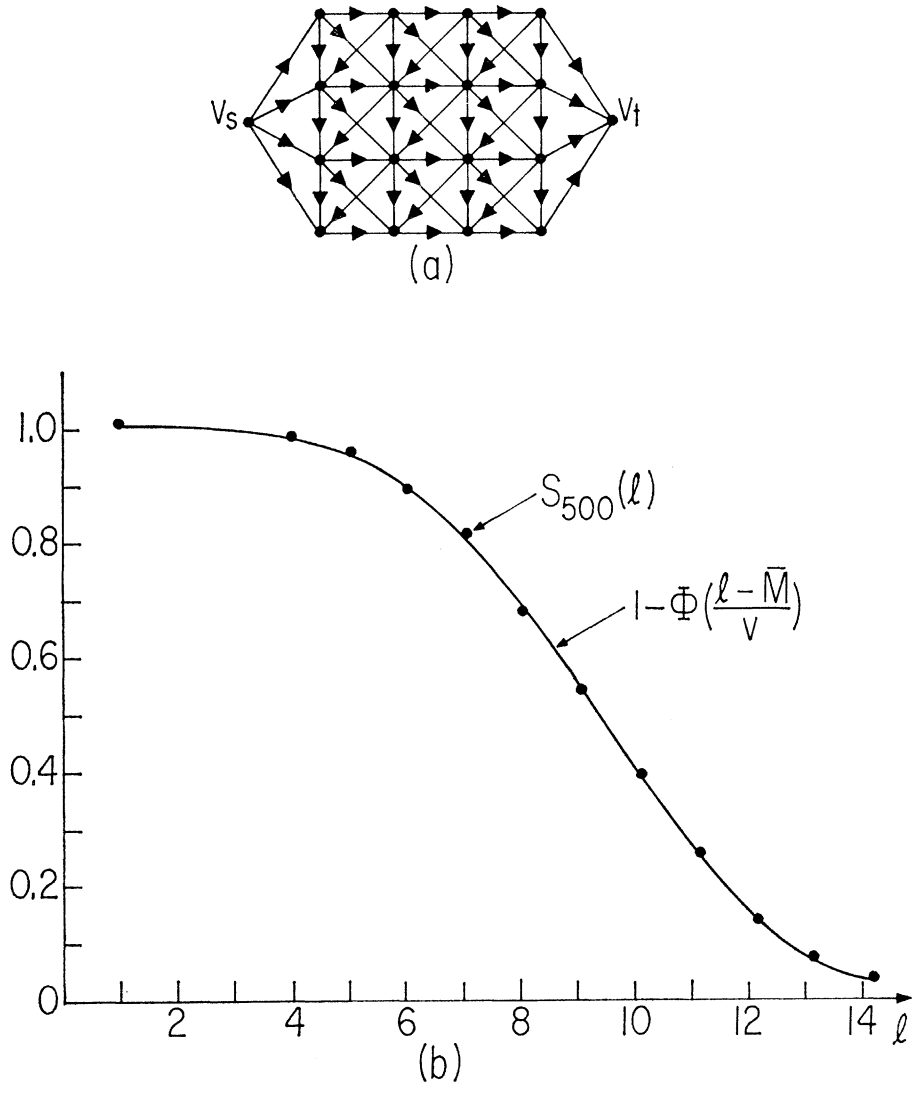


**Fig. 1.** Empirical distribution function and normal approximation for graph shown in (a).

in Fig. 2 was investigated as a function of the number of vertices in the graph. Examples were considered in which all branch lengths were uniform random variables, normal random variables, or exponential random variables. Graphs randomly generated by removing branches uniformly at random from a complete graph were also considered and the $\chi^2$ test evalu-

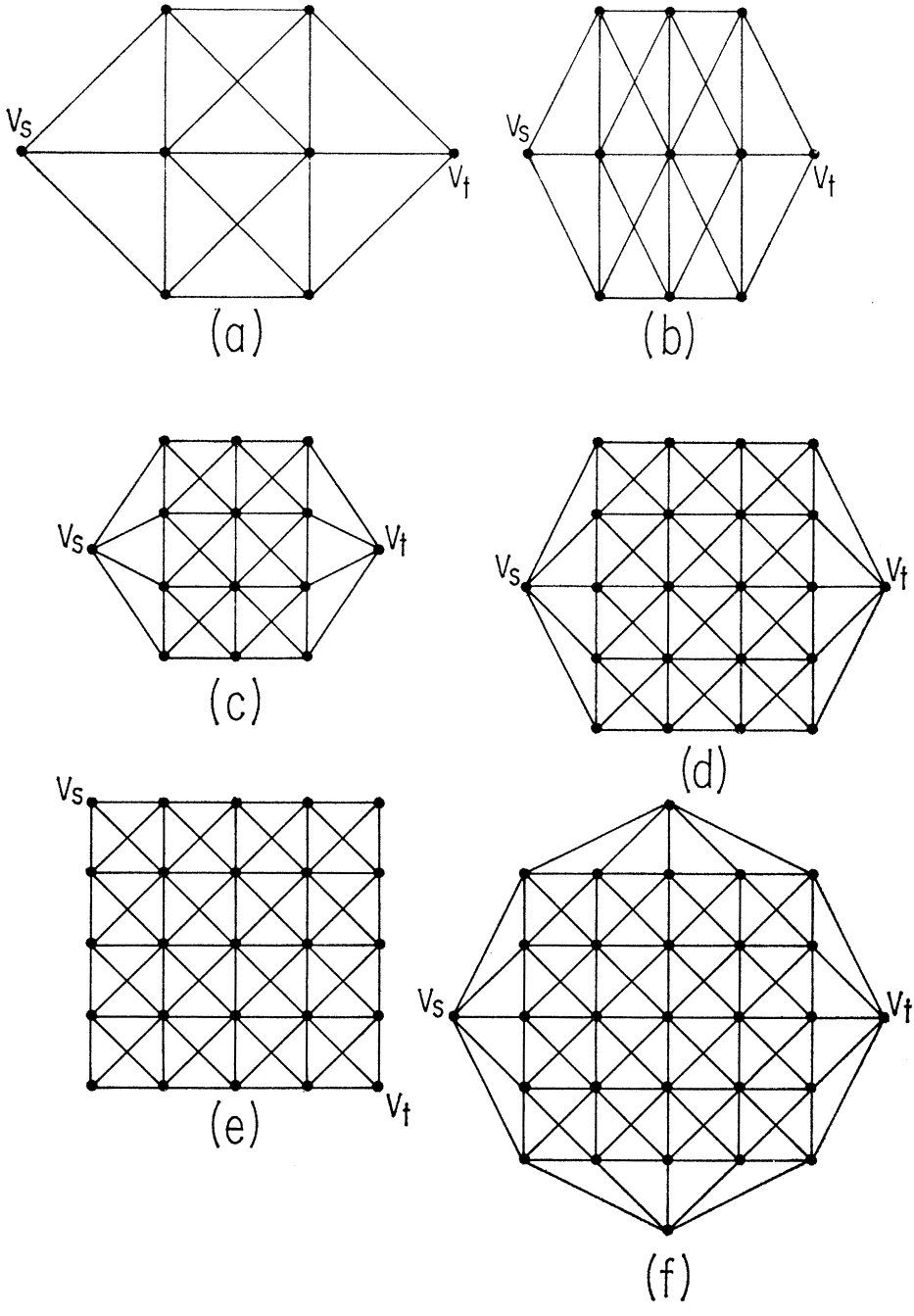


**Fig. 2.** Deterministic graphs used in simulation.

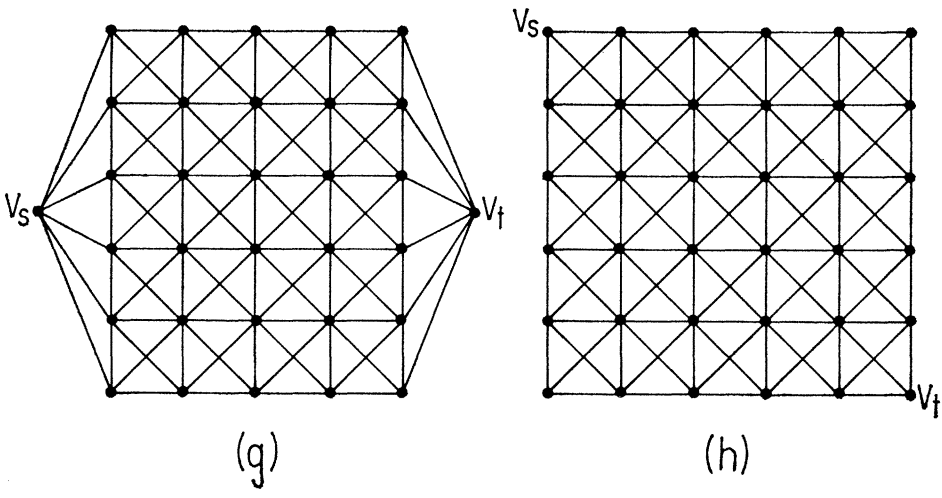


**Fig. 2**—*Continued.*

ated. Two typical random graphs are shown in Fig. 3. The results of this evaluation are shown in Tables I–VIII. From these results, it seems reasonable to conclude that the shortest path in a graph with a large number of vertices can often be approximated as a normal random variable.

## PARAMETRIC STATISTICAL ANALYSIS

IN THIS SECTION, we pursue the observation that $M$ is approximately normally distributed. Again, let $\boldsymbol{W}(k)=[W_1(k), \cdots, W_m(k)]$ be a measurement of the branch lengths at time $k$ for $k=1, \cdots, K$ and assume that $\boldsymbol{W}(1), \cdots, \boldsymbol{W}(K)$ are identically and independently distributed. We want to test the hypothesis

$$H_1: p=\text{prob}\{M \geqq l\} \geqq p_0 \tag{20a}$$

against the alternative

$$H_2: p<p_0 \tag{20b}$$

at level $\alpha$, where $\alpha$ is the probability of Type I error.

If we assume that $M$ is normally distributed with unknown mean $\nu$ and unknown variance $\sigma^2$, we have a parametric testing problem. Also, $M_1, M_2, \cdots, M_K$ are identically and independently distributed $N(\nu, \sigma^2)$ random variables, and the hypothesis $p \geqq p_0$ is equivalent to

$$1-\Phi[(l-\nu)/\sigma] \geqq p_0 \tag{21a}$$

or

$$\Phi[(l-\nu)/\sigma] \leqq 1-p_0, \tag{21b}$$

where $\Phi$ is the cumulative distribution function of the standard normal

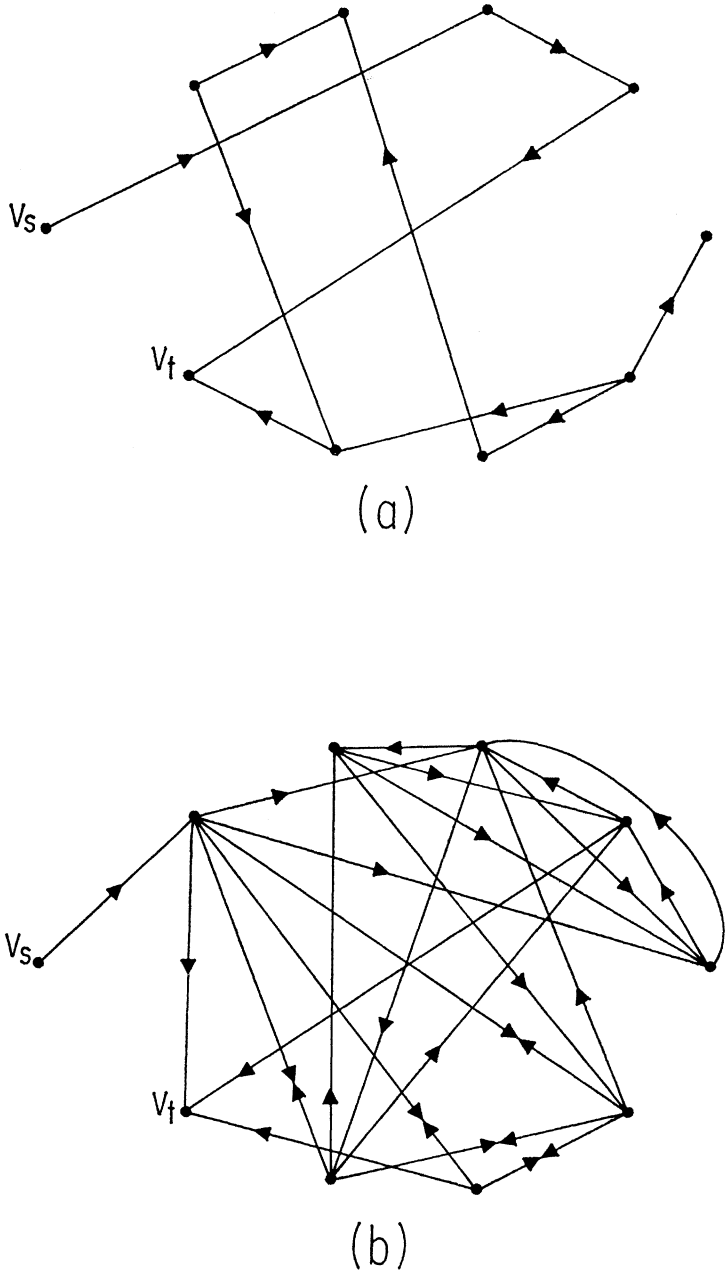


**Fig. 3.** Two random graphs used in simulation.

TABLE I

NORMAL DISTRIBUTION; REGULAR GRAPHS

| No. of vertices | 8 | 11 | 14 | 18 | 22 | 25 | 29 | 32 | 36 |
|---|---|---|---|---|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.18698 | 0.09775 | 0.23393 | 0.00404 | 0.67306 | 0.36426 | 0.39392 | 0.86190 | 0.54243 |
| | 0.00108 | 0.00066 | 0.36419 | 0.66745 | 0.12503 | 0.66040 | 0.35767 | 0.28739 | 0.56413 |
| | 0.12659 | 0.03912 | 0.14715 | 0.18127 | 0.09999 | 0.16080 | 0.36079 | 0.03967 | 0.50868 |
| | 0.31558 | 0.66967 | | | | | | | |
| Average $P(\chi^2 \geqq \chi_0^2)$ | 0.170 | 0.208 | 0.248 | 0.284 | 0.316 | 0.397 | 0.371 | 0.399 | 0.538 |

TABLE II

EXPONENTIAL DISTRIBUTION; REGULAR GRAPHS

| No. of vertices | 8 | 11 | 14 | 18 | 22 | 25 | 29 | 32 | 36 |
|---|---|---|---|---|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.00003 | 0.00041 | 0.00819 | 0.05326 | 0.00007 | 0.22908 | 0.00016 | 0.04692 | 0.18171 |
| | 0.00003 | 0.00006 | 0.00012 | 0.20935 | 0.02322 | 0.00036 | 0.00091 | 0.04016 | 0.03946 |
| | 0.00004 | 0.00092 | 0.00091 | 0.15015 | 0.04123 | 0.00002 | 0.07707 | 0.46283 | 0.10352 |
| | | 0.00000 | 0.00034 | 0.00000 | 0.04815 | 0.00041 | 0.79451 | 0.00562 | 0.00151 |
| | | | 0.00011 | 0.01188 | 0.00104 | 0.09802 | 0.93922 | | 0.04672 |
| | | | | 0.16496 | 0.02367 | | 0.00034 | | |
| Average $P(\chi^2 \geqq \chi_0^2)$ | 0.00003 | 0.00037 | 0.00193 | 0.098 | 0.0229 | 0.0656 | 0.15 | 0.14 | 0.075 |

TABLE III

UNIFORM DISTRIBUTION; REGULAR GRAPHS

| No. of vertices | 8 | 11 | 14 | 18 | 22 | 25 | 29 | 32 | 36 |
|---|---|---|---|---|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.32822 | 0.05838 | 0.73065 | 0.24385 | 0.77014 | 0.97190 | 0.87405 | 0.85665 | 0.91971 |
| | 0.36637 | 0.34084 | 0.57546 | 0.51149 | 0.44546 | 0.56509 | 0.47284 | 0.61410 | 0.38149 |
| | 0.00976 | 0.12746 | 0.55990 | 0.78015 | 0.32641 | 0.71028 | 0.42732 | 0.51219 | |
| | 0.08801 | 0.59367 | 0.23284 | 0.92012 | 0.20962 | 0.32637 | 0.47641 | 0.40032 | |
| | 0.32370 | | | 0.14964 | 0.99511 | | | | |
| Average $P(\chi^2 \geqq \chi_0^2)$ | 0.223 | 0.280 | 0.52 | 0.52 | 0.549 | 0.643 | 0.562 | 0.596 | 0.65 |

TABLE IV

NORMAL DISTRIBUTION; REGULAR GRAPHS; 18 VERTICES

| No. of branches | 32 | 41 | 50 | 56 |
|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.15184 | 0.35509 | 0.03419 | 0.42990 |
| | 0.19135 | 0.17135 | 0.99256 | 0.88456 |
| Average $P(\chi^2 \geqq \chi_0^2)$ | 0.1716 | 0.26322 | 0.51327 | 0.65723 |

TABLE V

UNIFORM DISTRIBUTION; REGULAR GRAPHS; 18 VERTICES

| No. of branches | 32 | 41 | 50 | 56 |
|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.34898<br>0.29829<br>0.17205 | 0.65336<br>0.28889<br>0.23148 | 0.20364<br>0.32431<br>0.78997 | 0.24376<br>0.82012<br>0.51620<br>0.31831 |
| Average $P(\chi^2 \geqq \chi_0^2)$ | 0.273 | 0.391 | 0.439 | 0.477 |

TABLE VI

NORMAL DISTRIBUTION; RANDOMLY GENERATED GRAPHS

| No. of vertices | 10 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.90667 | 0.97187 | 0.991189<br>0.14646 | 0.30533<br>0.46568 | 0.33010<br>0.41750 |

TABLE VII

UNIFORM DISTRIBUTION; RANDOMLY GENERATED GRAPHS

| No. of vertices | 5 | 10 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.03505 | 0.77400<br>0.17150<br>0.29039 | 0.34219<br>0.02600 | 0.16502 | 0.20516 | 0.07437 |

TABLE VIII

EXPONENTIAL DISTRIBUTION; RANDOMLY GENERATED GRAPHS

| No. of vertices | 25 | 30 | 35 | 40 |
|---|---|---|---|---|
| $P(\chi^2 \geqq \chi_0^2)$ | 0.03543<br>0.02126 | 0.01739<br>0.16625 | 0.03934 | 0.38263 |

variable; $\Phi$ and $\Phi^{-1}$ are strictly increasing functions. Therefore, (21a) and (21b) are also equivalent to

$$(l-\nu)/\sigma \leqq \Phi^{-1}(1-p_0), \tag{21c}$$

or, if we test in terms of the random variables $M_1^*, M_2^*, \cdots, M_K^*$ defined by $M_i^* = M_i - l$, $i=1, \cdots, K$, this is the same as

$$\xi/\sigma \geqq \theta_0 \stackrel{\Delta}{=} -\Phi^{-1}(1-p_0), \tag{22}$$

where $\xi = \nu - l$ is the mean of $M_i^*$ and $\sigma^2$ is the variance of $M_i^*$.

We now have the modified hypothesis and alternative which are

$$H_1^*: \xi/\sigma \geqq \theta_0 \tag{23a}$$

against

$$H_2^*: \xi/\sigma < \theta_0. \tag{23b}$$

This testing problem is discussed by LEHMANN (Sec. 6.4, reference 5). To test $H_i^*$ on the basis of the observations $m_k^* = m_k - l$, we need concern ourselves only with the variables

$$\bar{M} \stackrel{\Delta}{=} (1/K) \sum_{i=1}^{i=K} M_i^* \tag{24a}$$

and

$$S \stackrel{\Delta}{=} \sqrt{\sum_{i=1}^{i=K} (M_i^* - \bar{M})^2}. \tag{24b}$$

These variables are independent and are sufficient statistics for $(\xi, \sigma)$ (*see* Note 2). Furthermore, $\bar{M}$ is $N(\xi, \sigma^2/K)$, and the density $h(u)$ of the random variable $U \stackrel{\Delta}{=} S/\sigma$ is

$$h(u) = \begin{cases} [K^{(K-1)/2} u^{K-2} \exp\{-Ku^2/2\}]/K^{(K-1)/2}\Gamma[(K-1)/2], & (u>0) \\ 0. & (u \leqq 0) \end{cases} \tag{25}$$

We are testing $\theta \stackrel{\Delta}{=} \xi/\sigma$; a sufficient statistic that depends only on $\theta$ is

$$T_\theta(m_1, \cdots, m_K) = \sqrt{K}\bar{M}/(S/\sqrt{K-1}). \tag{26}$$

In fact, $T$ is a maximal invariant (*see* Note 3) under the transformation group 'multiplication by a positive constant.' The probability density $g_\theta(t)$ of $T$ is

$$g_\theta(t) = K^* \int_0^\infty \exp\left\{-\frac{1}{2}\left(t\sqrt{\frac{w}{K-1}} - \sqrt{K}\theta\right)^2\right\} w^{(K-2)/2} \exp\left(-\frac{w}{2}\right) dw, \tag{27}$$

where $$K^* = \{2^{K/2}\Gamma[(K-1)/2]\sqrt{\pi(K-1)}\}^{-1}.$$

Suppose we consider the class of all tests that depend on $M_1^*, \cdots, M_K^*$ only through the maximal invariant $T$. Then, it can be shown[5] that the uniformly most powerful invariant test (i.e., the UMP test among this class) for testing $H_1^*: \theta \geqq \theta_0$ against $H_2^*: \theta < \theta_0$, at level $\alpha$, is:

$$\text{Reject } H_1^* \text{ if } T_\theta(m_1, \cdots, m_K) \leqq \hat{K}. \tag{28a}$$

$$\text{Accept } H_1^* \text{ if } T_\theta(m_1, \cdots, m_K) > \hat{K}. \tag{28b}$$

Here $\hat{K}$ is a constant determined by

$$\int_{-\infty}^{\hat{K}}\int_{0}^{\infty} w^{(K-2)/2}\exp\left(-\frac{w}{2}\right)\exp\left\{-\frac{1}{2}\left(t\sqrt{\frac{w}{K-1}}-\sqrt{K}\theta_0\right)^2\right\} dw dt = \alpha 2^{K/2}\Gamma\left(\frac{K-1}{2}\right)\sqrt{\pi(K-1)}. \tag{28c}$$

Computation of $\hat{K}$ by means of the last equation appears formidable. However, it can be shown that $g_\theta(t)$ is the probability density of Student's noncentral $t$ distribution with $K-1$ degrees of freedom. A good approximation to $\int_{-\infty}^{\hat{K}} g_\theta(t)$ is given by[9]

$$\int_{-\infty}^{\hat{K}} g_\theta(t)\, dt \doteq \text{prob}\left\{X < \left[\hat{K}\left(1-\frac{1}{4(K-1)}\right)-\sqrt{K}\theta \Big/ \left(1+\frac{\hat{K}^2}{2(K-1)}\right)^{1/2}\right]\right\}, \tag{29}$$

where $X$ is the standard normal random variable.

Applying this approximation to (28c) gives

$$\alpha = \text{prob}\,\{X < \{\hat{K}[1-1/4(K-1)]-\delta\}/[1+\hat{K}^2/2(K-1)]^{1/2}\}, \tag{30}$$

where $\delta = -\sqrt{K}\Phi^{-1}(1-p_0)$. Thus, $\hat{K}$ can be approximately found by solving the quadratic equation.

$$\hat{K}^2\{[1-1/4(K-1)]^2-[\Phi^{-1}(\alpha)]^2/2(K-1)\} - 2\delta[1-1/4(K-1)]\hat{K}+\delta^2-[\Phi^{-1}(\alpha)]^2=0. \tag{31}$$

If $\alpha < \frac{1}{2}$, then the smaller of the two roots of this equation is approximately equal to $\hat{K}$.

To determine $\hat{K}$ more accurately, an iterative procedure based on the approximation given above has been developed by JOHNSON AND WELSH.[10] This procedure is fast, accurate, and computationally straightforward. For further information, the reader is referred to Johnson and Welsh's original paper.

Thus, we have a procedure to test on the basis of observations of branch length whether or not prob $\{M \geqq l\} \geqq p_0$. The procedure is optimum in the sense that the probability of rejecting $H_1$ when it is true is no greater than $\alpha$ and the probability of accepting $H_1$ when it is false is minimum among all tests that depend on the sufficient statistic $T$. If we want stronger control over the probability of accepting $H_1$ when it is false, we should test $H_1'\!: p \leqq p_0$ against $H_2'\!: p > p_0$. This is equivalent to a test discussed in detail by Lehmann.[5] We further note that to perform

either this test or the one described in the section on nonparametric statistical analysis, we do not have to know the path matrix $P_{s,t}$ of $G$. Given the observed flow vector $\boldsymbol{W}(k)=\boldsymbol{w}(k)$, we can form the graph $G_k$ that has the same structure as $G$ but has branch lengths $w_1(k), w_2(k), \cdots, w_m(k)$. Then, $m_k$ is the length of the minimum $s$-$t$ path in $G_k$ that can be found by standard shortest path algorithms.

## PAIRWISE PATH COMPARISON

IN THE preceding sections, we considered means of computing, estimating, or testing hypotheses concerning prob$\{M<l\}$. These methods may be used to compare subsets of paths with one another. A special case of this comparison is the comparison of a single path to one other path. This case could occur in the situation when it is necessary to choose between two given routes in a graph.

Let $\Pi_1$ and $\Pi_2$ be *disjoint* $s$-$t$ paths. As a first approach, we might say that $\Pi_1$ is 'better than' $\Pi_2$ if $E\{|\Pi_1|\}<E\{|\Pi_2|\}$. Since $E\{|\Pi_1|\}=\sum_{b_j\epsilon\Pi_i} E\{W_j\}$, this criterion is theoretically trivial to test. However, criteria based on expected values are often inadequate. For example, we may be concerned with guaranteeing that extremely 'long' paths are avoided as often as possible. Then, given a positive number $l_0$, we could say the $\Pi_1$ is '$l_0$ better than $\Pi_2$,' written $\Pi_1 l_0 > \Pi_2$ if

$$\text{prob}\{|\Pi_1|\geqq l_0\}<\text{prob}\{|\Pi_2|\geqq l_0\}. \tag{32}$$

This criterion is considerably more difficult to test than the preceding one, since to evaluate prob$\{|\Pi_1|\geqq l_0\}$ we must perform a convolution of random variables. Such a convolution is formidable if $\Pi_1$ and $\Pi_2$ contain many branches. If $\Pi_1$ and $\Pi_2$ contain only a few branches, prob$\{|\Pi_i|\geqq l_0\}$ can be found without great numerical difficulty. However, if $|\Pi_i|$ contains a large number of branches it is more reasonable to accept an approximate answer based on a limit distribution. Naturally, a normal approximation can be used, since under very general conditions sums of random variables are approximately normally distributed.[7]

Let $E\{|\Pi_i|\}=\mu_i$, var $\{|\Pi_i|\}=\sigma_i^2$ and assume that prob$\{|\Pi_i|\geqq l_0\}\doteq 1-\Phi([l_0-\mu_i]/\sigma_i)$ for $i=1, 2$. Then,

$$\Pi_1 l_0>\Pi_2 \text{ iff } \Phi([l_0-\mu_1]/\sigma_1)>\Phi([l_0-\mu_2]/\sigma_2). \tag{33}$$

Moreover, since $\Phi(\cdot)$ is a strictly increasing function,

$$\Pi_1 l_0>\Pi_2 \text{ iff } (l_0-\mu_1)/\sigma_1>(l_0-\mu_2)/\sigma_2, \tag{34}$$

or, in other words if $\sigma_2>\sigma_1$,

$$l_0>(\sigma_2\mu_1-\sigma_1\mu_2)/(\sigma_2-\sigma_1). \tag{35}$$

Hence, if we accept the normal approximation as valid, if $\sigma_2 \neq \sigma_1$, there exists a finite constant $l^*$ such that $\Pi_1 l > \Pi_2$ for $l > l^*$ and $\Pi_2 l > \Pi_1$ for $l < l^*$. Then, if $l^*$ is either very large or very small it is easy to decide between $\Pi_1$ and $\Pi_2$. However, if prob $\{|\Pi_i| \geqq l^*\}$ is not negligible, this decision is not straightforward.

The above discussion was based on the premise that $\Pi_1$ and $\Pi_2$ are disjoint. However, we did not take advantage of this fact in the mathematical development and so the previous discussion holds even if $\Pi_1$ and $\Pi_2$ contain common branches. An obvious method of comparing nondisjoint paths is to compare the noncommon branches. Let $\Pi_1 = \hat{\Pi}_1 \cup \Pi_0$ and $\Pi_2 = \hat{\Pi}_2 \cup \Pi_0$ where $\hat{\Pi}_1 \cap \hat{\Pi}_2$ is empty. Then, since

$$|\Pi_i| = \sum_{b_j \epsilon \hat{\Pi}_i} W_j + \sum_{b_j \epsilon \hat{\Pi}_0} W_j = |\hat{\Pi}_i| + |\hat{\Pi}_0|, \tag{36}$$

it is natural to conclude that $\Pi_1$ is better than $\Pi_2$ if $\hat{\Pi}_1$ is better than $\hat{\Pi}_2$. ($\hat{\Pi}_1$, $\hat{\Pi}_2$, and $\hat{\Pi}_0$ are not necessarily paths, but are in general disjoint unions of paths.)

Let $E\{|\hat{\Pi}_i|\} = \hat{\mu}_i$, $\text{var}\{|\hat{\Pi}_i|\} = \hat{\sigma}_i^2$ and assume that $\text{prob}\{|\hat{\Pi}_i| \geqq l\} \doteq 1 - \Phi([l - \hat{\mu}_i]/\hat{\sigma}_i)$ for $i = 0, 1, 2$. From the previous discussion, $\hat{\Pi}_1 l > \hat{\Pi}_2$ for all $l > \hat{l}^*$ where

$$\hat{l}^* = (\hat{\sigma}_2 \hat{\mu}_1 - \hat{\sigma}_1 \hat{\mu}_2)/(\hat{\sigma}_2 - \hat{\sigma}_1), \qquad \text{for } \hat{\sigma}_2 > \hat{\sigma}_1. \tag{37}$$

From this, it is tempting to conjecture that since $|\Pi_1| = |\hat{\Pi}_1| + |\hat{\Pi}_0|$ and $|\Pi_2| = |\hat{\Pi}_2| + |\hat{\Pi}_0|$, then $\Pi_1 l > \Pi_2$ for all $l > \hat{l}^*$. However, this conjecture is false. To see this, first note that $\mu_1 = \hat{\mu}_1 + \hat{\mu}_0$, $\mu_2 = \hat{\mu}_2 + \hat{\mu}_0$, $\sigma_1 = \sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_0^2}$ and $\sigma_2 = \sqrt{\hat{\sigma}_2^2 + \hat{\sigma}_0^2}$. From (43), $\Pi_1 l > \Pi_2$ iff $l > l^*$, where

$$l^* = [\sqrt{\hat{\sigma}_2^2 + \hat{\sigma}_0^2}(\hat{\mu}_1 + \hat{\mu}_0) - \sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_0^2}(\hat{\mu}_2 + \hat{\mu}_0)] / [\sqrt{\hat{\sigma}_2^2 + \hat{\sigma}_0^2} - \sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_0^2}], \tag{38}$$

and obviously there exist numerous combinations of $\hat{\sigma}_1, \hat{\sigma}_2, \hat{\mu}_1, \hat{\mu}_2$ and $\hat{\sigma}_0$ and $\hat{\mu}_0$ such that $\hat{l}^* \neq l^*$. It is also interesting to note that if $\hat{\sigma}_0 \to \infty$, $\Pi_1 l > \Pi_2$ iff $\mu_1 < \mu_2$ for all $l$, but $\hat{\Pi}_1 l > \hat{\Pi}_2$ only for $l > \hat{l}^*$. Hence, in comparing paths it is only reasonable to compare their disjoint subpaths.

## NOTES

1. The treatment of directed graphs is obtained trivially by replacing the word path by the words directed path throughout the paper.

2. $(\bar{M}, S)$ is a sufficient statistic for $(\xi, \sigma)$ if the conditional distribution of any other statistic $X$, given that $(\bar{M}, S) = (\bar{m}, s)$, is independent of $(\xi, \sigma)$ and hence does not give any information about $(\xi, \sigma)$.

3. If $t$ is a particular value of $T$ and $g$ is a transformation on the parameter space of $(\xi, \sigma)$, then $T$ is *invariant* under $g$ if for every value of $(\xi, \sigma)$, $T_{(\xi,\sigma)}(t) = T_{g(\xi,\sigma)}(t)$. If $\mathcal{G}$ is a group of transformations such that $T$ is invariant for any $g \epsilon \mathcal{G}$, $T$ is a *maximal* invariant if $T_{\theta_1}(t) = T_{\theta_2}(t)$ for all $t$ implies $\theta_1 = g(\theta_2)$ for some $g \epsilon \mathcal{G}$.

**ACKNOWLEDGMENTS**

THIS RESEARCH was sponsored by the Office of Emergency Preparedness, Executive Office of the President, and the Joint Services Electronics Program. The author is on leave from the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley.

**REFERENCES**

1. R. G. BUSACKER AND T. L. SAATY, *Finite Graphs and Networks*, McGraw-Hill, New York, 1965.
2. C. BERGE AND A. GHOUILA-HOURI, *Programming, Games and Transportation Networks*, Wiley, New York, 1965.
3. H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N.J., 1946.
4. H. FRANK AND S. L. HAKIMI, "Probabilistic Flows Through a Communication Network," *IEEE Trans. on Circuit Theory*, **CT-12,** 413–414 (1965).
5. E. L. LEHMANN, *Testing Statistical Hypotheses*, Wiley, New York, 1959.
6. H. FRANK AND S. L. HAKIMI, "On the Optimum Synthesis of Statistical Communication Nets," *J. Franklin Inst.* **284,** 407–416 (1967).
7. M. FISZ, *Probability Theory and Mathematical Statistics*, Wiley, New York, 1963.
8. R. M. VAN SLYKE, "Monte Carlo Methods and the PERT Problem," *Opns. Res.* **11,** 839–860 (1963).
9. M. ABRAMOWITZ AND L. A. STEGUN (Eds.), *Handbook of Mathematical Functions*, National Bureau of Standards AMS 55, p. 949 (1964).
10. N. L. JOHNSON AND B. L. WELSH, "Applications of the Noncentral $t$ Distribution," *Biometrika* **40,** 362–389 (1939).