THE ASTROPHYSICAL JOURNAL, 399:345–352, 1992 November 10

A SIMPLE TEST OF INDEPENDENCE FOR TRUNCATED DATA WITH APPLICATIONS TO REDSHIFT SURVEYS

BRADLEY EFRON
Department of Statistics, Stanford University, Stanford, CA 94305
AND
VAHÉ PETROSIAN[1]
Center for Space Science and Astrophysics, Stanford University, Stanford, CA 94305
Received 1992 February 26; accepted 1992 May 14

ABSTRACT

In analysis of astronomical data, one is often faced with determination of bivariate distributions from truncated data. This leads to the following statistical question: Is a truncated sample of observed points (x_i, y_i) consistent with the hypothesis H_0 that x and y are statistically independent? This paper presents an easily applied permutation test for H_0, closely related to Lynden-Bell's estimate of the marginal distribution of truncated data. The test is applied to two redshift-magnitude surveys, one of galaxies and one of quasars.

Analysis of the galaxy survey by Loh & Spillar shows that in the framework of a simple Hubble Law model, that is, distance proportional to redshift, or most conventional models with zero cosmological constant and density parameters $\Omega \sim O(1)$, the absolute magnitude or luminosity and redshift are statistically independent. Therefore, assuming statistical independence, testing H_0 amounts to testing validity of the cosmological model. Segal's chronomatic cosmological model is rejected under H_0. On the other hand, for the quasar sample H_0 is rejected strongly in a conventional cosmological model (and in a chronomatic model as well) indicating either incorrectness of the models or, as is more commonly assumed, indicating strong luminosity evolution.

Subject headings: cosmology: observations — galaxies: distances and redshifts — methods: statistical

1. INTRODUCTION

An important question that arises in investigations of astronomical phenomena is the variation with distance and angles in the sky of some physically interesting intrinsic parameters. Most such parameters exhibit dispersion, forcing this question to be answered statistically. Moreover, the parameters often are not directly measurable but can only be obtained from other measured variables which often include the distance to the sources. This paper concerns testing truncated bivariate data for statistical independence. We shall apply these tests to problems arising in analysis of data obtained from redshift surveys.

One of the goals in analysis of redshift survey data is determination of the distribution of the luminosity of the extragalactic sources and how it varies with distance or cosmological epoch. For this one needs the absolute luminosity L of a source, which is related to the observed apparent luminosity l as $L = 4\pi l\, d^2(z, \Omega)$, where the luminosity distance d depends on the redshift z and cosmological model (represented here by Ω). We ignore here K-correction, spatial absorption, or difficulties in dealing with extended sources. Therefore, the distribution of luminosity L or the so-called luminosity function, $\psi(L)$, is intimately connected to the spatial distribution of sources so that, in general, one is dealing with determination of joint distribution of L and distance, $\Psi(L, d)$ or $\Psi(L, z)$.

The main difficulty here is lack of a priori knowledge about the cosmological model [i.e., form of $d(z, \Omega)$]. As is commonly known, but often ignored, it is impossible to determine both cosmology of the universe and properties of the function Ψ from a redshift survey. One must assume a cosmological model to determine Ψ, or conversely, assume a form for Ψ to obtain values of the cosmological parameters (see below).

A redshift survey provides pairs of measurements (z_i, m_i) on redshift z_i and magnitude $m = -2.5 \log_{10} l + C_1$ of a sample of extragalactic sources with various observational biases. We shall ignore all but the most common bias, which is that introduced by the limiting magnitude m_0 of the survey. Such surveys are called magnitude-limited surveys. We will write the data for n objects as

$$(z_i, m_i) \quad \text{for} \quad i = 1, 2, \ldots, n \quad \text{with} \quad m_i < m_0 \,. \tag{1.1}$$

The absolute magnitude $M = -2.5 \log L + C_2$ can be computed if we assume a specific cosmological model $\Omega = \Omega_0$:

$$M_i(\Omega_0) = m_i - 5 \log d(z_i, \Omega_0) + C \,. \tag{1.2}$$

The data then can be reexpressed as (z_i, M_i) for $i = 1, 2, \ldots, n$, satisfying the truncation relationship

$$M_i(\Omega_0) \le m_0 - 5 \log d(z_i, \Omega_0) + C \,. \tag{1.3}$$

The constant C, which depends on definitional details, will be set equal to zero in what follows. This does not affect any of our calculations.

Figure 1 shows $n = 492$ galaxies from Loh & Spillar's (1988) survey. The plotted points are $(x_i, y_i) = (\log z_i, M_i)$, where M_i is the absolute magnitude calculated from equation (1.2) assuming luminosity distance $d(z_i, \Omega_0) \propto z_i$, so that $M_i = m_i - 5 \log_{10} (z_i)$. This is what is expected from the simple Hubble law. The magnitude limit $m_0 = 21.5$ of the survey leads to the diagonal truncation boundary

$$M_i \le 21.5 - 5 \log z_i \tag{1.4}$$

apparent in Figure 1.

[1] Also Departments of Physics and Applied Physics.

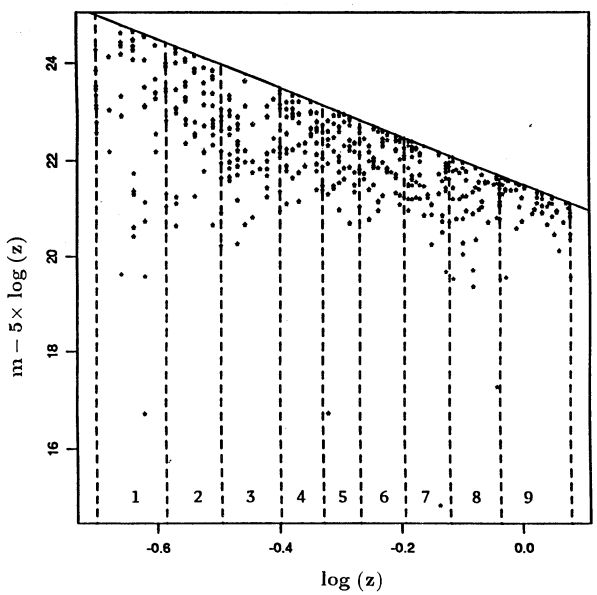

FIG. 1.—Distribution (x_i, y_i) from Loh & Spillar's (1988) redshift survey of 492 galaxies; $x_i = \log(z_i)$ and $y_i = m_i - 5 \log(z_i)$, for z_i the redshift and m_i the apparent magnitude of the ith galaxy. This assumes that the simple Hubble law is valid out to the maximum redshift of the survey. The 492 points are those satisfying $m_i < 21.5$ and $z_i \geq 0.2$. The nine vertical strips each contain 55 points, except for the rightmost which contains 52. Diagonal line indicates truncation boundary $21.5 - 5x$.

This kind of truncation causes trouble for standard statistical methods. A survey limited by absolute magnitude, rather than apparent magnitude, would give a horizontal truncation boundary, which causes no statistical difficulties. A similar statement applies to volume-limited surveys, which include all sources up to a limiting distance or redshift. The difficulty caused by the truncation due to the magnitude limit can be overcome if the luminosity distribution is independent of redshift which means that z and L (or M) behave as two independent variables, so that their joint density $\Psi(z, L)$ can be written as

$$\Psi(z, L) = \rho(z)\psi(L) . \tag{1.5}$$

This behavior is sometimes referred to as pure density evolution. In this case, as described by Petrosian (1992), there exists a unique nonparametric technique for estimating $\rho(z)$ and $\psi(L)$, based on Lynden-Bell's (1971) c^- method. The key assumption in the c^- method is independence. In this paper we provide a nonparametric test of the independence hypothesis, applied to truncated data. The test is closely related to Lynden-Bell's methods, and also to the theory in Bhattacharya, Chernoff, & Yang (1983).

Section 2 describes a class of nonparametric permutation tests which are both powerful and easy to apply. Section 3 describes rank-based statistics for testing independence. Section 4 describes two applications of the method. Section 4.1 shows that the Loh & Spillar (1988) data set of Figure 1 successfully passes our permutation tests for independence in the simplified cosmological model obeying Hubble's law. Conversely, assuming independence, it is possible to use the permutation tests to obtain confidence intervals for cosmological model parameter Ω. Each choice of Ω gives a data set $\{[z_i, M_i(\Omega)], i = 1, \ldots, n\}$, which can be tested for independence. The set of Ω-values that pass the test form the confidence interval for the cosmological model. We show, for example, that the chronomatic model (e.g., Segal & Nicoll 1986) falls outside this confidence interval.

Section 4.2 shows that the permutation tests strongly reject the independence of z_i and M_i for a subsample of quasar survey from Boyle et al. (1990). This could indicate either a failure of the assumed cosmological model, or it can be ascribed to evolutionary trends in the quasar luminosity function (see Boyle et al. 1987; Caditz & Petrosian 1990). Such trends can be searched for with the tests discussed here. If the evolution can be described by the so-called pure luminosity evolution so that the shape of the luminosity function is invariant but all luminosities vary by some factor $\Psi(L) \to \Psi[L/L_*(z)]/L_*(z)$, then the independence test can be used to determine the rate of evolution $L_*(z)$. For example, assuming a cosmological model, Ω_0, and an evolutionary form, $L_* \propto (1 + z)^\gamma$, one can then apply the permutation test described below to a family of evolutionary models

$$M(\Omega_0, \gamma) = m - 5 \log d(z, \Omega_0) + 2.5\gamma \log(1 + z) . \tag{1.6}$$

Each choice of γ gives a data set $\{[z_i, M_i(\gamma)], i = 1, \ldots, n\}$ which can then be tested for independence. The set of γ-values that pass the test form the confidence interval for luminosity evolution in the assumed cosmological model. This is not done here. Our primary purpose here is to describe the test statistics and not to draw astronomical inference from the above-mentioned data. The latter will be dealt with in more detail in subsequent works.

The test described here can be applied not only to redshift surveys but also to any other situation where a bivariate (or a more generally multivariate) distribution is required. A prime example in astrophysics is the spatial and luminosity distribution of other kinds of sources and joint distributions of luminosities at various wavelengths.

2. PERMUTATION TESTS FOR INDEPENDENCE

2.1. *Nontruncated Data*

We first describe permutation tests for independence when the data is not truncated. Suppose x and y are the two random variables to be tested for independence. The data consist of a random sample of n pairs from the joint distribution of (x, y),

$$\text{data} = \{(x_i, y_i) , \quad i = 1, 2, \ldots, n\} . \tag{2.1}$$

This data is to be used to test the null hypothesis of independence

$$H_0 : x \quad \text{and} \quad y \quad \text{independent} . \tag{2.2}$$

Denote the n ordered x-values by $\boldsymbol{x} = (x_{(n)}, x_{(n-1)}, \ldots, x_{(1)})$, where

$$x_{(n)} < x_{(n-1)} \cdots < x_{(1)} , \tag{2.3}$$

and likewise $\boldsymbol{y} = [y_{(n)}, y_{(n-1)}, \ldots, y_{(1)}]$,

$$y_{(n)} < y_{(n-1)} \cdots < y_{(1)} . \tag{2.4}$$

The convenient assumption of no ties among the x- or y-values has no real effect on the permutation theory, except for a minor point that will be noted later. The data (2.1) can be described without loss of information by the observed pairings between the $x_{(i)}$ and $y_{(j)}$, say

$$\text{data} = \{(x_{(i)}, y_{(j_i)}) , \quad i = 1, 2, \ldots, n\} . \tag{2.5}$$

There are $n!$ possible ways to choose the pairing vector $\boldsymbol{j} =$

$(j_1, j_2, \ldots, j_n)$. Under the null hypothesis of independence it is easy to show that all $n!$ ways are equally likely (see eqs. [2.14] and [2.15] below).

Let t (data) be a test statistic, for instance, the sample correlation between the x_i and y_i or the test statistics described in § 3. A permutation test of H_0 compares the observed value of the statistic t (data) with the *permutation distribution* of t, that is, with the $n!$ values of t obtained by all possible pairings of $\boldsymbol{x}$ and $\boldsymbol{y}$. It is conventional to reject H_0 if t (data) is extremely large or small compared to the permutation distribution, say in the upper or lower 5% tails (see chap. 6 of Hajek 1969).

2.2. *Truncated Data*

We now describe a modification of this theory to deal with truncated data. We assume that pairs (x, y) are observable only if they satisfy the truncation relationship

$$y \le u(x) . \tag{2.6}$$

Here $u(x)$ is a monotonic function of x, and for definiteness we assume $u(x)$ decreasing in x. In Figure 1, $u(x) = 21.5 - 5x$.

For each i, the upper limit

$$u_{(i)} = u(x_{(i)}) \tag{2.7}$$

defines the set of values $y_{(j)} \le u_{(i)}$ in $\boldsymbol{y}$ that could possibly be paired with $x_{(i)}$. Let n_i indicate the number of such values,

$$n_i = \text{Number } \{j : y_{(j)} \le u_{(i)}\} , \quad i = 1, 2, \ldots, n . \tag{2.8}$$

Imagine that we have observed $(x_{(1)}, y_{(j_1)})$, $(x_{(2)}, y_{(j_2)})$, $\ldots$, $(x_{(i-1)}, y_{(j_{i-1})})$. Define J_i, the *eligible* set, as the index j contributing to n_i minus the indices j_i from the first $i - 1$ pairings in the sample:

$$J_i = \{j : y_{(j)} \le u_{(i)}\} - \{j_1, j_2, \ldots, j_{i-1}\} . \tag{2.9}$$

Figure 2 illustrates these definitions for a hypothetical data set of $n = 7$ pairs. In this example $J_4 = \{2, 3, 6\}$ because $y_{(1)} > u_{(4)}$ and $y_{(5)}$, $y_{(7)}$, $y_{(4)}$ are already paired with $x_{(1)}$, $x_{(2)}$, $x_{(3)}$, respectively. The number of points in J_i is defined to be N_i,

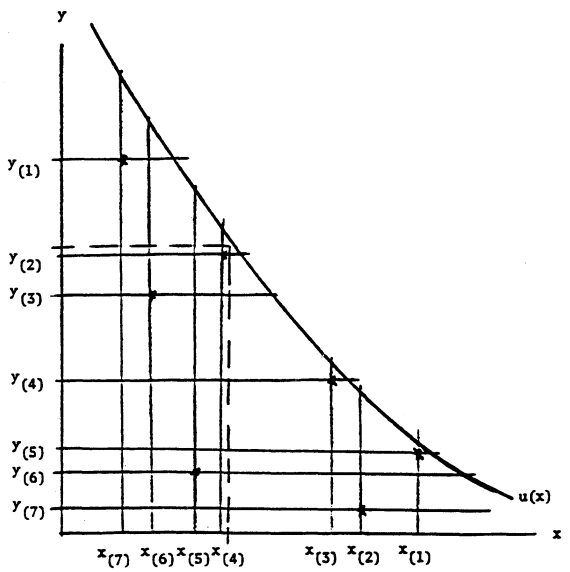

FIG. 2.—Hypothetical truncated sample of size 7; stars indicate the seven observed points; $u(x)$ indicates the truncation boundary, above which points are not observed. Boundary of the elligible set for $J_4 = \{2, 3, 6\}$ is shown by the heavy solid lines. It is then clear that for $i = 1, 2, \ldots, 7$ the number of points in the eligible sets are $N_i = 3, 2, 2, 3, 2, 2, 1$, respectively. The product of the N_i, $N = 144$ is the number of ways the $y_{(j)}$ can be paired with the $x_{(i)}$ to give an observable set of seven points.

$$N_i = \text{Number } J_i = n_i - i + 1 . \tag{2.10}$$

For each value i, the rectangular region $x \le x_{(i)}$ and $y \le u_{(i)}$ contains the points indexed by J_i, the *comparable set* in terminology of Bhattacharya et al. (1983). In the context of Figure 1, it could be thought of as a subsample of the full survey, in which the observational limits were $z \le z_{(i)}$ and $M \le 21.5 - 5 \log (z_{(i)})$. This is an untruncated subsample, untruncated really meaning truncated parallel to the coordinate axis. It is the largest possible absolute (rather than apparent) magnitude-limited subsample associated with the ith point. The number $N_i - 1$ is same as the number C^- of Lynden-Bell (1971) and the numbers N and M defined in Petrosian (1968).

The following two facts determine the permutation theory for truncated data.

Fact 1. Given $\boldsymbol{x}$ and $\boldsymbol{y}$, there are

$$N = \prod_{i=1}^{n} N_i \tag{2.11}$$

pairings $\{(x_{(i)}, y_{(j_i)}), i = 1, 2, \ldots, n\}$ satisfying the truncation relationship

$$y_{(j_i)} \le u_{(i)} \quad \text{for} \quad i = 1, 2, \ldots, n . \tag{2.12}$$

We call a pairing vector $\boldsymbol{j} = (j_1, j_2, \ldots, j_n)$ *observable* if relationship (2.12) is satisfied.

Fact 2. Each observable pairing vector has equal conditional probability given $\boldsymbol{x}$ and $\boldsymbol{y}$, assuming that the null hypothesis of independence is true,

$$\text{Prob}_{H_0} \{\boldsymbol{j} | \boldsymbol{x}, \boldsymbol{y}\} = 1/N \quad \text{for} \quad \boldsymbol{j} \text{ observable} . \tag{2.13}$$

The proof of equation (2.11) is immediate from definitions (2.9)–(2.10). There are N_1 observable ways to select $y_{(j_1)}$, N_2 observable ways to select $y_{(j_2)}$ having already selected $y_{(j_1)}$, etc. Notice that the numbers n_i and N_i, and hence N, are determined entirely by $\boldsymbol{x}$ and $\boldsymbol{y}$, even though the eligible sets J_i are not. We could compute $N = 144$ in Figure 1 even if the stars indicating the observed data points were erased.

Result (2.13) is almost equally immediate. Suppose the independence hypothesis H_0 is true, so that, ignoring truncation, the joint density of (x, y) can be represented as

$$h(x, y) = f(x)g(y) . \tag{2.14}$$

The unconditional probability of observing data $\{(x_i, y_i), i = 1, 2, \ldots, n\}$ is proportional to

$$\prod_{i=1}^{n} [f(x_{(i)})g(y_{(j_i)})] = \prod_{i=1}^{n} f(x_{(i)}) \prod_{j=1}^{n} g(y_{(j)}) , \tag{2.15}$$

where the proportionality constant depends on the differential elements dx_i, dy_j and on the sample selection criteria but not on the pairings $j_1, j_2, \ldots, j_n$. Therefore the probability in equation (2.15) is the same for all observable pairing vectors $\boldsymbol{j}$ so that the conditional probabilities $\text{Prob}_{H_0} \{\boldsymbol{j} | \boldsymbol{x}, \boldsymbol{y}\}$ must all be equal.

It is now easy to describe a permutation test of independence for truncated data. First of all choose a test statistic t(data), where data $= \{(x_{(i)}, y_{(j_i)}), i = 1, 2, \ldots, n\}$ as in equation (2.5). Compute all N observable pairing vectors corresponding to $\boldsymbol{x}$ and $\boldsymbol{y}$. Each such vector, say $\boldsymbol{j}^*$, gives a permuted data set

$$\text{data}^* = \{(x_{(i)}, y_{(j_i^*)}) , \quad i = 1, 2, \ldots, n\} . \tag{2.16}$$

Compute t(data*) for each j^*. Finally, compare the observed value t(data) with the permutation distribution, that is with the N permutation values t(data*). Reject the null hypothesis of independence if t(data) is in the extreme tails of the permutation distribution.

Exactness is the great virtue of permutation tests. A level 0.05 permutation test rejects H_0 exactly 5% of the time when H_0 is true. The next section discusses the choice of the test statistic t(data). A good choice leads to a powerful test, one that has high probability of rejecting H_0 when it is false. It is also helpful to choose t so that easy approximations are available for the permutation distribution, making it unnecessary to actually compute the N values of t(data*).

3. TEST STATISTICS FOR INDEPENDENCE

This section discusses rank-based statistics for testing independence, the goal being powerful, robust, and easily administered tests. We consider truncated data, but the methods as written apply to untruncated cases as well, the difference being only in the way we define subset J_i for each data point.

3.1. *Normalized Rank Statistics*

Corresponding to the eligible set of indices J_i, (eq. [2.9]), there is

$$Y_i = \{y_{(j)} : j \in J_i\} , \tag{3.1}$$

the set of $y_{(j)}$ values that can possibly be paired with $x_{(i)}$ in an observable way after the pairings $(x_{(1)}, y_{(j_1)}) \ldots (x_{(i-1)}, y_{(j_{i-1})})$ have already been made. Define

$$R_i = \text{rank of } y_{(j_i)} \text{ in } Y_i , \tag{3.2}$$

so R_i ranges from 1 if $y_{(j_i)}$ is the smallest member of Y_i to N_i if $y_{(j_i)}$ is the largest.

It follows from equation (2.13) that if the hypothesis of independence is true, then R_i equals any of its possible values with equal probability,

$$\text{Prob}_{H_0}\{R_i = r \,|\, \boldsymbol{x}, \boldsymbol{y}\} = \frac{1}{N_i} \quad \text{for} \quad r = 1, 2, \ldots, N_i . \tag{3.3}$$

In other words, $y_{(j_i)}$ is chosen completely at random from Y_i if H_0 is true. Moreover the R_i are independent of each other under H_0, given $\boldsymbol{x}$ and $\boldsymbol{y}$. Using $U(1:N)$ to represent a discrete uniform distribution on the integers 1, 2, ..., N, we can write all of this as

$$R_i \,|\, \boldsymbol{x}, \boldsymbol{y} \sim U(1:N_i) \quad \text{independently for} \quad i = 1, 2, \ldots, n \tag{3.4}$$

under H_0.

The discrete distribution $U(1:N)$ has expectation $(N+1)/2$ and variance $(N^2-1)/12$. The *normalized rank statistic*

$$T_i = (R_i - E_i)/V_i^{1/2} , \tag{3.5}$$

where

$$E_i = (N_i + 1)/2 \quad \text{and} \quad V_i = (N_i^2 - 1)/12 , \tag{3.6}$$

has expectation 0 and variance 1 under H_0. Because of equation (3.4) the vector $\boldsymbol{T} = (T_1, T_2, \ldots, T_n)$ has expectation vector $\boldsymbol{0}$ and covariance matrix the $n \times n$ identity $\boldsymbol{I}_n$ under H_0. We will use the vector $\boldsymbol{T}$ to construct test statistics for independence. Note that computer ranking programs usually assign the average rank in case of ties. This does not change the mean E_i in equation (3.6), but decreases the variance V_i. The effect is negligible unless there is a very large number of ties.

3.2. *Normal Approximation*

It is convenient to base test statistics for independence on weighted linear combinations of the normalized ranks T_i. Let $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ be any vector of weights. The test statistic

$$t_{\boldsymbol{w}}(\text{data}) = \sum_{i=1}^{n} w_i T_i \Big/ \left(\sum_{i=1}^{n} w_i^2 \right)^{1/2} \tag{3.7}$$

has mean 0 and variance 1 under H_0. In practice it will not be necessary to evaluate the values or distribution of $t_{\boldsymbol{w}}$ for all possible permutations. The normal approximation

$$t_{\boldsymbol{w}}(\text{data}) \dot{\sim} N(0, 1) \quad (\text{under } H_0) \tag{3.8}$$

will be quite accurate in most circumstances because the T_i have short-tailed distributions symmetrical about 0. Even for n as small as 10, equation (3.8) approximates the relevant parts of $t_{\boldsymbol{w}}$(data) with an error of less than 1% in the untruncated case. Higher-order Edgeworth expansions for the distribution of $t_{\boldsymbol{w}}$(data) are easily derived, but the corrections to equation (3.8) tend to be small. One argument for using rank-based tests is the increased reliability of normal approximations. Another is improved invariance properties: $t_{\boldsymbol{w}}$(data) retains the same value if we make monotonic increasing transformations on the x- and y-axes.

The normal approximation (3.8) is better expressed as $t_{\boldsymbol{w}}(\text{data}) \,|\, \boldsymbol{x}, \boldsymbol{y} \dot{\sim} N(0, 1)$ under H_0, to emphasize the conditioning on $\boldsymbol{x}$ and $\boldsymbol{y}$. Remembering that the conditional distribution of the data given $\boldsymbol{x}$ and $\boldsymbol{y}$, under H_0, is the same as the permutation distribution (2.13), we can carry out the permutation test of independence using the approximation

$$t_{\boldsymbol{w}}(\text{data*}) \sim N(0, 1) . \tag{3.9}$$

Following convention, we can accept H_0 if $|t\ (\text{data})| \leq 1.645$, and reject H_0 otherwise. The rejection probability of the permutation test would be approximately 0.10 according to equation (3.9).

A simple and reasonable choice for the weight vector $\boldsymbol{w}$ is $\boldsymbol{w} = (1, 1, \ldots, 1)$, which assigns equal weights to all of the T_i. Then $t_{\boldsymbol{w}}(\text{data}) = n^{1/2}\bar{T}$ where $\bar{T} = \sum_{i=1}^{n} T_i/n$. We will call this statistic t_1(data). It is essentially the same as the statistic β^{**} suggested in Bhattacharya et al. (1983), the primary difference being the use here of permutation tests.

3.3. *Tests with Good Power*

Permutation theory allows us to keep the rejection probability of our test small when H_0 is true. However, we also want a test that has good power, that is, a big probability of rejection, under a likely alternative hypothesis H_1 to H_0. This will be true for the test based on t_1(data), which simply adds up the T_i if the H_1 expectations of the statistic T_i, say $E_{H_1}\{T_i\}$, all have the same sign. As mentioned in § 1, the alternative hypothesis could be independence of y and x in another cosmological model and/or for some kind of luminosity evolution.

Suppose we believe that a particular choice of cosmological model, say $\Omega = \Omega_0$, makes y independent of x. Then we can calculate ranks $R_i(\Omega_0)$ (eq. [3.2]) and normalized ranks $T_i(\Omega_0)$ (eq. [3.5]) based on the data set

$$(x_i, y_i)\ i = 1, 2, \ldots, n; \quad x_i = \log z_i \quad \text{and} \quad y_i = M_i(\Omega_0) , \tag{3.10}$$

where $M_i(\Omega_0)$ is defined in equation (1.2). If in fact $\Omega = \Omega_0$, then the $T_i(\Omega_0)$ all have expectation zero. However if the correct value of Ω is actually Ω_1 so that it is data (Ω_1) that satisfies H_0, then often the $E_{\Omega_1}\{T_i(\Omega_0)\}$ will all tend to be either

positive or negative. Both cases are favorable to the use of the test statistic t_1(data), although more powerful permutation tests may be available, as discussed next.

The locally most powerful rank statistic (3.7) for testing that $\Omega = \Omega_0$ is the correct value for independence in equation (3.10) uses weights

$$w_i = \left.\frac{\partial E_\Omega\{T_i(\Omega_0)\}}{\partial\Omega}\right|_{\Omega=\Omega_0}, \tag{3.11}$$

where $E_{\Omega_1}\{T_i(\Omega_0)\}$ indicates the expectation of $T_i(\Omega_0)$ when Ω_1 is correct (see Hajek 1969). These weights are optimal for testing nearby alternatives to $\Omega = \Omega_0$. Standard probability calculations give the optimum weights (3.11), the answer depending on the marginal densities $f(x)$ and $g(y)$ in equation (2.14).

Suppose that $f(x)$ and $g(y)$ have algebraic tails in the observable region, say

$$f(x) = k_1(x - x_{\min})^{p_1} \quad \text{and} \quad g(y) = k_2(y - y_{\min})^{p_2} \tag{3.12}$$

for $x \geq x_{\min}$ and $y \geq y_{\min}$. Here $x_{\min}$ and $y_{\min}$ are lower bounds for x and y, p_1 and p_2 are positive numbers, and k_1 and k_2 are arbitrary normalizing constants. Then it can be shown that the optimal weights (3.11) are given by

$$w_i = \frac{x_i - x_{\min}}{u_i - y_{\min}}. \tag{3.13}$$

Basing our test statistics on ranks helps protect the test against occasional outlying data points, like the four lowest points in Figure 1. However, as discussed below, outliers can still cause trouble. A legitimate and useful tactic is to *trim* the data by simply removing outlying points. Everything said so far applies just as well to the trimmed data sets as to the original data.

A theoretically interesting choice of the weights w_i in equation (3.7), which will not be pursued here, is $w_i = V_i^{1/2}$. Then equation (3.5) gives t_w(data) $= (\Sigma R_i - \Sigma E_i)/(\Sigma V_i)^{1/2}$. In the untruncated case this is equivalent to the nonparametric correlation measure called Kendall's tau statistic (Hajek 1969). It can be shown that this choice of weights gives the same value of t_w(data) if the roles of x and y are interchanged in the definitions, even with truncated data.

Suppose we are considering several test statistics t_1(data), t_2(data), ..., t_p(data), corresponding to different weight vectors $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_p$ in equation (3.7). These can be combined into an omnibus test statistic

$$t_w^2(\text{data}) = \boldsymbol{T}'\boldsymbol{W}(\boldsymbol{W}'\boldsymbol{W})^{-1}\boldsymbol{W}'\boldsymbol{T}, \tag{3.14}$$

where $\boldsymbol{T}' = (T_1, T_2, \ldots, T_n)$ and $\boldsymbol{W}$ is the $n \times p$ matrix with columns $\boldsymbol{w}_j$, assumed to be of full rank. An approximate level α test rejects the independence hypothesis H_0 if t_w^2(data) exceeds $\chi_p^{2(1-\alpha)}$, the $100 \times (1 - \alpha)$th percentile point of a χ^2 distribution with p degrees of freedom. Omnibus statistics test the independence hypothesis H_0 against a broader array of alternatives then do 1 degree of freedom tests like equation (3.7), but they often pay for this by having less power against alternatives of particular interest.

4. APPLICATIONS OF THE TESTS

This section applies the truncated permutation tests to two redshift surveys. The first is the data of Loh & Spillar (1988) on galaxies shown in Figure 1 where there is strong Hubble relation while weak correlation between luminosity and redshift. The second data set on qusars from Boyle et al. (1990) shows essentially no Hubble relation indicating a strong correlation between redshift and luminosity under conventional assumption about the nature of their redshifts.

4.1. *The Loh-Spillar Survey*

We apply the permutation tests of § 3 to the 492 points from the redshift survey of Loh & Spillar (1988) shown in Figure 1 with $m < m_0 = 21.5$ and $z \geq 0.20$. Here we assume the simple Hubble law, with $d(z) \propto z$ so that $M = m - 5 \log z$. In what follows we apply the permutation tests not only to this data but on data obtained for a whole family of ad hoc cosmological models where $d(z) \propto z^{0.2c}$. The data sets are then

$$\text{data } (c) = \{[x_i, y_i(c)],\ i = 1, 2, \ldots, 492\}, \tag{4.1}$$

where

$$x_i = \log z_i \quad \text{and} \quad y_i(c) = m_i - c \log z_i. \tag{4.2}$$

The set data (5), corresponding to the simplest form of the Hubble law, are displayed in Figure 1.

Figure 3 shows three different permutation statistics evaluated for $3 \leq c \leq 10$. The three statistics are t_w[data (c)] for the equally weighted choice $\boldsymbol{w} = (1, 1, \ldots, 1)$,

$$t_1[\text{data } (c)] = \sqrt{492}\,\bar{T}(c) = \sum_{i=1}^{492} T_i(c)/\sqrt{492}; \tag{4.3}$$

the trim-5 version of t_1,

$$t_1[\text{data}_{(5)}\ (c)] = \sqrt{487}\,\bar{T}(c), \tag{4.4}$$

where data$_{(5)}$ (c) is the subset of data (c) having the points with the five smallest values of $y_i(c)$ removed; and finally the "optimal" choice applied to the trimmed data sets,

$$t_{\text{opt}} = t_w[\text{data}_{(5)}\ (c)] \tag{4.5}$$

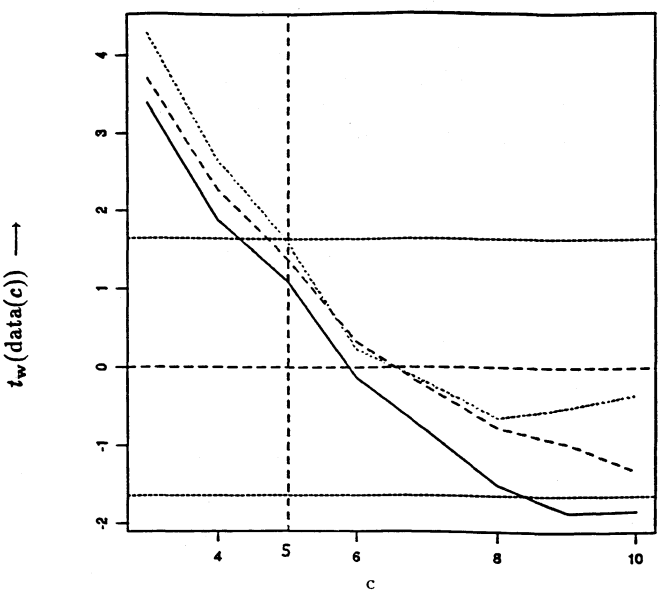

FIG. 3.—Three permutation statistics applied to the Loh-Spillar galactic redshift survey shown in Fig. 1 with $y = m - c \log (z)$. *Solid line*: t_1[data$_{(5)}$ (c)], the equal weight statistic applied after trimming the five points with smallest $y(c)$ values; *dashed line*: t_1(data), the equal weight statistic $(492)^{1/2}\bar{T}$; *dotted line*: t_w[data$_{(5)}$ (c)] with $\boldsymbol{w}$ equal optimum weights (eq. [3.13]); data $(c) = \{[x_i, y_i(c)],\ i = 1, 2, \ldots, 492\}$, $x_i = \log z_i$, and $y_i(c) = m_i - cx_i$. In this case the statistic t_1[data$_{(5)}$ (c)] performs best, giving point estimate $c = 5.89$ as the zero-crossing value, and 90% confidence interval [4.30, 8.82] for c.

where the weights w_i are given by equation (3.14). More explicitly, the weights for t_{opt} equal

$$w_i = \frac{x_i + 0.7}{2 - 5.89x_i}, \tag{4.6}$$

as discussed below.

The trimmed version of the equal weight statistic, $t_1[\text{data}_{(5)}(c)]$, performs best here. It gives point estimate $c = 5.89$, this being the value of c for which $t_1[\text{data}_{(5)}(c)] = 0$. Also

$$t_1[\text{data}_{(5)}(4.30)] = 1.645 \quad \text{and} \quad t_1[\text{data}_{(5)}(8.82)] = -1.645, \tag{4.7}$$

giving the 90% two-sided confidence interval [4.30, 8.82] for c, based on approximation (3.9). The Hubble law value $c = 5$ has $t_1[\text{data}_{(5)}(c)] = 1.09$. On the basis of this test, the Hubble law under the hypothesis H_0 is accepted for the Loh-Spillar data.

The value $c = 2.5$ connected with the chronometric cosmology, $d(z) \propto z^{1/2}$ for small z (Segal & Nicoll 1980) is rejected assuming H_0. The more precise relationship of the chronometric cosmology, $y = m - 2.5 \log [z/(1 + z)]$, yields $t_1(\text{data}_{(5)}) = 5.84$, for the trim-5 version of the t_1 statistics, and is also strongly rejected.

We have also tested more realistic cosmological models. For example, the model with $\Omega = 1$ and with zero cosmological constant has y_i or $M_i = m_i - 5x_i$ where

$$x_i = \log [r_i(1 + z_i)], \quad \text{with} \quad r = 2(1 - 1/\sqrt{1 + z}). \tag{4.8}$$

Using these definitions, we find $t_1(\text{data}_{(5)}) = 0.46$ so that the hypothesis of independence H_0 is accepted in this model with greater than 90% confidence level.

Further analysis using these tests is required for drawing inferences on cosmological models or evolution of the luminosity function of galaxies beyond what has already been discussed by Bahcall & Tremaine (1988) and Caditz & Petrosian (1989). This is beyond the scope of our paper.

4.1.1. *Comparison of Tests*

The three test statistics described above give essentially the same result, but there are some interesting differences. We now consider the sources of these differences.

Figure 3 shows that the untrimmed t_1 statistic (4.3) is systematically larger than the trim-5 version (4.4). The size of the difference may seem surprising since only five points have been trimmed.

What happens is this. A low-lying point $[x_i, y_i(c)]$ gives a small value of $T_i(c)$, the smallest possible value being $-3^{1/2}$ obtained for $R_i = 1$ in equation (3.5). However, this same point makes a positive contribution of about $3^{1/2}/N_j$ to every $T_j(c)$ with $x_{(j)} > x_{(i)}$, that is, with $j < i$. The net contribution of the point to equation (4.3) is about

$$\sqrt{\frac{3}{n}}\left[(i - 1)\,\text{mean}_{j<i}\left\{\frac{1}{N_j}\right\} - 1\right]. \tag{4.9}$$

This can be quite large, especially for large values of c, where the N_j become small. For the Loh-Spillar data, mean (N_i) was 123 for $c = 5$ but only 69 for $c = 10$. The point at the lower left in Figure 1 is especially influential, with statistic (4.9) equaling 0.23 at $c = 5$.

The weights (4.6) for the theoretically optimum statistic t_{opt} are based on (3.13), with $u_i = 21.5 - 5.89x_i$, $x_{min} = -0.7$, and $y_{min} = 19.5$. Here $\hat{c} = 5.89$ is the point estimate for c based on the trim-5 t_1 statistic, and 19.5 is a lower bound for the 487 untrimmed y-values. The question arises as to why t_{opt} performed so poorly in Figure 3. Table 1 helps isolate the problem.

Each choice of c yields 492 values of the normalized rank statistics $T_i(c)$, (eq. [3.5]). Table 1 gives normed averages of the $T_i(c)$ for the nine strips shown in Figure 1; the entries are $A_j(c) = [n_j(c)]^{1/2}\bar{T}_j(c)$, where $n_j(c)$ is the number of untrimmed points $[x_i, y_i(c)]$ in jth strip, and $\bar{T}_j(c)$ is the average of the $T_i(c)$ values for those points. We expect approximately

$$A_j(c_0) \sim N(0, 1) \tag{4.10}$$

for the correct independence value c_0 of c. The $A_j(c)$ should be generally negative for $c > c_0$ and positive for $c < c_0$.

Strip 3 looks definitely anamolous in Table 1. All of the entries $A_3(c)$ are significantly negative, even for c near $\hat{c}_0 = 5.89$. The anomaly can be seen in Figure 1 as a deficiency of dim (high-magnitude) galaxies, rather than an excess of bright ones.

TABLE 1

COMPARISON OF t_{opt} AND t_1

STRIP	c = 3	4	5	6	7	8	9	10	DIFFERENCE 7–5	AVERAGE (w)
1	0.27	0.38	0.10	0.06	0.05	−0.06	−0.21	−0.33	−0.05	0.01
2	0.03	−0.06	0.32	0.25	0.17	0.06	−0.07	−0.08	−0.15	0.03
3	−2.61	−2.69	−2.90	−3.12	−3.13	−3.03	−3.22	−2.88	−0.23	0.05
4	3.14	2.19	1.92	1.84	1.78	1.42	1.09	0.99	−0.14	0.08
5	1.47	0.88	0.77	0.34	−0.03	−0.54	−1.36	−1.75	−0.80	0.11
6	2.53	2.03	1.71	0.80	0.24	−0.36	−0.37	−0.48	−1.47	0.14
7	1.14	0.62	−0.01	−0.62	−1.12	−1.47	−1.57	−1.63	−1.11	0.19
8	0.66	−0.44	−0.52	−1.17	−1.85	−2.10	−1.78	−1.57	−1.33	0.26
9	3.53	2.76	1.91	1.28	1.57	1.65	2.05	2.31	−0.34	0.40
$t_1[\text{data}_{(5)}(c)]$	3.39	1.88	1.09	−0.13	−0.80	−1.51	−1.86	−1.84	...	...
$t_1[\text{data}(c)]$	3.71	2.27	1.38	0.33	−0.24	−0.78	−0.98	−1.32	...	...
$t_{opt}[\text{data}_{(5)}(c)]$	4.28	2.65	1.59	0.23	−0.18	−0.65	−0.51	−0.35	...	...

NOTES.—Normed averages of normalized rank statistics $T_i(c)$ for the nine strips shown in Fig. 1. Each entry is $[n_j(c)]^{1/2}\bar{T}(c)$, where $n_j(c)$ is the number of untrimmed points in strip j, and $\bar{T}_j(c)$ is the average of $T_i(c)$ for these points. Shown at bottom are the overall permutation statistics plotted in Fig. 3. Theoretically the average differences between $c = 7$ and 5 should be approximately proportional to the average of the optimal weights w_i (eq. [4.5]). The last two columns show that this is roughly true for strips 1–8, but strip 9 is anamolous. The theoretically optimal statistic t_{opt} greatly overweights the data in strip 9.

Strip 9 is anomolous in a different way that causes particular problems for t_{opt}. The entries $A_j(c)$ actually increase as c increases from 6 to 10. For large c the number of eligible points N_i in the rank comparison gets small for $x_{(i)}$ in the ninth strip, and a majority of the comparisons are made with points in the eighth strip. The eighth strip has an exess of low-lying (low-magnitude) galaxies, leading to positive values of $T_i(c)$ in the ninth strip.

The column marked "average (w)" in Table 1 gives the strip averages of the optimal weights (eq. [4.6]). We see that t_{opt} puts its greatest weight on the points in strip nine. This causes the poor behavior of t_{opt} for large c seen in Figure 3. If the Loh-Spillar data is further truncated to exclude the ninth strip (the highest redshift bin where there could be some incompleteness of data), then t_{opt} performs a little better than t_1: the 90% confidence interval based on t_{opt} becomes [3.74, 6.96], compared to [3.54, 7.33] for t_1.

All of this is a reminder that redshift surveys are censuses of rather small finite populations, and not random samples of the universe. Our analysis is based on an assumption of spatial homogeneity and invariance of the distribution of absolute magnitudes. Even if this is true in the large, there may be significant deviations in the relatively small regions covered by a survey. Test statistics like t_1 are designed to reveal large-scale discrepancies in an hypothesized cosmology, like a wrong choice of c, and to ignore local deviations from independence like that in strip 3.

The next to rightmost column of Table 1 is the difference of averages in the jth strip, $A_j(7) - A_j(5)$. If the probabilistic assumptions leading to equation (3.13) are met, then these differences should be nearly proportional to the column of average optimal weights (cf. eq. [3.11]). This is roughly true for strips 1–8, but not for strip 9, which is another way to explain the failure of t_{opt}. A statistical truism is that very simple statistics, like t_1, often perform better in practice than theoretically more efficient, but more sensitive competitors.

As a compromise between simplicity and theoretical efficiency, we can calculate both t_1 and t_{opt}, using equation (3.14) to evaluate the joint significance level. Figure 4 traces $\mathbf{t}(c) = (t_1[\text{data}_{(5)}(c)], t_{opt}[\text{data}_{(5)}(c)])$ as a function of c. Vectors $\mathbf{t}(c)$ inside the ellipse indicate values of c with

$$t_w^2[\text{data}_{(5)}(c)] \leq \chi_2^{2(1-\alpha)}, \tag{4.11}$$

for $\alpha = 0.10$, that is, values of c passing the 2 degrees of freedom hypothesis test (3.14) at the 0.10 level. The point estimate $\hat{c} = 6.2$ is obtained by minimizing $t_w^2[\text{data}_{(5)}(c)]$. An approximate 90% confidence interval for c is given by those values satisfying $t_w^2 < \chi_2^{2(0.90)}$, $c \in [4.5, 8.7]$. A more conservative argument uses instead those values of c satisfying

$$t_w^2[\text{data}_{(5)}(c)] - t_w^2[\text{data}_{(5)}(\hat{c})] \leq \chi_2^{2(0.90)}. \tag{4.12}$$

This interval equals [4.4, 8.9] for the Loh-Spiral data.

4.1.2. *The Distribution Functions*

For the cosmological models where the independence hypothesis is acceptable (in our context), we can find the distribution (or density) functions $f(x)$ and $g(y)$ using Lynden-Bell's (1971) method. This nonparametric method gives an estimate of the cumulative distribution functions (cdf). The cdf of x is

$$\hat{F}(x) = \exp\left[-\sum_{j \leq i} 1/(N_i - \tfrac{1}{2})\right] \tag{4.13}$$

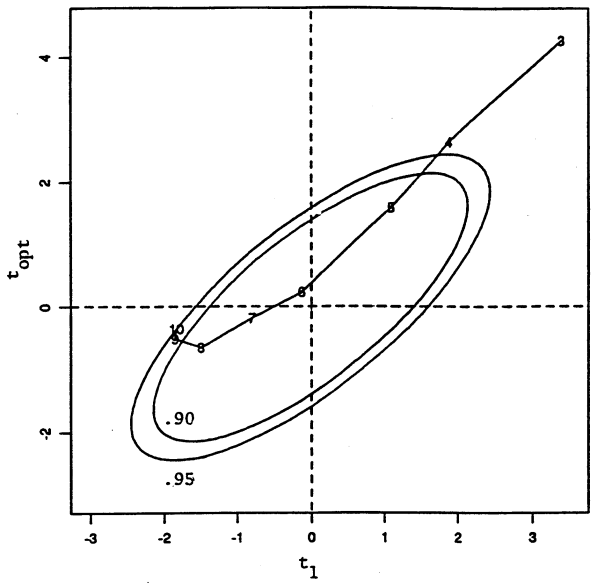

FIG. 4.—The vectors $\mathbf{t}(c) = (t_1[\text{data}_{(5)}(c)], t_{opt}[\text{data}_{(5)}(c)])$, plotted as a function of c. Values of c indicated along curve $\mathbf{t}(c)$; point estimate $\hat{c} = 6.2$ minimizes $t_w^2[\text{data}_{(5)}(c)]$ in eq. (3.14); ellipses indicate regions $t_w^2 < \chi_2^{2(1-\alpha)}$ for $1 - \alpha = 0.90$ and 0.95.

in our notation. An analogous expression gives $\hat{G}(y)$, the estimated cdf of y. Subtracting half, instead of one, from the N_i improves the correspondence between equation (4.13) and the usual nonparametric cdf estimate when there is no truncation. See Remark C of Efron (1977).

$\hat{F}(x)$ and $\hat{G}(y)$ were obtained for the data of Figure 1 and differentiated to give estimates $\hat{f}(x)$ and $\hat{g}(y)$ for the corresponding probability densities. Here we have used the cosmological model $\Omega = 1$, $\Lambda = 0$ where the comoving distance $r = 2(1 - 1/[1 + z]^{1/2})$ with $z = 10^x$. The comoving density estimate, $\hat{f} = \hat{f}(x)dz/dr$, is shown in Figure 5. As expected the

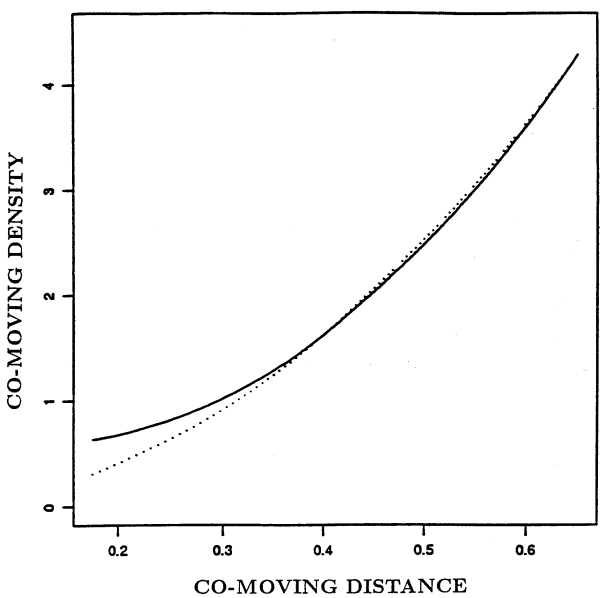

FIG. 5.—Density estimate for the number of galaxies as a function of comoving coordinate $r = 2[1 - 1/(1 + z)^{1/2}]$. Except for small r, this nearly matches r^2 (*dotted line*), predicted from volume considerations without any evolution.

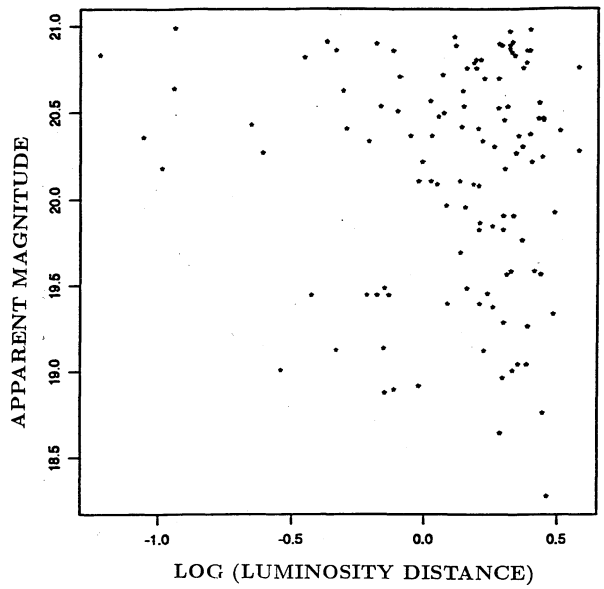

FIG. 6.—Redshift survey of 118 quasars, Boyle et al. (1990); vertical axis is apparent magnitude; horizontal axis is $x = \log(d)$ where luminosity distance $d = \{r(1+z)\}$, with $r = 2[1 - 1/(1+z)^{1/2}]$ for $\Omega = 1$, $\Lambda = 0$ model. Truncation boundary $m < m_0 = 21.0$. The permutation statistic $t_1 = (118)^{1/2}\bar{T}$ was applied to data sets $\{(x_i, m_i - cx_i), i = 1, 2, \ldots, 118\}$, giving point estimate $\hat{c} = 0.12$ and 90% confidence interval $c \in [-0.40, 0.11]$.

comoving density increases quadratically with comoving distance r except at small values.

4.2. *Quasar Data*

As another example we consider a small subsample of the quasar survey data by Boyle et al. (1990). Figure 6 shows the data for 118 quasars with the magnitude limit $m \leq m_0 = 21.0$. The plotted points are (x_i, m_i), where $x_i = \log d(z_i\ \Omega)$ for $\Omega = 1$, $\Lambda = 0$ model. The equally weighted permutation statistic $t_1[\text{data } (c)] = (118)^{1/2}\bar{T}(c)$ was applied to data sets

$$\text{data } (c) = \{(x_i, m_i - cx_i)\ ,\quad i = 1, 2, \ldots, 118\}\ . \tag{4.14}$$

In this case the cosmological "correct" value $c = 5$ is strongly rejected by the permutation test, with $t_1[\text{data } (5)] = -3.16$. The 90% confidence interval for c, that is, those values giving $|t_1[\text{data } (c)]| \leq 1.645$, is

$$c \in [-0.40, 0.11]\ . \tag{4.15}$$

This means that $c = 0$ is an acceptable hypothesis for the quasar data. In other words, the redshift and apparent (not absolute) magnitude of the quasars are independent so that there is no obvious Hubble relation as stressed by G. Burbidge frequently. In most conventional cosmological models this requires a strong evolution of luminosities (or, more generally, a strong evolution of the luminosity function) of quasars. The unsettling fact is that there seems to be a conspiracy such that the dimming of the sources due to their cosmological distances is cancelled out by the evolution of their luminosities (see Fig. 4 of Caditz & Petrosian 1990).

5. SUMMARY

In analysis of astronomical data, it is often necessary to give a statistical description of bivariate distribution of two physically important parameters from a truncated data. If the two parameters are statistically independent, then there is a unique maximum likelihood method based on Lynden-Bell's (1971) c^- method. It is, therefore, important to establish independence for truncated data.

We have described modification of the permutation test of independence for truncated data related to this method and based on a theory similar to that in Bhattacharya et al. (1983). These tests will be useful in analysis of any data when a bivariate (or more generally a multivariate) distribution is required. A very common example is the distribution in space and luminosities of the sources.

We have applied this test to magnitude-limited galaxy and quasar surveys using rank statistics. We have demonstrated how this test allows us to determine whether or not redshift and absolute magnitudes (or luminosities) of sources are statistically independent in an assumed cosmological model. Or, conversely, we show that it is possible to set limits on cosmological parameters and on the evolution of the luminosities or densities of the sources if we assume independence. For example, from analysis of the Loh & Spillar (1988) data on galaxies, we show that absolute magnitude and redshift are very nearly independent in conventional cosmological models which at this redshift of the survey can be approximated by a simple Hubble expansion law. Thus assuming independence, which is a reasonable approximation for galaxies of redshifts $z < 1$, one can rule out unconventional models such as the chronomatic model of Segal, at least to the Loh & Spillar data.

On the other hand, similar analysis of a small subsample of a magnitude-limited quasar data by Boyle et al. (1990) shows that independence of redshift and absolute luminosities can be rejected in most conventional models so that if such models are correct, there must be strong luminosity evolution (consistent with previous results of Boyle et al. 1987, Caditz & Petrosian 1990) in the sense of higher luminosities at larger redshift. This evolution corrects the dimming due to distance of the source resulting in statistically independent distribution of redshift and *apparent* magnitudes.

V. P. was supported partially by NASA grant NAGW 2290. B. E. was supported by grants NIH 5GM21215-17 and NSF DMS 89-05-874.

REFERENCES

Bahcall, S. R., & Tremaine, S. 1988, ApJ, 326, L1
Bhattacharya, P., Chernoff, H., & Yang, S. 1983, Ann. Statist., 11, 505
Boyle, B. J., Fong, R., Shanks, T., & Peterson, B. A. 1987, MNRAS, 227, 717
———. 1990, MNRAS, 243, 1
Caditz, D. M., & Petrosian, V. 1989, ApJ, 337, L65
———. 1990, ApJ, 357, 326
Efron, B. 1977, J. Am. Stat. Assoc., 72, 557
Hajek, J. 1969, Nonparametric Statistics (San Francisco: Holden-Day)
Loh, E. D., & Spillar, E. J. 1986, ApJ, 303, 154
Lynden-Bell, D. 1971, MNRAS, 155, 95
Petrosian, V. 1986, in Structure and Evolution of Active Galactic Nuclei, ed. G. Giunian (Dordrecht: Reidel), 355
———. 1992, in Proc. Statistical Challenges in Modern Astronomy, Penn State University, in press
Segal, I., & Nicoll, J. 1986, ApJ, 300, 224