A TEST OF LINEARITY VERSUS CONVEXITY OF A MEDIAN REGRESSION CURVE[1]

By Bruce Marvin Hill[2]

Stanford University

1. Introduction and summary. In this paper we shall propose a test of linearity of a median regression curve against an alternative of convexity. To be specific, we shall test

$$H_0: Y_i = \alpha + \beta X_i + \epsilon_i, \qquad i = 0, 1, \cdots, n,$$

against

$$H_1: Y_i = \phi(X_i) + \epsilon_i, \qquad i = 0, 1, \cdots, n,$$

where α, β and ϕ are unspecified, and $\phi(x)$ is a nonlinear convex function. The basic assumption underlying the test is that the ϵ_i are independent identically distributed random variables with median zero and with a continuous density function $f(\epsilon)$ such that $f(0) > 0$. The X_i are fixed and known.

The test consists in estimating a line by the Mood-Brown procedure (using medians) from a central subset of the observations, making a weighted count of the number of remaining observations lying above the line, and rejecting H_0 if this number, R_n, is large. The test can easily be adapted to a one-sided alternative of concavity or to a two-sided alternative of either convexity or concavity.

Section 2 is devoted to a discussion of the line estimation procedure, and in particular, the asymptotic distribution of the estimator is obtained under the null and alternative hypotheses. In Section 3 the R test of convexity is introduced, the asymptotic null and alternative hypothesis distributions of the test statistic R_n are obtained, and a formula for the asymptotic power is given. The test is shown to be consistent against twice differentiable convex alternatives.

In Section 4 we obtain the relative asymptotic efficiency of the R test as compared to the least squares test for parabolic alternatives with errors normally distributed, and make recommendations for the use of the test. Finally, in the Appendix, results of some Monte Carlo experiments used to investigate the small sample behavior of R_n under H_0 are presented.

The author is not aware of the existence of other tests of linearity against general convex alternatives. In the case where the alternatives in mind can be expressed in a linear regression scheme, both the least squares test and the median test suggested by Mood in [5] are possible competitors of the convexity test presented in this paper. The least squares test is to be preferred when errors are

Received April 13, 1961; revised April 16, 1962.

[1] The report is based upon the author's doctoral dissertation, and was sponsored in part by a Public Health Service Training Grant in Biometry.

[2] Now at the University of Michigan.

known to be normally distributed with common variance, but for more general types of errors there is at present no obvious way of determining a "best" test, and consequently the choice of test to be used in a given situation must be based largely upon subjective considerations.

2. The Mood-Brown estimation procedure. The proposed test of linearity is defined in terms of a line estimated from a set of points (X_i, Y_i) by means of the Mood-Brown procedure, [1] and [5]. Because of the crucial role this procedure is to play we shall defer the introduction of the test of linearity until Section 3, and in the present section we shall be concerned with the distribution of the line estimator.

2.1. *Definition of the estimator*. For completeness we first define the median of $n + 1$ numbers $X_0 \leqq X_1 \leqq X_2 \leqq \cdots \leqq X_n$ to be

$$M_n = \begin{cases} X_k & \text{if } n = 2k \\ \tfrac{1}{2}(X_k + X_{k+1}) & \text{if } n = 2k + 1 \end{cases}$$

The Mood-Brown line corresponding to a set of points $(X_i, Y_i), i = 0, 1, \cdots, n$, is then defined to be the line $\hat{\alpha}_n + \hat{\beta}_n x$ (denoted $(\hat{\beta}_n, \hat{\alpha}_n)$) satisfying the equations

$$\text{median}_{\{i:X_i \leqq M_n\}}[Y_i - \hat{\alpha}_n - \hat{\beta}_n X_i] = \text{median}_{\{i:X_i > M_n\}}[Y_i - \hat{\alpha}_n - \hat{\beta}_n X_i] = 0.$$

To justify this definition we prove

THEOREM 1. *Corresponding to any set of points (X_i, Y_i) with at least two of the X_i distinct there exists a unique Mood-Brown estimate $(\hat{\beta}_n, \hat{\alpha}_n)$.*

PROOF. Without loss of generality we can assume that the line $x = 0$ strictly separates the points with $X_i \leqq M_n$ and those with $X_i > M_n$. We first show that for each point $(0, \alpha)$ on the vertical line $x = 0$ there exist unique lines having the parameters $(\beta_1(\alpha), \alpha)$ and $(\beta_2(\alpha), \alpha)$ such that

$$\text{median}_{\{X_i \leqq M_n\}}[Y_i - \alpha - \beta_1(\alpha)X_i] = \text{median}_{\{X_i > M_n\}}[Y_i - \alpha - \beta_2(\alpha)X_i] = 0.$$

For we can choose a line passing through $(0, \alpha)$ and lying below all the (X_i, Y_i) with $X_i \leqq M_n$. By rotating this line about $(0, \alpha)$ as center we can pass through each of these points until finally all lie below the rotated line. Since the function $m_1(\beta, \alpha) = \text{median}_{\{X_i \leqq M_n\}}[Y_i - \alpha - \beta X_i]$ is a continuous and strictly monotonic function of β for fixed α, which takes on both positive and negative values by the above rotation argument, it follows that there exists a unique slope $\beta_1(\alpha)$ for which $m_1(\beta_1(\alpha), \alpha) = 0$. Similarly we obtain a unique slope $\beta_2(\alpha)$ such that $m_2(\beta_2(\alpha), \alpha) = 0$, where $m_2(\beta, \alpha) = \text{median}_{\{X_i > M_n\}}[Y_i - \alpha - \beta X_i]$. Noting that $\beta_1(\alpha)$ is a continuous and strictly increasing function of α such that $\lim_{\alpha \to -\infty} \beta_1(\alpha) = -\infty$, $\lim_{\alpha \to +\infty} \beta_1(\alpha) = +\infty$, and that $\beta_2(\alpha)$ is a continuous and strictly decreasing function of α such that $\lim_{\alpha \to -\infty} \beta_2(\alpha) = +\infty$, $\lim_{\alpha \to +\infty} \beta_2(\alpha) = -\infty$, it follows that there exists a unique line (β_0, α_0) such that $m_1(\beta_0, \alpha_0) = 0$, and so $(\hat{\beta}_n, \hat{\alpha}_n) = (\beta_0, \alpha_0)$.

COROLLARY. *Let the X_i be fixed with at least two distinct, and let $(Y_0, \cdots, Y_n)$ be* *an observation from an absolutely continuous $(n+1)$-dimensional distribution. Then with probability one the number of points (X_i, Y_i) with $X_i \leqq M_n$ which lie upon the line $(\hat{\beta}_n, \hat{\alpha}_n)$ is one or zero according as the total number of points with $X_i \leqq M_n$ is odd or even. The same statement holds for the number of points lying upon $(\hat{\beta}_n, \hat{\alpha}_n)$ with $X_i > M_n$.*

Proof. The proof follows from the assumption of absolute continuity, which implies that the probability of any linear relationship holding between three or more of the (X_i, Y_i) is zero.

2.2. *Asymptotic distribution under H_0.* We assume that the set of points (X_i, Y_i) is such that $X_i = h(i/n)$, $i = 0, 1, \cdots, n$, where $h(t)$ is a continuous and strictly monotonically increasing function defined on the interval $0 \leqq t \leqq 1$, with $h(0) = c$, $h(1) = d$. The function $h(t)$ is a spacing function, allowing the points X_i to be spaced in any designated manner such that $-\infty < X_0 < X_1 < \cdots < X_n < +\infty$. Under the null hypothesis of the test of linearity we assume $Y_i = \alpha + \beta X_i + \epsilon_i$, where the ϵ_i are independent identically distributed random variables having common density function $f(\epsilon)$ and cumulative distribution function $F(x) = \int_{-\infty}^{x} f(\epsilon)\,d\epsilon$ such that $f(0) \neq 0$, $F(0) = \frac{1}{2}$, and f is continuous. We shall now prove

Theorem 2. *Under the above assumptions, the random vector $(\hat{\eta}_n, \hat{\xi}_n) = m^{\frac{1}{2}}(\hat{\beta}_n - \beta, \hat{\alpha}_n - \alpha)$ is asymptotically normally distributed with mean $(0, 0)$ and covariance matrix*

$$\Sigma = \left(f(0)\left[\int_{\frac{1}{2}}^{1} h(t)\,dt - \int_{0}^{\frac{1}{2}} h(t)\,dt\right]\right)^{-2} \cdot \begin{pmatrix} \frac{1}{8} & -\frac{1}{8}\int_0^1 h(t)\,dt \\ -\frac{1}{8}\int_0^1 h(t)\,dt & \frac{1}{4}\left[\left(\int_0^{\frac{1}{2}} h(t)\,dt\right)^2 + \left(\int_{\frac{1}{2}}^1 h(t)\,dt\right)^2\right] \end{pmatrix}$$

where $m = (n+1)/2$ tends to ∞.

Proof. Without loss of generality we can take $\alpha = \beta = 0$. For let $(\hat{\beta}_n', \hat{\alpha}_n')$ be the Mood-Brown estimate that would be obtained from the (unobservable) sample (X_i, Y_i'), where $Y_i' = Y_i - \alpha - \beta X_i = \epsilon_i$, and let $Z_i = Y_i' - \hat{\alpha}_n' - \hat{\beta}_n' X_i$. Hence $\text{median}_{\{X_i \leqq M_n\}} Z_i = \text{median}_{\{X_i > M_n\}} Z_i = 0$. But $Z_i = Y_i - (\hat{\alpha}_n' + \alpha) - (\hat{\beta}_n' + \beta)X_i$, and so by definition of $(\hat{\beta}_n, \hat{\alpha}_n)$ we have $(\hat{\beta}_n, \hat{\alpha}_n) = (\hat{\beta}_n' + \beta, \hat{\alpha}_n' + \alpha)$. Since the distribution of $(\hat{\beta}_n', \hat{\alpha}_n') = (\hat{\beta}_n - \beta, \hat{\alpha}_n - \alpha)$ does not depend upon (β, α) it thus suffices to take $\alpha = \beta = 0$.

Now let (b, a) be a fixed line and define

$V_{1i}(b, a)$ = number of (X_k, Y_k) lying above (b, a), $X_k \leqq M_n$ and $k \neq i$,

$V_{2j}(b, a)$ = number of (X_k, Y_k) lying above (b, a), $X_k > M_n$ and $k \neq j$,

$V_1(b, a)$ = number of (X_k, Y_k) lying above (b, a), $X_k \leqq M_n$ and

$V_2(b, a)$ = number of (X_k, Y_k) lying above (b, a), $X_k > M_n$.

In order to avoid trivial complications we shall first assume that m tends to ∞ through a sequence of odd integers, and later show that this assumption can be dropped. Since m is odd it follows from the Corollary to Theorem 1 that with probability one the line $(\hat{\beta}_n, \hat{\alpha}_n)$ passes through exactly one point (X_i, Y_i) with $X_i \leqq M_n$ and exactly one point (X_j, Y_j) with $X_j > M_n$. Let $g_n(b, a)$ be the density function of $(\hat{\beta}_n, \hat{\alpha}_n)$ evaluated at (b, a), and $k_n(b, a)$ be the density function of $(\hat{\eta}_n, \hat{\xi}_n) = m^{\frac{1}{2}}(\hat{\beta}_n, \hat{\alpha}_n)$ evaluated at (b, a). Hence

$$k_n(b, a)\, db\, da = (1/m) g_n(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}})\, db\, da.$$

In Lemma 1 of the Appendix we prove that

$$g_n(b, a)\, db\, da = \sum_{(i,j)\varepsilon\mathfrak{A}} \Pr\{V_{1i}(b, a) = \tfrac{1}{2}(m-1),\ V_{2j}(b, a) = \tfrac{1}{2}(m-1)\}$$

$$\cdot f(a + bX_i)\, f(a + bX_j)\, d(a + bX_i)\, d(a + bX_j),$$

where $\mathfrak{A} = \{(i, j): X_i \leqq M_n, X_j > M_n\}$. This equation is a consequence of the above remark concerning the Corollary to Theorem 1.

Let $\mathfrak{A}_1 = \{i: X_i \leqq M_n\}$, $\mathfrak{A}_2 = \{j: X_j > M_n\}$. We then have

$$g_n(b, a)\, db\, da = \sum_{i\varepsilon\mathfrak{A}_1} \Pr\{V_{1i}(b, a) = \tfrac{1}{2}(m-1)\} f(a + bX_i)\, d(a + bX_i)$$

$$\cdot \sum_{j\varepsilon\mathfrak{A}_2} \Pr\{V_{2j}(b, a) = \tfrac{1}{2}(m-1)\} f(a + bX_j)\, d(a + bX_j)$$

because of the mutual independence of the random variables $V_{1i}(b, a)$ and $V_{2j}(b, a)$. It follows that

$$k_n(b, a)\, db\, da$$

$$= \frac{1}{m} \sum_{i\varepsilon\mathfrak{A}_1} \Pr\left\{V_{1i}\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{1}{2}(m-1)\right\} f\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right) d(a + bX_i)$$

$$\cdot \sum_{j\varepsilon\mathfrak{A}_2} \Pr\left\{V_{2j}\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{1}{2}(m-1)\right\} f\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_j\right) d(a + bX_j).$$

The next step in the proof of asymptotic normality will be to show that $k_n(b, a)$ converges to a normal density function $k(b, a)$ at each (b, a).

By a generalization of the Liapounoff version of the Central Limit Theorem to sequences of sums of independent random variables we obtain that

$$V_1^*\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{V_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) - EV_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right)}{\left(\operatorname{Var} V_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right)\right)^{\frac{1}{2}}}$$

is asymptotically normally distributed with mean zero and unit variance for each (b, a). This result is an immediate consequence of Lemma 4 of the Appendix, where we note that $V_1(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}})$ is a sum of indicator variables

$$Z_i(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}}) = \begin{cases} 1 & \text{if } \quad Y_i > a/m^{\frac{1}{2}} + b/m^{\frac{1}{2}} X_i \\ 0 & \text{if not} \end{cases}$$

and

$$\Pr\left\{Z_i\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = 1\right\} = 1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right) = 1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} h\left(\frac{i}{n}\right)\right).$$

We then have

$$\begin{aligned} EV_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) &= \sum_{i \varepsilon \mathfrak{A}_1} \left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right)\right] \\ &\sim n \int_0^{\frac{1}{2}} \left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} h(t)\right)\right] dt \\ &\sim n \int_0^{\frac{1}{2}} \left[\frac{1}{2} - (a/m^{\frac{1}{2}})f(0) - (b/m^{\frac{1}{2}})h(t)f(0)\right] dt \\ &\sim \frac{n}{4} - [(n/2)^{\frac{1}{2}}f(0)]a - \left[(2n)^{\frac{1}{2}}f(0) \int_0^{\frac{1}{2}} h(t)\, dt\right] b \quad \text{as} \quad n \text{ tends to } \infty, \end{aligned}$$

and similarly,

$$\begin{aligned} \operatorname{Var}\left[V_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right)\right] &= \sum_{i \varepsilon \mathfrak{A}_1} F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right)\left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right)\right] \\ &\sim n \int_0^{\frac{1}{2}} F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} h(t)\right)\left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} h(t)\right)\right] dt \sim \frac{n}{8}. \end{aligned}$$

It then follows that

$$\begin{aligned} &\Pr\left\{V_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{1}{2}(m - 1)\right\} \\ &\qquad = \Pr\left\{V_1^*\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{\frac{1}{2}(m - 1) - EV_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right)}{\left[\operatorname{Var} V_1\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right)\right]^{\frac{1}{2}}}\right\} \\ &\sim \Pr\left\{V_1^*\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{\frac{1}{2}(m - 1) - \left(\frac{1}{4}n - \left[(\frac{1}{2}n)^{\frac{1}{2}}f(0)\right]a - \left[(2n)^{\frac{1}{2}}f(0)\int_0^{\frac{1}{2}} h(t)\, dt\right]b\right)}{(n/8)^{\frac{1}{2}}}\right\} \\ &\sim \Pr\left\{V_1^*\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = 2f(0)a + 4f(0)\int_0^{\frac{1}{2}} h(t)\, dt\, b\right\} \end{aligned}$$

$$\sim (2\pi(n/8))^{-\frac{1}{2}} \exp\left(-\frac{1}{2} y_1^2\right),$$

where $y_1 = 2f(0)\ a + 4f(0) \int_0^{\frac{1}{2}} h(t)\ dt\ b$.

In a similar manner we obtain

$$\Pr\{V_2(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}}) = \tfrac{1}{2}(m-1)\} \sim (2\pi(n/8))^{-\frac{1}{2}} \exp(-\tfrac{1}{2}y_2^2),$$

where $y_2 = 2f(0)\ a + 4f(0) \int_{\frac{1}{2}}^1 h(t)\ dt\ b$.

It can then easily be shown that the factors

$$\frac{1}{m^{\frac{1}{2}}} \sum_{i \epsilon \alpha_1} \Pr\left\{V_{1i}\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{1}{2}(m-1)\right\} f\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right)$$

and

$$\frac{1}{m^{\frac{1}{2}}} \sum_{j \epsilon \alpha_2} \Pr\left\{V_{2j}\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \frac{1}{2}(m-1)\right\} f\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_j\right)$$

of $k_n(b, a)$ converge to $2f(0)(2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}y_1^2)$ and $2f(0)(2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2}y_2^2)$, respectively, as n tends to ∞. Hence

$$\lim_{n\to\infty} k_n(b, a)\ db\ da = (2\pi)^{-1} \exp\{-\tfrac{1}{2}(y_1^2 + y_2^2)\}\ dy_1\ dy_2 = k(b, a)\ db\ da.$$

It is straightforward to show that $k(b, a)$ is a bivariate normal density with mean and covariance matrix as claimed in Theorem 2, so to complete the proof in the case where m is odd it remains only to show

$$\lim_{n\to\infty} \int_B k_n(b, a)\ db\ da = \int_B k(b, a)\ db\ da$$

for all Borel sets B in the plane. Now Scheffé in [6] proves that sufficient conditions to yield the above result are that $k_n(b, a)$ and $k(b, a)$ be densities, and that $\lim_{n\to\infty} k_n(b, a) = k(b, a)$ for almost all (b, a) in the plane. Since these conditions are satisfied the proof of Theorem 2 is complete for the case where m is odd.

Next let us consider the case where the number of points on one side of the median M_n is even, say $2K$. Let $d_j = Y_j - a - bX_j$ for some fixed (b, a), and let $d_{(j)}^{(2K)}$ be the jth order statistic from these $2K$ numbers. Then the condition that median $(d_i) = 0$ is equivalent to the condition $\frac{1}{2}[d_{(K)}^{(2K)} + d_{(K+1)}^{(2K)}] = 0$, and so each of these conditions is the defining condition that $(\hat\beta_n, \hat\alpha_n) = (b, a)$ for the $2K$ points under consideration. (An analogous condition is, of course, required for the points on the other side of M_n in order that $(\hat\beta_n, \hat\alpha_n) = (b, a)$.) On the other hand, in the case where the given set of points is odd, say $2K + 1$, the defining condition for the set of points is that

$$\text{median}(d_i) = d_{(K+1)}^{(2K+1)} = 0,$$

where $d_{(K+1)}^{(2K+1)}$ is the $K + 1$ order statistic from the $2K + 1$ numbers. But it is evident that the asymptotic distributions of the three order statistics $d_{(K)}^{(2K)}$, $d_{(K+1)}^{(2K)}$ and $d_{(K+1)}^{(2K+1)}$ are the same, and, moreover, that $d_{(K)}^{(2K)}$ and $d_{(K+1)}^{(2K)}$ taken from the same sample are asymptotically perfectly correlated as K tends to ∞. It follows that the asymptotic distribution of $\frac{1}{2}[d_{(K)}^{(2K)} + d_{(K+1)}^{(2K)}]$ is the same as that of $d_{(K+1)}^{(2K+1)}$. Since the asymptotic distribution of these two medians determines the asymptotic distribution of $(\hat{\beta}_n, \hat{\alpha}_n)$ it follows that the latter must be independent of the manner in which the number of points tends to ∞, and so the proof of Theorem 2 is complete.

A special case of interest occurs when $h(t)$ is linear with positive slope. This case corresponds to equally spaced X_i in an interval.

Taking $h(t) = c + (d - c)t$ with $c < d$ we obtain $(\hat{\eta}_n, \hat{\xi}_n)$ is distributed asymptotically like $(\hat{\eta}, \hat{\xi})$, where

$$\operatorname{Var} \hat{\eta} = \frac{2}{f^2(0)(d-c)^2}$$

$$\operatorname{Var} \hat{\xi} = \frac{10c^2 + 12cd + 10d^2}{16f^2(0)(d-c)^2}$$

$$\operatorname{Cov}(\hat{\xi}, \hat{\eta}) = \frac{-(c+d)}{f^2(0)(d-c)^2}.$$

2.3. *Asymptotic distribution under the alternative.* In 2.2 we found under the null hypothesis $Y_i = \alpha + \beta X_i + \epsilon_i$ and the assumptions of Theorem 2 that the Mood-Brown estimate $(\hat{\beta}_n, \hat{\alpha}_n)$ is such that $[\frac{1}{2}(n+1)]^{\frac{1}{2}} (\hat{\beta}_n - \beta, \hat{\alpha}_n - \alpha)$ is asymptotically normally distributed with mean $(0, 0)$ and covariance matrix Σ. It follows that $(\hat{\beta}_n, \hat{\alpha}_n)$ is a consistent estimator of (β, α).

In order to investigate the asymptotic power of the test of linearity proposed in Section 3 it is necessary also to know the asymptotic distribution of $(\hat{\beta}_n, \hat{\alpha}_n)$ when the alternative hypothesis holds, i.e., when $Y_i = \phi(X_i) + \epsilon_i$, $i = 0, 1, \cdots, n$, and $\phi(x)$ is a strictly convex function.

Under mild assumptions we shall show in this subsection that corresponding to the convex function $\phi(x)$ there exists a unique line $(\beta_0(\phi), \alpha_0(\phi))$ such that $[\frac{1}{2}(n+1)]^{\frac{1}{2}} (\hat{\beta}_n - \beta_0, \hat{\alpha}_n - \alpha_0)$ is asymptotically normally distributed with mean $(0, 0)$, and hence that $(\hat{\beta}_n, \hat{\alpha}_n)$ is a consistent estimate of $(\beta_0, \alpha_0) = (\beta_0(\phi), \alpha_0(\phi))$. The line (β_0, α_0) will depend upon the spacing function $h(t)$, the distribution function F of the ϵ_i, and the convex function ϕ specified as alternative. Defining $V_1(b, a)$ and $V_2(b, a)$ as in 2.2, we shall find that (β_0, α_0) is the unique line such that $EV_1(\beta_0, \alpha_0) \sim \frac{1}{4}n$ and $EV_2(\beta_0, \alpha_0) \sim \frac{1}{4}n$, or equivalently such that the equations

$$\text{(2.3.1)} \quad \int_0^{\frac{1}{2}} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] \, dt = \int_{\frac{1}{2}}^1 [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] \, dt = \tfrac{1}{4}$$

are satisfied.

We now state Lemma 6, the proof of which is given in the Appendix.

LEMMA 6. *Let F be a cumulative distribution function with $F(x) = \int_{-\infty}^{x} f(\epsilon)\, d\epsilon$ for all x, $F(0) = \frac{1}{2}$, $f(0) = F'(0) > 0$, and f continuous. Let $h(t)$ be continuous* *and strictly monotonically increasing for* $0 \leqq t \leqq 1$. *Then corresponding to each strictly convex twice differentiable function* $\phi(x)$ *there exists a unique solution* $(\beta_0(\phi), \alpha_0(\phi))$ *satisfying equations* (2.3.1).

The main result of this subsection is stated in

THEOREM 3. *Let* $Y_i = \phi(X_i) + \epsilon_i$, $i = 0, 1, \cdots, n$, *where* $\phi(x)$ *is strictly convex and twice differentiable, and let* $X_i = h(i/n)$, *where* $h(0) = c$, $h(1) = d$, *and* $h(t)$ *is strictly monotonically increasing and continuous for* $0 \leqq t \leqq 1$. *Let the* ϵ_i *be independent identically distributed random variables with common density* $f(\epsilon)$ *and cumulative distribution function* $F(x) = \int_{-\infty}^{x} f(\epsilon)\, d\epsilon$ *such that* $F(0) = \frac{1}{2}$, $f(0) > 0$, *and* f *is continuous. We assume that*

$$F[\sup_{0 \leqq t \leqq 1}(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] < 1$$

and

$$F[\inf_{0 \leqq t \leqq 1}(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] > 0,$$

where (β_0, α_0) *is the unique solution of* (2.3.1). *Let* $m = (n + 1)/2$ *tend to* ∞. *Then the random vector* $(\hat{\eta}_n, \hat{\xi}_n) = m^{\frac{1}{2}}(\hat{\beta}_n - \beta_0, \hat{\alpha}_n - \alpha_0)$ *is asymptotically normally distributed with mean* (0, 0) *and covariances given by* (2.3.2) *below.*

PROOF. The proof is very similar to that of Theorem 2, and consequently we shall only indicate the necessary modifications of that proof.

We define $V_{1i}(b, a)$, $V_{2j}(b, a)$, $V_1(b, a)$, $V_2(b, a)$, $\mathfrak{A}_1$ and $\mathfrak{A}_2$, as in 2.2. Let $k_n(b, a)$ and $g_n(b, a)$ be the density functions of $(\hat{\eta}_n, \hat{\xi}_n)$ and $(\hat{\beta}_n, \hat{\alpha}_n)$, respectively, evaluated at (b, a).

From Lemma 5 of the Appendix it follows that

$$V_1^*(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = \frac{V_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) - EV_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}})}{[\text{Var } V_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}})]^{\frac{1}{2}}}$$

is asymptotically normally distributed with mean zero and variance one for each (b, a). Here we note that

$$V_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = \sum_{i \in \mathfrak{A}_1} Z_i(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}),$$

where

$$Z_i(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = \begin{cases} 1 & \text{if } \; Y_i > (\alpha_0 + a/m^{\frac{1}{2}}) + (\beta_0 + b/m^{\frac{1}{2}})X_i \\ 0 & \text{if not,} \end{cases}$$

and so

$$\begin{aligned} \Pr\{Z_i(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = 1\} &= 1 - F(\alpha_0 + a/m^{\frac{1}{2}} + (\beta_0 + b/m^{\frac{1}{2}})X_i - \phi[X_i]) \\ &= 1 - F(\alpha_0 + a/m^{\frac{1}{2}} + (\beta_0 + b/m^{\frac{1}{2}})\, h(i/n) - \phi[h(i/n)]). \end{aligned}$$

Using the Riemann integral approximation to a sum, and expanding $F(\alpha_0 +$

$a/m^{\frac{1}{2}} + (\beta_0 + b/m^{\frac{1}{2}})\, h(t) - \phi[h(t)])$ in a Taylor Series about $\alpha_0 + \beta_0 h(t) - \phi[h(t)]$, we obtain

$$EV_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) \sim n \int_0^{\frac{1}{2}} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)]) - [(a/m^{\frac{1}{2}}) + (b/m^{\frac{1}{2}})h(t)]f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]\, dt,$$

and

$$\operatorname{Var} V_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) \sim n \int_0^{\frac{1}{2}} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt.$$

Hence

$$\Pr\{V_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = \tfrac{1}{2}(m-1)\} = \Pr\left\{V_1^*(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = \frac{\tfrac{1}{2}(m-1) - EV_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}})}{[\operatorname{Var} V_1(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}})]^{\frac{1}{2}}}\right\} \sim \Pr\{V_1^*(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}}) = c_n(b, a)\},$$

where

$$c_n(b,a) = \frac{\tfrac{1}{2}(m-1) - n \int_0^{\frac{1}{2}} \{1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)]) - [(a/m^{\frac{1}{2}}) + (b/m^{\frac{1}{2}})h(t)]f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\}\, dt}{n^{\frac{1}{2}}\left(\int_0^{\frac{1}{2}} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right)^{\frac{1}{2}}}.$$

The reason for the equations (2.3.1) defining (β_0, α_0) is now apparent. In order that $\frac{1}{2}(m-1) - n \int_0^{\frac{1}{2}}[1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]dt = O(1)$, and hence that $\lim_{n\to\infty} c_n(b, a)$ exist and be finite, we must have $\int_0^{\frac{1}{2}}[1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]\, dt = \frac{1}{4}$.

Deriving the corresponding result for $V_2^*(\beta_0 + b/m^{\frac{1}{2}}, \alpha_0 + a/m^{\frac{1}{2}})$, and utilizing an expression for $k_n(b, a)$ similar to that used in the proof of Theorem 2, we then obtain

$$\lim_{n\to\infty} k_n(b, a)\, db\, da = (2\pi)^{-1} \exp[-\tfrac{1}{2}(y_1^2 + y_2^2)]\, dy_1\, dy_2 = k(b, a)\, db\, da,$$

where

$$y_1 = \sqrt{2}\left(\int_0^{\frac{1}{2}} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right)^{-\frac{1}{2}} \cdot \int_0^{\frac{1}{2}} (a + bh(t))f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt,$$

$$y_2 = \sqrt{2}\left(\int_{\frac{1}{2}}^{1} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right)^{-\frac{1}{2}} \cdot \int_{\frac{1}{2}}^{1} (a + bh(t))f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt.$$

It is straightforward to show that the limiting distribution of $(\hat{\eta}_n, \hat{\xi}_n)$ is that of a bivariate normal distribution having the density function $k(b, a)$, and after some calculation we obtain that $(\hat{\eta}_n, \hat{\xi}_n)$ is distributed asymptotically like $(\hat{\eta}, \hat{\xi})$, where

$$\operatorname{Var} \hat{\eta} = (2/D^2)\left(\left[D_1 \int_0^{\frac{1}{2}} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right]^2 + \left[D_2 \int_{\frac{1}{2}}^{1} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right]^2\right),$$

$$\operatorname{Var} \hat{\xi} = (2/D^2)\left(\left[D_1 \int_0^{\frac{1}{2}} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])h(t)\, dt\right]^2 + \left[D_2 \int_{\frac{1}{2}}^{1} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])h(t)\, dt\right]^2\right),$$

(2.3.2)

$$\operatorname{Cov}(\hat{\xi}, \hat{\eta}) = -(2/D^2)\left(D_1^2 \int_0^{\frac{1}{2}} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt \cdot \int_0^{\frac{1}{2}} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])h(t)\, dt + D_2^2 \cdot \int_{\frac{1}{2}}^{1} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt \cdot \int_{\frac{1}{2}}^{1} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])h(t)\, dt\right),$$

and where

$$D_1 = \left(\int_0^{\frac{1}{2}} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right)^{-\frac{1}{2}},$$

$$D_2 = \left(\int_{\frac{1}{2}}^{1} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right)^{-\frac{1}{2}},$$

and

$$D^2 = \left[2D_1 D_2 \int_{\frac{1}{2}}^{1} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])h(t)\, dt \cdot \int_0^{\frac{1}{2}} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt\right.$$

$$-2D_1 D_2 \int_0^{\frac{1}{2}} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])h(t)\, dt$$

$$\cdot \int_{\frac{1}{2}}^{1} f(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt \Big]^2.$$

This completes the sketch of the proof of Theorem 3.

In the special case $h(t) = c + (d - c)t$ and $\phi(x) = \gamma x^2$ it is easy to show $\beta_0 = \beta_0(\phi) = \phi'\frac{1}{2}(c + d) = \gamma(c + d)$. Hence the difference between the line $(\beta_0, \alpha_0) = (\beta_0(\phi), \alpha_0(\phi))$ and the function $\phi(x)$ is given by $d(x) = \alpha_0 + \beta_0 x - \phi(x) = \alpha_0 + \gamma(c + d)x - \gamma x^2 = \alpha_0 + \gamma\frac{1}{4}(c + d)^2 - \gamma(x - \frac{1}{2}(c + d))^2$, which is symmetric about the line $x = \frac{1}{2}(c + d)$.

If F is uniform on $[-K, K]$ for sufficiently large K, we further obtain $\alpha_0 = \alpha_0(\phi) = -(\gamma/6)[(c + d)^2 + 2cd]$, and $d(x) = (\gamma/12)(d - c)^2 - \gamma(x - \frac{1}{2}(c + d))^2$. In this case (β_0, α_0) is the least squares line to the curve $\phi(x)$ in the interval $[c, d]$.

2.4. *Summary and remarks.* In the present section we have proved the existence and uniqueness of the Mood-Brown line estimate $(\hat{\beta}_n, \hat{\alpha}_n)$, and have obtained its limiting distribution. Under mild assumptions we have shown that $(\hat{\beta}_n, \hat{\alpha}_n)$ is an asymptotically normal and consistent estimator of the true line (β, α) under the null hypothesis, and of a unique line (β_0, α_0) determined by the convex function ϕ (which is the true regression curve) and the distribution F under the alternative hypothesis.

We close this section with the remark that although we have assumed heretofore that the function $h(t)$ is strictly monotonic, thus restricting the number of Y observations at a fixed X to be at most one, this assumption is not essential, and the results go through for non-strictly monotonic $h(t)$ with only minor modifications.

3. A test of linearity versus convexity. The R test is a statistical test of the null hypothesis

$$H_0: Y_i = \alpha + \beta X_i + \epsilon_i, \qquad i = 0, 1, \cdots, n,$$

against the alternative

$$H_1: Y_i = \phi(X_i) + \epsilon_i, \qquad i = 0, 1, \cdots, n,$$

where $\phi(x)$ is strictly convex. The parameters α and β, and the function ϕ, are unspecified. We assume that the ϵ_i are independent identically distributed random variables with common density $f(\epsilon)$ and cumulative distribution function $F(x) = \int_{-\infty}^{x} f(\epsilon)\, d\epsilon$ such that $f(0) \neq 0$, $F(0) = \frac{1}{2}$, and f is continuous. The X_i are fixed and defined by the relation $X_i = h(i/n)$, where $h(t)$ is continuous and strictly monotonically increasing for $0 \leqq t \leqq 1$, $h(0) = c$, and $h(1) = d$.

The form of the R test is as follows: To the points in some subinterval $[c_1, d_1]$ of $[c, d]$ a straight line is fitted by the Mood-Brown procedure discussed in the previous section. Corresponding to each of the remaining points indicator variables are defined, taking on the value $+1$ if the point is above the estimated line, and 0 otherwise. A weighted sum of the indicators, R, is used as the test statistic.

Against the alternative H_1 the rejection criterion is $R > r_\theta$, where r_θ is chosen to make the test of level θ.

The logic underlying the test depends upon the fact that the line joining two points of a strictly convex function lies above the function in the interval between the abscissae of the points, and lies below elsewhere. Now we have seen in the previous section that when the convex function $\phi(x)$ is the true regression curve, the Mood-Brown estimate $(\hat{\beta}_n, \hat{\alpha}_n)$ converges to a line $(\beta_0(\phi), \alpha_0(\phi))$ satisfying equations (2.3.1). But it is evident from these equations that if $[c_1, d_1]$ is the interval used for line estimation, then $(\beta_0(\phi), \alpha_0(\phi))$ must intersect $\phi(x)$ in two points with abscissae in this interval. Hence for large n the line $(\hat{\beta}_n, \hat{\alpha}_n)$ will tend to lie below the curve $\phi(x)$ in the intervals $c \leqq x \leqq c_1$ and $d_1 \leqq x \leqq d$; and so if X_i lies in either of these intervals then Y_i will tend to have probability greater than $\frac{1}{2}$ of lying above $(\hat{\beta}_n, \hat{\alpha}_n)$. Consequently $R = R_n$ will tend to be large.

In this section we shall obtain the asymptotic distribution of R_n under the null and alternative hypotheses, and give an asymptotic expression for the power of the test.

3.1. *Definition of the R test.* The line $(\hat{\beta}_n, \hat{\alpha}_n)$ is obtained by the Mood-Brown procedure from the points (X_i, Y_i) with $h(\delta_1) = c_1 \leqq X_i \leqq d_1 = h(\delta_2)$, where $0 \leqq \delta_1 \leqq \delta_2 \leqq 1$. Corresponding to each of the remaining points we define the indicators

$$Z_i = Z_i(\hat{\beta}_n, \hat{\alpha}_n) = \begin{cases} +1 & \text{if} \quad Y_i > \hat{\alpha}_n + \hat{\beta}_n X_i \\ 0 & \text{if not, where} \quad X_i < c_1 \text{ or } X_i > d_1 . \end{cases}$$

The test statistic is defined to be

$$R_n = R_n(\hat{\beta}_n, \hat{\alpha}_n) = \sum_{i \varepsilon \mathcal{A}} a_i Z_i(\hat{\beta}_n, \hat{\alpha}_n),$$

where $\mathcal{A} = \{i : X_i < c_1 \text{ or } X_i > d_1\}$, $a(t)$ is a given weighting function defined for $0 \leqq t \leqq 1$, and $a_i = a(i/n)$. The test of H_0 against H_1 at the level θ is to reject H_0 if $R_n > r_\theta^{(n)}$, where $\Pr\{R_n > r_\theta^{(n)} \mid H_0\} = \theta$.

3.2. *Asymptotic distribution of R_n under H_0.* We make the same assumptions as in Theorem 2 concerning the function $h(t)$ and the distribution of the ϵ_i. Before considering the asymptotic distribution of R_n we shall state a modified version of the conclusion of Theorem 2, where the modification is necessary to allow for the fact that only the points (X_i, Y_i) with $h(\delta_1) \leqq X_i \leqq h(\delta_2)$ are now being used to obtain the line $(\hat{\beta}_n, \hat{\alpha}_n)$. Here we define $M_n^* =$ median $[X_i]$ over the set of X_i with $h(\delta_1) \leqq X_i \leqq h(\delta_2)$.

Assuming $Y_i = \alpha + \beta X_i + \epsilon_i$ we define $(\hat{\eta}_n, \hat{\xi}_n) = m^{\frac{1}{2}}(\hat{\beta}_n - \beta, \hat{\alpha}_n - \alpha)$, where $m = (n + 1)\frac{1}{2}(\delta_2 - \delta_1)$. We note that asymptotically the number of points now used for line estimation is $2m$. Then $(\hat{\eta}_n, \hat{\xi}_n)$ is asymptotically normally distributed with mean (0, 0) and covariance matrix

$$\Sigma^* = \left(f(0)\left[\int_{\bar\delta}^{\delta_2} h(t)\,dt - \int_{\delta_1}^{\bar\delta} h(t)\,dt\right]\right)^{-2}$$

$$(3.2.1)\qquad \cdot\begin{pmatrix} \frac{1}{8}(\delta_2-\delta_1)^2 & -\frac{1}{8}(\delta_2-\delta_1)\int_{\delta_1}^{\delta_2} h(t)\,dt \\ -\frac{1}{8}(\delta_2-\delta_1)\int_{\delta_1}^{\delta_2} h(t)\,dt & \frac{1}{4}\left[\left(\int_{\delta_1}^{\bar\delta} h(t)\,dt\right)^2 + \left(\int_{\bar\delta}^{\delta_2} h(t)\,dt\right)^2\right] \end{pmatrix}$$

where $\bar\delta = \frac{1}{2}(\delta_1 + \delta_2)$.

We shall now derive the asymptotic distribution of R_n. Under H_0 we have $Y_i = \alpha + \beta X_i + \epsilon_i$, $i = 0, 1, \cdots, n$. Without loss of generality we may take $\alpha = \beta = 0$. For let $Y_i' = Y_i - \alpha - \beta X_i$, and $(\hat\beta_n', \hat\alpha_n')$, Z_i', and R_n' be the statistics corresponding to the (unobservable) sample (X_i, Y_i'). We readily obtain $Z_i' = 1$ if and only if $Z_i = 1$, and so $R_n' = R_n$. Hence it suffices to take $\alpha = \beta = 0$.

We shall now obtain the asymptotic mean and variance of R_n, given that $(\hat\eta_n, \hat\xi_n) = (b, a)$. We have

$$E[R_n \mid (\hat\eta_n, \hat\xi_n) = (b, a)] = E\,[\sum_{i\varepsilon\mathfrak{a}} a_i Z_i(\hat\beta_n, \hat\alpha_n) \mid (\hat\eta_n, \hat\xi_n) = (b, a)]$$

$$= E\sum_{i\varepsilon\mathfrak{a}} a_i Z_i\left(\frac{b}{m^{\frac{1}{2}}}, \frac{a}{m^{\frac{1}{2}}}\right) = \sum_{i\varepsilon\mathfrak{a}} a_i\left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} X_i\right)\right]$$

$$\sim n\left[\int_0^{\delta_1}\left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} h(t)\right)\right] a(t)\,dt + \int_{\delta_2}^{1}\left[1 - F\left(\frac{a}{m^{\frac{1}{2}}} + \frac{b}{m^{\frac{1}{2}}} h(t)\right)\right] a(t)\,dt\right]$$

$$\sim n\left[\int_0^{\delta_1}\left(\frac{1}{2} - (a/m^{\frac{1}{2}})f(0) - (b/m^{\frac{1}{2}})f(0)h(t)\right) a(t)\,dt + \int_{\delta_2}^{1}\left(\frac{1}{2} - (a/m^{\frac{1}{2}})f(0) - (b/m^{\frac{1}{2}})f(0)h(t)\right) a(t)\,dt\right]$$

$$\equiv V_n(b, a) = \frac{1}{2} n\left[\int_0^{\delta_1} a(t)\,dt + \int_{\delta_2}^{1} a(t)\,dt\right] + \mu_n(b, a).$$

Similarly,

$$\text{Var}\,[R_n \mid (\hat\eta_n, \hat\xi_n) = (b, a)] \sim \frac{1}{4} n\left[\int_0^{\delta_1} a^2(t)\,dt + \int_{\delta_2}^{1} a^2(t)\,dt\right] \equiv \sigma_n^2.$$

Now let $a(t)$ be continuous and not identically zero on at least one of the sets $0 \leqq t \leqq \delta_1$ and $\delta_2 \leqq t \leqq 1$. It follows immediately from Lemma 4 of the Appendix that the conditional distribution of R_n, given that $(\hat\eta_n, \hat\xi_n) = (b, a)$, is asymptotically normal with mean $V_n(b, a)$ and variance σ_n^2 defined above.

We now proceed to find the unconditional distribution of R_n. Let $\tau_n = V_n(b, a) - \mu_n(b, a) = \frac{1}{2}n[\int_0^{\delta_1} a(t)\,dt + \int_{\delta_2}^1 a(t)\,dt]$, B be a bounded rectangle set in the plane, $k_n(b, a)$ be the density function of $(\hat{\eta}_n, \hat{\xi}_n)$ evaluated at (b, a), and $k(b, a) = \lim_{n\to\infty} k_n(b, a)$ be the limiting bivariate normal density of $(\hat{\eta}_n, \hat{\xi}_n)$. It can be shown that $k_n(b, a) < K$, constant, for sufficiently large n and $(b, a) \varepsilon B$.

It follows from the Dominated Convergence Theorem that for each r we have

$$\begin{aligned}\lim_{n\to\infty} \int_B \Pr\left\{\frac{R_n - \tau_n}{\sigma_n} < r \,\middle|\, (\hat{\eta}_n, \hat{\xi}_n) = (b, a)\right\} k_n(b, a)\, db\, da \\ = \lim_{n\to\infty} \int_B \Pr\left\{\frac{R_n - V_n(b, a)}{\sigma_n} < r - \frac{\mu_n(b, a)}{\sigma_n} \,\middle|\, (\hat{\eta}_n, \hat{\xi}_n) = (b, a)\right\} \\ \cdot k_n(b, a)\, db\, da \\ = \int_B \Phi(r - \mu(b, a)) k(b, a)\, db\, da,\end{aligned} \tag{3.2.2}$$

where

$$\mu(b, a) = \lim_{n\to\infty}\left[\frac{\mu_n(b, a)}{\sigma_n}\right]$$

$$= \frac{-f(0)}{\left(\frac{\delta_2 - \delta_1}{2}\right)^{\frac{1}{2}}} \frac{\left[a\left(\int_0^{\delta_1} a(t)\, dt + \int_{\delta_2}^1 a(t)\, dt\right) + b\left(\int_0^{\delta_1} a(t)h(t)\, dt + \int_{\delta_2}^1 a(t)h(t)\, dt\right)\right]}{\frac{1}{2}\left(\int_0^{\delta_1} a^2(t)\, dt + \int_{\delta_2}^1 a^2(t)\, dt\right)^{\frac{1}{2}}},$$

and Φ is the cumulative distribution function of a standardized normal random variable.

We shall now show

$$\lim_{n\to\infty} \Pr\{[(R_n - \tau_n)/\sigma_n] < r\} = \int \Phi(r - \mu(b, a)) k(b, a)\, db\, da.$$

Let $\{B_i\}$ be a monotone sequence of bounded rectangle sets converging to the entire plane. Then

$$\begin{aligned}\lim_{n\to\infty} \Pr\left\{\frac{R_n - \tau_n}{\sigma_n} < r\right\} &= \lim_{n\to\infty}\lim_{i\to\infty} \Pr\left\{\frac{R_n - \tau_n}{\sigma_n} < r \,\middle|\, (\hat{\eta}_n, \hat{\xi}_n) \varepsilon B_i\right\} \\ &= \lim_{n\to\infty}\lim_{i\to\infty} \int_{B_i} \Pr\left\{\frac{R_n - \tau_n}{\sigma_n} < r \,\middle|\, (\hat{\eta}_n, \hat{\xi}_n) = (b, a)\right\} k_n(b, a)\, db\, da \\ &= \lim_{i\to\infty}\lim_{n\to\infty} \int_{B_i} \Pr\left\{\frac{R_n - \tau_n}{\sigma_n} < r \,\middle|\, (\hat{\eta}_n, \hat{\xi}_n) = (b, a)\right\} k_n(b, a)\, db\, da \\ &= \lim_{i\to\infty} \int_{B_i} \Phi(r - \mu(b, a))\, k(b, a)\, db\, da = \int \Phi(r - \mu(b, a)) k(b, a)\, db\, da.\end{aligned}$$

The only step in these equations which is not immediate is the interchanging of the limits with respect to n and i, and this step can be verified in a straightforward fashion.

Now consider abstract random variables R, B, and A, having a trivariate normal distribution for which the marginal density of (B, A) is $k(b, a)$, and for which the conditional distribution of R, given that $(B, A) = (b, a)$, is normal with mean $\mu(b, a)$, variance one, and is otherwise independent of (B, A). Then

$$\Pr\{R < r\} = \int \Phi(r - \mu(b, a))k(b, a)\, db\, da.$$

But clearly R has the same distribution as the random variable $R' = \mu(B, A) + \omega$, where ω is normally distributed with mean 0, variance 1, and is independent of (B, A). Hence $ER = ER' = E\mu(B, A)$ and $\operatorname{Var} R = \operatorname{Var} R' = \operatorname{Var} \mu(B, A) + 1$.

From this discussion it follows that $(R_n - \tau_n)/\sigma_n$ is asymptotically normally distributed with mean $E\mu(\hat{\eta}_n, \hat{\xi}_n) = 0$ and variance $= \operatorname{Var}[\mu(\hat{\eta}_n, \hat{\xi}_n)] + 1$. Evaluating $\operatorname{Var}[\mu(\hat{\eta}_n, \hat{\xi}_n)]$ we thus obtain

THEOREM 4. *Let $a(t)$ be continuous and not identically zero on at least one of the sets $0 \leqq t \leqq \delta_1$ and $\delta_2 \leqq t \leqq 1$. Assume the hypothesis of Theorem 2 concerning $h(t)$ and the distribution of the ϵ_i. Then under H_0, R_n has an asymptotically normal distribution with mean τ_n and variance s_n^2, where*

$$\tau_n = \frac{1}{2} n \left[\int_0^{\delta_1} a(t)\, dt + \int_{\delta_2}^1 a(t)\, dt \right],$$

and

$$\begin{aligned} s_n^2 = {} & \frac{2nf^2(0)}{\delta_2 - \delta_1} \left(\int_0^{\delta_1} a(t)\, dt + \int_{\delta_2}^1 a(t)\, dt \right)^2 \operatorname{Var} \hat{\xi}_n \\ & + \frac{2nf^2(0)}{\delta_2 - \delta_1} \left(\int_0^{\delta_1} h(t)a(t)\, dt + \int_{\delta_2}^1 h(t)a(t)\, dt \right)^2 \operatorname{Var} \hat{\eta}_n \\ & + \frac{4nf^2(0)}{\delta_2 - \delta_1} \left(\int_0^{\delta_1} a(t)\, dt + \int_{\delta_2}^1 a(t)\, dt \right) \left(\int_0^{\delta_1} h(t)a(t)\, dt \right. \\ & \left. + \int_{\delta_2}^1 h(t)a(t)\, dt \right) \operatorname{Cov}(\hat{\xi}_n, \hat{\eta}_n) + \frac{1}{4} n \left[\int_0^{\delta_1} a^2(t)\, dt + \int_{\delta_2}^1 a^2(t)\, dt \right]. \end{aligned}$$

Here $(\hat{\eta}_n, \hat{\xi}_n) = [\frac{1}{2}(n + 1)(\delta_2 - \delta_1)]^{\frac{1}{2}}(\hat{\beta}_n - \beta, \hat{\alpha}_n - \alpha)$, and the covariance matrix of $(\hat{\eta}_n, \hat{\xi}_n)$ is given by (3.2.1).

A special case of interest occurs when $h(t) = c + (d - c)t$, $\delta_1 = 1 - \delta_2 = \delta$, and $a(t) = a(1 - t)$, where $0 < \delta < \frac{1}{2}$, $0 \leqq t \leqq 1$. In this case equal weights $a_i = a(i/n) = a(1 - (i/n)) = a_{n-i}$ are given to the indicators Z_i and Z_{n-i} whose abscissae are equidistant from the median

$$M_n = \underset{\{i=0,1,\dots,n\}}{\operatorname{median}} [X_i] \sim M_n^* = \underset{\{i : h(\delta_1) \leq X_i \leq h(\delta_2)\}}{\operatorname{median}} [X_i] \sim \tfrac{1}{2}(c + d).$$

From Theorem 4 we then obtain that the distribution of R_n is asymptotically normal with mean and variance

$$ER_n \sim n \int_0^\delta a(t)\, dt,$$

$$\text{Var } R_n \sim [n/(1 - 2\delta)] \left(\int_0^\delta a(t)\, dt \right)^2 + \frac{1}{2} n \int_0^\delta a^2(t)\, dt.$$

3.3. *Asymptotic distribution of R_n under the alternative.* We shall here state a result concerning the asymptotic distribution of R_n when the true regression curve is the convex function $\phi(x)$. The proof is very similar to the proof of Theorem 4, and will not be presented.

THEOREM 5. *Assume the hypothesis of Theorem 3 concerning the functions $\phi(x)$ and $h(t)$. Let the ϵ_i be independent identically distributed random variables with common density $f(\epsilon)$ and cumulative distribution function $F(x) = \int_{-\infty}^x f(\epsilon)\, d\epsilon$ such that $F(0) = \frac{1}{2}$, $f(0) \neq 0$, f is continuous, and*

$$F[\sup_{\delta_1 \leq t \leq \delta_2} (\alpha_0 + \beta_0 h(t) - \phi[h(t)])] < 1,$$

$$F[\inf_{\delta_1 \leq t \leq \delta_2} (\alpha_0 + \beta_0 h(t) - \phi[h(t)])] > 0,$$

where $(\beta_0, \alpha_0) = (\beta_0(\phi), \alpha_0(\phi))$ is the unique solution of the equations

$$\int_{\delta_1}^{(\delta_1+\delta_2)/2} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]\, dt = \int_{\frac{1}{2}(\delta_1+\delta_2)}^{\delta_2} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]\, dt = \frac{1}{4}(\delta_2 - \delta_1). \tag{3.3.1}$$

Then R_n is asymptotically normally distributed with mean

$$\tau_n(\phi) = n\left[\int_0^{\delta_1} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]\, a(t)\, dt + \int_{\delta_2}^1 [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] a(t)\, dt\right]$$

and variance $s_n^2(\phi) = O(n)$. Here $s_n^2(\phi)$ can be expressed by an equation analogous to that of Theorem 4.

3.4. *The asymptotic power of the test.* In this subsection we shall obtain an expression for the asymptotic power of the R test when the true regression curve is the convex function $\phi(x)$.

In Theorem 4 we have shown that under H_0, R_n is asymptotically normally distributed with mean $\tau_n = \frac{1}{2}n[\int_0^{\delta_1} a(t)\, dt + \int_{\delta_2}^1 a(t)\, dt]$ and variance s_n^2. We see that under H_0 the asymptotic distribution of R_n depends only upon $h(t)$, $a(t)$, δ_1, δ_2, and n. Letting r_θ be the upper $(1 - \theta)$ percentile of the standardized normal distribution, we have $\Pr\{R_n > \tau_n + s_n r_\theta \mid H_0\} \sim \theta$, and so for "large" n the level θ test of H_0 against H_1 is to reject H_0 when $R_n > r_\theta^{(n)} = \tau_n + s_n r_\theta$.

The asymptotic power of the R test when the convex function ϕ is the true regression curve is then given by

$$\Pr\{R_n > r_\theta^{(n)} \mid \phi\} = \Pr\left\{\frac{R_n - \tau_n(\phi)}{s_n(\phi)} > \frac{r_\theta^{(n)} - \tau_n(\phi)}{s_n(\phi)} \middle| \phi\right\}$$
$$\sim 1 - \Phi\left(\frac{r_\theta^{(n)} - \tau_n(\phi)}{s_n(\phi)}\right)$$

where Φ is the cumulative distribution function of the standardized normal distribution, and $\tau_n(\phi)$ and $s_n^2(\phi)$ are defined in Theorem 5. Here

$$[r_\theta^{(n)} - \tau_n(\phi)] \sim \frac{n}{2}\left(\int_0^{\delta_1} a(t)\,dt + \int_{\delta_2}^1 a(t)\,dt\right)$$
$$+ r_\theta s_n - n\left[\int_0^{\delta_1} [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]a(t)\,dt\right.$$
$$\left. + \int_{\delta_2}^1 [1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])]a(t)\,dt\right]$$

Now it was remarked at the beginning of Section 3 that the line (β_0, α_0) necessarily lies strictly below the function $\phi(x)$ for $c \leqq x \leqq c_1$ and $d_1 \leqq x \leqq d$, or equivalently, $\alpha_0 + \beta_0 h(t) - \phi[h(t)] < 0$ for $0 \leqq t \leqq \delta_1$ and $\delta_2 \leqq t \leqq 1$. Consequently $1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)]) > \frac{1}{2}$ in this range of t, and since $s_n(\phi) = 0(n^{\frac{1}{2}})$, we have $[(r_\theta^{(n)} - \tau_n(\phi))/s_n(\phi)] \sim n^{\frac{1}{2}}G$, where $G < 0$. Hence $\Pr\{R_n > r_\theta^{(n)} \mid \phi\} \sim 1 - \Phi(n^{\frac{1}{2}}G)$, tending to 1 as n tends to ∞, and it follows that the R test is consistent against all alternatives satisfying the hypothesis of Theorem 5. Here

$$G = G(\delta_1, \delta_2, \phi, a(t), h(t), F)$$
$$= \left[\int_0^{\delta_1}\left[F(\alpha_0 + \beta_0 h(t) - \phi[h(t)]) - \frac{1}{2}\right]a(t)\,dt\right.$$
$$\left. + \int_{\delta_2}^1 \left[F(\alpha_0 + \beta_0 h(t) - \phi[h(t)]) - \frac{1}{2}\right]a(t)\,dt\right](s_n^2(\phi)/n)^{-\frac{1}{2}}.$$

In the following section we shall be particularly interested in the case where $h(t) = c + (d - c)t$, $a(t) = a(1 - t)$, $\delta_1 = \delta$, $\delta_2 = 1 - \delta$, and $0 \leqq t \leqq 1$, $0 \leqq \delta \leqq \frac{1}{2}$. We note that in this case

$$r_\theta^{(n)} = n\int_0^\delta a(t)\,dt + \left\{[n/(1 - 2\delta)]\left[\int_0^\delta a(t)\,dt\right]^2 + \frac{1}{2}n\int_0^\delta a^2(t)\,dt\right\}^{\frac{1}{2}} r_\theta,$$

which reduces further to $r_\theta^{(n)} = n\delta + (n\delta/2(1 - 2\delta))^{\frac{1}{2}} r_\theta$ when $a(t) \equiv 1$.

4. Efficiency and use of the R test. In this section we obtain the relative asymptotic efficiency of the R test as compared to the least squares test for a sequence of parabolic alternatives converging to a line. Errors are here assumed normally distributed with common variance. In the final subsection we make recommendations for the use of the R test.

4.1. *Relative asymptotic efficiency.* In this subsection we shall obtain the relative asymptotic efficiency of the proposed test of convexity as compared with the least squares test when errors are independently normally distributed with common variance. Under such assumptions it is known that the least squares method yields normally distributed minimum variance linear unbiased estimates of the parameters of a linear regression scheme, and the least squares test (here identical with the likelihood ratio test) for values of the parameters has certain optimal properties. For discussions of the least squares test and of the notion of relative asymptotic efficiency we refer to [7] and [3], respectively.

Since the general convex alternative $\phi(x)$ of the R test cannot be expressed as a linear function of a finite set of parameters, in order to make comparisons we shall restrict ourselves to quadratic alternatives $\phi(x) = \gamma x^2$ in the present discussion. Specifically, we shall then be comparing the R test of the null hypothesis of linearity against the alternative of convexity with the least squares test of the null hypothesis $\gamma = 0$ against the alternative $\gamma > 0$.

We shall take $X_i = h(i/n)$ as in the previous sections, and (without loss of generality) take the common variance of the ϵ_i to be one. The asymptotic efficiency will then be obtained for the sequence $\phi_n(x) = \gamma n^{-\frac{1}{2}} x^2$ of alternatives converging to the line $y = 0$.

The least squares level θ test of the null hypothesis $\gamma = 0$ against the alternative $\gamma > 0$ (assuming it is known that $\sigma^2 = 1$) is to reject if

$$\hat{\gamma}_n > E(\hat{\gamma}_n \mid H_0) + r_\theta[\text{Var}\,(\hat{\gamma}_n \mid H_0)]^{\frac{1}{2}},$$

where $\hat{\gamma}_n$ is the least squares estimate of γ, and r_θ is the upper $(1 - \theta)$ percentile of the standardized normal distribution $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp\,(-\frac{1}{2}t^2)\, dt$. We readily obtain $E(\hat{\gamma}_n) = \gamma$, and

$$\text{Var}\,(\hat{\gamma}_n \mid H_0) = \frac{(n+1) \sum_{i=0}^{n} (X_i - \bar{X})^2}{(n+1)\left[\sum_{i=0}^{n} (X_i - \bar{X})^2 \sum_{i=0}^{n} (X_i - \bar{X})^4 - \left(\sum_{i=0}^{n} (X_i - \bar{X})^3\right)^2\right] - \left(\sum_{i=0}^{n} (X_i - \bar{X})^2\right)^3},$$

where $\bar{X} = (n+1)^{-1} \sum_{i=0}^{n} X_i$.

From [3] it is easy to verify that the relative asymptotic efficiency of the R test as compared to the least squares test for the sequence $\{\phi_n(x) = n^{-\frac{1}{2}} \gamma x^2\}$ is given by the equation

$$E = \lim_{n \to \infty} \left[\frac{(d/d\gamma) E_\gamma\{R_n\} \mid \gamma = 0}{(d/d\gamma) E_\gamma\{\hat{\gamma}_n\} \mid \gamma = 0}\right]^2 \frac{\text{Var}\,[\hat{\gamma}_n \mid \gamma = 0]}{\text{Var}\,[R_n \mid \gamma = 0]}.$$

From Theorem 5 R_n is asymptotically normally distributed with mean

$$\tau_n(\phi) = n\left[\int_0^{\delta_1} [1 - \Phi(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] a(t)\, dt + \int_{\delta_2}^{1} [1 - \Phi(\alpha_0 + \beta_0 h(t) - \phi[h(t)])] a(t)\, dt\right],$$

where $\beta_0 = \beta_0(\gamma)$ and $\alpha_0 = \alpha_0(\gamma)$ are defined by equations (3.3.1), and $\phi(x) = \gamma x^2$. Hence

$$(d/d\gamma)E_\gamma[R_n] \sim -n(2\pi)^{-\frac{1}{2}} \Bigg[\int_0^{\delta_1} \exp\left[-\frac{1}{2}(\alpha_0(\gamma) + \beta_0(\gamma)h(t) - \gamma h^2(t))^2\right] \cdot (\alpha_0'(\gamma) + \beta_0'(\gamma)h(t) - h^2(t))a(t)\,dt + \int_{\delta_2}^{1} \exp\left[-\frac{1}{2}(\alpha_0(\gamma) + \beta_0(\gamma)h(t) - \gamma h^2(t))^2\right] \cdot (\alpha_0'(\gamma) + \beta_0'(\gamma)h(t) - h^2(t))a(t)\,dt\Bigg],$$

and so

$$(d/d\gamma)E_\gamma\{R_n\} \mid \gamma = 0 \sim -n(2\pi)^{-\frac{1}{2}} \Bigg[\int_0^{\delta_1} (\alpha_0'(0) + \beta_0'(0)h(t) - h^2(t))a(t)\,dt + \int_{\delta_2}^{1} (\alpha_0'(0) + \beta_0'(0)h(t) - h^2(t))a(t)\,dt\Bigg],$$

because $\alpha_0(0) = \beta_0(0) = 0$. Since $[(d/d\gamma)E_\gamma\{\hat{\gamma}_n\} \mid \gamma = 0] = 1$, we thus have

$$E = \lim_{n\to\infty}\Bigg[n(2\pi)^{-1/2}\Bigg(\int_0^{\delta_1} [\alpha_0'(0) + \beta_0'(0)h(t) - h^2(t)]a(t)\,dt + \int_{\delta_2}^{1} [\alpha_0'(0) + \beta_0'(0)h(t) - h^2(t)]a(t)\,dt\Bigg)\Bigg]^2 \frac{\operatorname{Var}[\hat{\gamma}_n \mid \gamma = 0]}{\operatorname{Var}[R_n \mid \gamma = 0]}.$$

In deriving explicit values of E we shall make the further assumptions: $h(t) = c + (d - c)t$, $\delta_1 = \delta$, $\delta_2 = 1 - \delta$, and $a(t) = a(1 - t)$, where $0 \leqq t \leqq 1$, $0 < \delta < \frac{1}{2}$. In this case we have

$$\operatorname{Var}[R_n \mid \gamma = 0] \sim [n/(1 - 2\delta)]\left(\int_0^{\delta} a(t)\,dt\right)^2 + \tfrac{1}{2}n\left(\int_0^{\delta} a^2(t)\,dt\right),$$

and

$$\operatorname{Var}[\hat{\gamma}_n \mid \gamma = 0] \sim [180/n(d - c)^4].$$

Next we shall obtain $\alpha_0'(0)$ and $\beta_0'(0)$. We have pointed out earlier that $\beta_0(\gamma) = (c + d)\gamma$ when $\phi(x) = \gamma x^2$. Hence $\beta_0'(0) = c + d$. To obtain $\alpha_0(\gamma)$ it is necessary then only to solve the equation

$$\int_\delta^{1/2} [1 - \Phi(\alpha_0(\gamma) + (c + d)\gamma h(t) - \gamma h^2(t))]\,dt = \tfrac{1}{4}(1 - 2\delta).$$

Differentiating with respect to γ yields

$$\int_\delta^{1/2} \exp[-\tfrac{1}{2}(\alpha_0(\gamma) + (c + d)\gamma h(t) - \gamma h^2(t))^2] \cdot (\alpha_0'(\gamma) + (c + d)h(t) - h^2(t))\,dt = 0,$$

from which we obtain $\alpha_0'(0) = \frac{1}{3}[(d - c)^2(\frac{1}{2} - \delta)^2] - \frac{1}{4}(c + d)^2$. After some algebraic manipulation we finally obtain

$$E = (180/\pi)\Big([\tfrac{2}{3}(\tfrac{1}{2} - \delta)^2 - \tfrac{1}{2}]\int_0^\delta a(t)\,dt + 2\int_0^\delta ta(t)\,dt - 2 \cdot \int_0^\delta a(t)t^2\,dt\Big)^2\Big[(1 - 2\delta)^{-1}\Big(\int_0^\delta a(t)\,dt\Big)^2 + \tfrac{1}{2}\int_0^\delta a^2(t)\,dt\Big]^{-1}. \tag{4.1.1}$$

We note that E does not depend upon the interval $[c, d]$ in which the X_i fall.

It is easy to show that for the particular weighting function $a(t) \equiv 1$, the efficiency E takes on its maximum with respect to δ for $\delta = (7 - 17^{\frac{1}{2}})/16 \approx \frac{3}{16}$ and this maximal efficiency is $E \approx 1.5/\pi$. Under the restrictions $a(t) = a(1 - t)$ and $\delta_1 = 1 - \delta_2 = \delta$ it is not difficult to show that the optimal weighting function using the criterion of relative asymptotic efficiency is quadratic of the form $a(t) = K_1 + \frac{1}{2}(t - 1)^2$, where $K_1 < 0$; and using the optimal δ for this function $a(t)$ we have obtained by numerical computations an efficiency of .636 which appears to be an approximation to $2/\pi$.

4.2. *Use of the R test.* Here we shall discuss briefly some of the considerations involved in performing the R test.

From 3.4 we have $\Pr\{R_n > r_\theta^{(n)} \mid H_0\} \sim \theta$, and it follows that for "large" n the level θ test of H_0 against convex alternatives is to reject H_0 when $R_n > r_\theta^{(n)}$. For "small" values of n the use of this criterion will alter the significance level of the test, but can nevertheless be instructive. In the future, tables of the exact null distribution of R_n will be given for small values of n and certain distributions of the ϵ_i. During the remainder of this discussion we shall assume, however, that n is large enough for use of the asymptotic criterion $R_n > r_\theta^{(n)}$. (In Appendix 5.2 the small sample behavior of R_n will be discussed.)

Now the R test is defined in terms of a spacing function $h(t)$, a weighting function $a(t)$, and parameters δ_1 and δ_2 which determine the location and proportion of observations to be used for estimating the line $(\hat{\beta}_n, \hat{\alpha}_n)$.

The function $h(t)$ giving the location of the X_i may or may not be at the choice of the statistician. When it is, the criterion of simplicity would suggest that $h(t)$ be chosen linear, corresponding to equally spaced X_i in an interval. Assuming $a(t) = a(1 - t)$ and $\delta_1 = 1 - \delta_2 = \delta$, where $0 \leqq t \leqq 1$ and $0 < \delta < \frac{1}{2}$, we then have the simple rejection criterion

$$R_n > r_\theta^{(n)} = n\int_0^\delta a(t)\,dt + r_\theta\Big[(n/(1 - 2\delta))\Big(\int_0^\delta a(t)\,dt\Big)^2 + \tfrac{1}{2}n\int_0^\delta a^2(t)\,dt\Big]^{1/2}.$$

When $h(t)$ cannot be chosen by the statistician, but is nevertheless a known function (in other words the X_i fall in a known pattern), the theory of the previous sections applies, and $r_\theta^{(n)}$ can be determined from Theorem 4 of Section 3.

Finally, suppose that no spacing function $h(t)$ is known to the statistician and that the values $X_0 < X_1 < \cdots < X_n$ appear quite arbitrary. Since the asymptotic variance of R_n under H_0, and hence also the rejection value $r_\theta^{(n)}$, depends upon the unknown function $h(t)$, it follows that the test cannot yet be defined. Moreover, since there exist infinitely many continuous monotonic functions $h^*(t)$ such that $h^*(i/n) = X_i$, $i = 0, 1, \cdots, n$, any choice of a particular such function to use in obtaining the rejection value $r_\theta^{(n)}$ must be arbitrary. However, for large n such functions $h^*(t)$ will yield approximately the same rejection value $r_\theta^{(n)}$, and for simplicity of computation we recommend the use of the piecewise linear function $h^*(t)$ which satisfies $h^*(i/n) = X_i$, $i = 0, 1, \cdots, n$, and which is linear over the subintervals $[(i - 1)/n, (i/n)]$, $i = 1, \cdots, n$. Using this function $h^*(t)$ we can then obtain $r_\theta^{(n)}$ as before.

The choice of the weighting function $a(t)$ and the parameters δ_1 and δ_2 must be made on a subjective basis. In general we recommend taking $a(t) \equiv 1$ so that R_n is simply the number of points above the line $(\hat{\beta}_n, \hat{\alpha}_n)$ among those points not used in line estimation. The values $\delta_1 = .20$ and $\delta_2 = .80$ are also recommended, in part because they yield approximately maximal efficiency for the function $a(t) \equiv 1$ in the situation considered in 4.1. None of these choices are expected to be optimal in any general sense, however, and are recommended because they seem reasonable and simplify computations. In the future it is hoped that the use of Monte Carlo techniques will shed light on the optimal $a(t)$, δ_1, and δ_2, for particular kinds of alternatives in the small sample case.

When there are more than one Y observations, say $Y_{i1}, \cdots, Y_{ik_i}$, at each X_i, then slight modifications in the statement and proof of the theorems concerning the asymptotic distributions of R_n and $(\hat{\beta}_n, \hat{\alpha}_n)$ are necessary. This situation corresponds to a spacing function $h(t)$ which is constant within certain subintervals of $[0, 1]$.

We shall close this section with a brief mention of the computations involved in using the R test. For small samples, say $n < 20$, the line $(\hat{\beta}_n, \hat{\alpha}_n)$ can easily be determined with the use of a ruler from a plotting of the sample points on graph paper. R_n is then obtained by simply counting the number of points above $(\hat{\beta}_n, \hat{\alpha}_n)$ in the appropriate subintervals. For larger samples the only difficulty in computation will be in obtaining the line $(\hat{\beta}_n, \hat{\alpha}_n)$, and this can be done using an iterative procedure suggested by Mood in [5].

5. Appendix.

5.1. *A series of lemmas.* We shall here state and prove a series of lemmas referred to in the previous sections.

LEMMA 1. *Under the hypothesis of Theorem* 2, *and with* m *odd*,

$$g_n(b, a)\, db\, da = \sum_{(i,j)\in \mathfrak{a}} \Pr\{V_{1i}(b, a) = \tfrac{1}{2}(m - 1),\ V_{2j}(b, a) = \tfrac{1}{2}(m - 1)\}$$
$$\cdot f(a + bX_i) f(a + bX_j)\, d(a + bX_i)\, d(a + bX_j).$$

PROOF. Let (b_0, a_0) be a fixed line, let $B_\delta = \{(b, a) : |b - b_0| < \delta, |a - a_0| < \delta\}$, and let

$$B_{ij}(b, a) = \{V_{1i}(b, a) = \tfrac{1}{2}(m - 1), V_{2j}(b, a) = \tfrac{1}{2}(m - 1), Y_i = a + bX_i, Y_j = a + bX_j\}.$$

Then

$$\Pr\{(\hat{\beta}_n, \hat{\alpha}_n) \varepsilon B_\delta\} = \Pr \bigcup_{(i,j)\varepsilon \mathfrak{A}} \bigcup_{(b,a)\varepsilon B_\delta} B_{ij}(b, a).$$

This equation follows from the Corollary to Theorem 1, since with probability one we must have exactly one point (X_i, Y_i) with $X_i \leqq M_n$ lying on $(\hat{\beta}_n, \hat{\alpha}_n)$, and exactly one point (X_j, Y_j) with $X_j > M_n$ lying on $(\hat{\beta}_n, \hat{\alpha}_n)$.

Clearly we have

$$\Pr\{(\bigcup_{(b,a)\varepsilon B_\delta} B_{ij}(b, a)) \cap (\bigcup_{(b,a)\varepsilon B_\delta} B_{i'j'}(b, a))\} = 0$$

unless $i = i'$ and $j = j'$, because of the same Corollary. Hence

$$\Pr\{\hat{\beta}_n, \hat{\alpha}_n) \varepsilon B_\delta\} = \sum_{(i,j)\varepsilon \mathfrak{A}} \Pr \bigcup_{(b,a)\varepsilon B_\delta} B_{ij}(b, a).$$

We next obtain

$$\begin{aligned}
\Pr \bigcup_{(b,a)\varepsilon B_\delta} B_{ij}(b, a) &= \int_{B_\delta} \Pr\{\bigcup_{(\beta,\alpha)\varepsilon B_\delta} B_{ij}(\beta, \alpha) \mid Y_i = a + bX_i, Y_j = a + bX_j\} \\
&\qquad \cdot f(a + bX_i) f(a + bX_j)\, d(a + bX_i)\, d(a + bX_j) \\
&= \int_{B_\delta} \Pr\{B_{ij}(b, a) \mid Y_i = a + bX_i, Y_j = a + bX_j\} \\
&\qquad \cdot f(a + bX_i) f(a + bX_j)\, d(a + bX_i)\, d(a + bX_j) \\
&= \int_{B_\delta} \Pr\{V_{1i}(b, a) = \tfrac{1}{2}(m - 1), V_{2j}(b, a) = \tfrac{1}{2}(m - 1)\} \\
&\qquad \cdot f(a + bX_i) f(a + bX_j)\, d(a + bX_i)\, d(a + bX_j),
\end{aligned}$$

where the last step follows from the mutual independence of Y_i, Y_j, $V_{1i}(b, a)$ and $V_{2j}(b, a)$. Hence

$$\begin{aligned}
\Pr\{(\hat{\beta}_n, \hat{\alpha}_n)\varepsilon B_\delta\} &= \sum_{(i,j)\varepsilon \mathfrak{A}} \int_{B_\delta} \Pr\{V_{1i}(b, a) = \tfrac{1}{2}(m - 1), V_{2j}(b, a) \\
&= \tfrac{1}{2}(m - 1)\} f(a + bX_i) f(a + bX_j)\, d(a + bX_i)\, d(a + bX_j),
\end{aligned}$$

and letting δ go to zero, we obtain

$$\begin{aligned}
&g_n(b_0, a_0)\, db_0\, da_0 \\
&\quad = \sum_{(i,j)\varepsilon \mathfrak{A}} \Pr\{V_{1i}(b_0, a_0) = \tfrac{1}{2}(m - 1), V_{2j}(b_0, a_0) = \tfrac{1}{2}(m - 1)\} \\
&\qquad \cdot f(a_0 + b_0 X_i) f(a_0 + b_0 X_j)\, d(a_0 + b_0 X_i)\, d(a_0 + b_0 X_j)
\end{aligned}$$

as claimed.

We shall next prove four Lemmas which yield the asymptotic normality of certain statistics discussed earlier. These statistics are all of the form

$\sum_i a_i Z_i(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}})$, where

$$Z_i(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}}) = \begin{cases} +1 & \text{if} \quad Y_i > (a/m^{\frac{1}{2}}) + (b/m^{\frac{1}{2}})X_i, \\ 0 & \text{if not,} \end{cases}$$

the a_i are constants, and (b, a) is fixed. We note that, strictly speaking, two subscripts should be used in these expressions, since $X_i = h(i/n)$ depends upon both i and n. The four Lemmas will be stated for variables with double subscripts.

LEMMA 2. *Let Z^*_{nk}, $k = 1, \cdots, n$, be a sequence of independent random variables such that $EZ^*_{nk} = 0$ for all n and k. Let $Z^*_n = \sum_{k=1}^n Z^*_{nk}$ and $s_n = [\text{Var } Z^*_n]^{\frac{1}{2}}$. Then $\lim_{n\to\infty} s_n^{-3} \sum_{k=1}^n E|Z^*_{nk}|^3 = 0$ implies that Z^*_n/s_n is asymptotically normally distributed with mean zero and variance one.*

PROOF. Lemma 2 is a generalization of Liapounoff's Theorem to sequences of sums of independent random variables, and its proof is essentially the same as the proof of Liapounoff's Theorem given in [2].

LEMMA 3. *Let Z_{nk} be a sequence of independent random variables such that*

$$Z_{nk} = \begin{cases} 1 \text{ with probability } p_{nk} \\ 0 \text{ with probability } q_{nk} = 1 - p_{nk}, \end{cases}$$

and $Z^*_{nk} = a_{nk}(Z_{nk} - EZ_{nk})$, where $\{a_{nk}\}$ is a given set of constants such that $|a_{nk}| < A < \infty$ for all n and k, and such that there exists $\delta > 0$ for which the number of a_{nk} among $a_{n1}, \cdots, a_{nn}$ with $|a_{nk}| > \delta$ tends to ∞ as n tends to ∞. Then the condition $0 < \epsilon_1 < p_{nk} < 1 - \epsilon_2 < 1$ for some N, all $n > N$, and $k = 1, \cdots, n$, implies that *$Z^*_n/s_n = \sum_{k=1}^n Z^*_{nk}/s_n$ is asymptotically normal with mean zero and variance one. Here $s_n = [\text{Var } Z^*_n]^{\frac{1}{2}}$.*

PROOF. We have

$$s_n^{-3} \sum_{k=1}^n E|Z^*_{nk}|^3 = \sum_{k=1}^n |a_{nk}|^3 p_{nk} q_{nk}(p_{nk}^2 + q_{nk}^2)(s_n)^{-3} \leqq \sum_{k=1}^n |a_{nk}|^3 p_{nk} q_{nk}(s_n)^{-3}$$

$$\leqq A \sum_{k=1}^n a_{nk}^2 p_{nk} q_{nk} \left(\sum_{k=1}^n a_{nk}^2 p_{nk} q_{nk}\right)^{-3/2} = A \left(\sum_{k=1}^n a_{nk}^2 p_{nk} q_{nk}\right)^{-1/2}.$$

Hence by Lemma 2 it suffices to show $\sum_{k=1}^n a_{nk}^2 p_{nk} q_{nk}$ tends to ∞ as n tends to ∞, which follows immediately from the assumptions on a_{nk} and p_{nk}.

LEMMA 4. *Let Z_{nk}, Z^*_{nk}, Z^*_n, s_n and a_{nk} be defined as in Lemma 3. Let $h(t)$ be bounded for $0 \leqq t \leqq 1$, let F be a cumulative distribution function continuous at zero with $0 < F(0) < 1$, and let (b, a) be fixed. If*

$$p_{nk} = 1 - F[(a/n^{\frac{1}{2}}) + (b/n^{\frac{1}{2}})h(k/n)],$$

then Z^*_n/s_n is asymptotically normally distributed with mean zero and variance one.

PROOF. The proof follows immediately from Lemma 3.

REMARK. From Lemma 4 it is easy to obtain, under H_0, the asymptotic normality of the random variables

$$V_1[(b/m^{\frac{1}{2}}), (a/m^{\frac{1}{2}})] \quad \text{and} \quad V_2(b/m^{\frac{1}{2}}, a/m^{\frac{1}{2}})$$

discussed in Section 2, and of the conditional distribution of R_n, given that $(\hat{\eta}_n, \hat{\xi}_n) = (b, a)$, discussed in Section 3. As remarked earlier each of these variables has the form $\sum_i a_{ni}Z_{ni}[(b/m^{\frac{1}{2}}), (a/m^{\frac{1}{2}})]$, where

$$Z_{ni}[(b/m^{\frac{1}{2}}),(a/m^{\frac{1}{2}})] = \begin{cases} +1 & \text{if} \quad Y_{ni} > (a/m^{\frac{1}{2}}) + (b/m^{\frac{1}{2}})X_{ni} \\ 0 & \text{if not} \end{cases}$$

Hence

$$\Pr\{Z_{ni}((b/m^{\frac{1}{2}}), (a/m^{\frac{1}{2}})) = 1\}$$
$$= 1 - F((a/m^{\frac{1}{2}}) + (b/m^{\frac{1}{2}})X_{ni}) = 1 - F((a/m^{\frac{1}{2}}) + (b/m^{\frac{1}{2}})h(i/n))$$

under H_0. If $a_{nk} = a(k/n)$, where $a(t)$ is a continuous function for $0 \leqq t \leqq 1$ then we have $|a_{nk}| \leqq \sup_{0 \leqq t \leqq 1} |a(t)| < \infty$ for all n and k. Moreover, if $a(t)$ is not identically zero, then in some interval about a point t_0 for which $a(t_0) \neq 0$ we must have $|a(t)| > \delta > 0$ for some δ, and consequently the number of a_{nk} for which $|a_{nk}| > \delta$ tends to ∞ as n tends to ∞. Hence we may apply Lemma 4 to yield the desired result.

The asymptotic normality of these statistics under the alternative hypothesis follows readily from

LEMMA 5. *Let Z_{nk}, Z_{nk}^*, Z_n^*, s_n and a_{nk} be as in Lemma 3. Let $h(t)$ be bounded for $0 \leqq t \leqq 1$, let F be a continuous cumulative distribution function, and let ϕ be a convex function such that*

$$0 < \epsilon < 1 - F(\alpha_0 + \beta_0 h(t) - \phi[h(t)]) < 1 - \epsilon < 1 \qquad \text{for } 0 \leqq t \leqq 1,$$

where (β_0, α_0) is fixed. For constant (b, a) let

$$p_{nk} = 1 - F[\alpha_0 + (a/n^{\frac{1}{2}}) + [\beta_0 + (b/n^{\frac{1}{2}})]h(k/n) - \phi[h(k/n)]].$$

Then Z_n^*/s_n is asymptotically normal with mean zero and variance one.

PROOF. The proof follows immediately from Lemma 3.

Finally we shall prove

LEMMA 6. *Let F be a cumulative distribution function with $F(x) = \int_{-\infty}^{x} f(\epsilon)\, d\epsilon$ for all x, $F(0) = \frac{1}{2}$, $f(0) = F'(0) > 0$, and f continuous. Let $h(t)$ be continuous and strictly monotonically increasing for $0 \leqq t \leqq 1$. Then corresponding to each strictly convex twice differentiable function $\phi(x)$ there exists a unique solution $(\beta_0(\phi), \alpha_0(\phi)) = (\beta_0, \alpha_0)$ to the equations*

$$(*) \int_0^{1/2} [1 - F(\alpha_0 + \beta_0\, h(t) - \phi[h(t)])]\, dt$$

$$= \int_{1/2}^{1} [1 - F(\alpha_0 + \beta_0\, h(t) - \phi[h(t)])]\, dt = \tfrac{1}{4}.$$

PROOF. Let

$$\psi_1(\beta, \alpha) = \int_0^{\frac{1}{2}} F(\alpha + \beta h(t) - \phi[h(t)])\, dt,$$

$$\psi_2(\beta, \alpha) = \int_{\frac{1}{2}}^1 F(\alpha + \beta h(t) - \phi[h(t)])\, dt,$$

and $0 \leqq t_1 \leqq \frac{1}{2} \leqq t_2 \leqq 1$. Corresponding to any two such points t_1 and t_2 we associate the secant crossing $\phi(h)$ at the abscissae $h(t_1)$ and $h(t_2)$. Such a secant intersects $\phi[h(t)]$ in the points $U_1 = (h(t_1), \phi[h(t_1)])$ and $U_2 = (h(t_2), \phi[h(t_2)])$, and has the equation

$$\alpha(t_1, t_2) + \beta(t_1, t_2)h(t) = \begin{cases} \phi[h(t_1)] + \dfrac{\phi[h(t_2)] - \phi[h(t_1)]}{h(t_2) - h(t_1)}(h(t) - h(t_1)) & \text{if } (t_1, t_2) \neq \left(\dfrac{1}{2}, \dfrac{1}{2}\right) \\ \phi\left[h\left(\dfrac{1}{2}\right)\right] + \phi'\left[h\left(\dfrac{1}{2}\right)\right]\left(h(t) - h\left(\dfrac{1}{2}\right)\right) & \text{if } (t_1, t_2) = \left(\dfrac{1}{2}, \dfrac{1}{2}\right). \end{cases}$$

Using this line we define

$$\tilde{\psi}_1(t_1, t_2) = \psi_1(\beta(t_1, t_2), \alpha(t_1, t_2)) \quad \text{and} \quad \tilde{\psi}_2(t_1, t_2) = \psi_2(\beta(t_1, t_2), \alpha(t_1, t_2)).$$

Since $\alpha(t_1, \frac{1}{2}) + \beta(t_1, \frac{1}{2})h(t) < \phi[h(t)]$, for $\frac{1}{2} < t \leqq 1$, it follows that

$$\tilde{\psi}_2\left(t_1, \frac{1}{2}\right) = \int_{\frac{1}{2}}^1 F\left(\alpha\left(t_1, \frac{1}{2}\right) + \beta\left(t_1, \frac{1}{2}\right)h(t) - \phi[h(t)]\right) dt < \frac{1}{2}\int_{\frac{1}{2}}^1 dt = \frac{1}{4},$$

and similarly $\tilde{\psi}_2(t_1, 1) > \frac{1}{4}$, where $0 \leqq t_1 \leqq \frac{1}{2}$. Since $\tilde{\psi}_2(t_1, t_2)$ is continuous in t_2 for fixed t_1 it follows that corresponding to each t_1 with $0 \leqq t_1 \leqq \frac{1}{2}$, there exists a value $t_2 = g_2(t_1)$ such that $\tilde{\psi}_2(t_1, g_2(t_1)) = \frac{1}{4}$. Clearly $\frac{1}{2} < g_2(t_1) < 1$ for $0 \leqq t_1 \leqq \frac{1}{2}$. We shall now show that for each such t_1 the function $g_2(t_1)$ gives the unique solution t_2 of the equation $\tilde{\psi}_2(t_1, t_2) = \frac{1}{4}$, and is continuous.

It is easy to verify that

$$\frac{\partial\tilde{\psi}_2(t_1, t_2)}{\partial t_2} = \int_{\frac{1}{2}}^1 f(\alpha(t_1, t_2) + \beta(t_1, t_2)h(t) - \phi[h(t)]) \cdot(\partial/\partial t_2)(\alpha(t_1, t_2) + \beta(t_1, t_2)h(t) - \phi[h(t)])\, dt.$$

The first factor of the integrand equals $f(0) > 0$ when $t = t_2$, while the second factor is positive for all $0 \leqq t_1 \leqq \frac{1}{2} \leqq t_2 \leqq 1$ and $\frac{1}{2} \leqq t \leqq 1$ because of the strict convexity of $\phi(x)$ and the strict monotonicity of $h(t)$. It follows that $[\partial\tilde{\psi}_2(t_1, t_2)]/\partial t_2 > 0$ for such values of (t_1, t_2). From the Theorem of Implicit Functions we then obtain the uniqueness and continuity of the function $g_2(t_1)$ satisfying $\tilde{\psi}_2(t_1, g_2(t_1)) = \frac{1}{4}$ for $0 \leqq t_1 \leqq \frac{1}{2}$. Similarly we obtain a continuous function $g_1(t_2)$ which is the unique solution of the equation $\tilde{\psi}_1(g_1(t_2), t_2) = \frac{1}{4}$ for $\frac{1}{2} \leqq t_2 \leqq 1$, and is such that $0 < g_1(t_2) < \frac{1}{2}$ for $\frac{1}{2} \leqq t_2 \leqq 1$. Consideration of the functions $g_2(t_1)$ and $g_1(t_2)$ in the t_1, t_2 plane readily shows that they must intersect in at least one point (t_1^0, t_2^0) in the interior of the rectangle $0 \leqq t_1 \leqq \frac{1}{2}, \frac{1}{2} \leqq t_2 \leqq 1$, and hence that $(\beta_0, \alpha_0) = (\beta(t_1^0, t_2^0), \alpha(t_1^0, t_2^0))$ is a solution of $(^*)$.

To show the uniqueness of (β_0, α_0) let us assume that (β_1, α_1) is also a solution of $(^*)$. Then one of the lines $\alpha_0 + \beta_0 x$ and $\alpha_1 + \beta_1 x$ must lie entirely above the other in at least one of the sets $\{x:x > h(\frac{1}{2})\}$ and $\{x:x < h(\frac{1}{2})\}$. Suppose, for example, $\alpha_1 + \beta_1 x > \alpha_0 + \beta_0 x$ when $x > h(\frac{1}{2})$. Hence

$$F(\alpha_1 + \beta_1 h(t) - \phi[h(t)]) \geqq F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])$$

when $t > \frac{1}{2}$. We shall prove the existence of a value t^* with $\frac{1}{2} < t^* < 1$ for which the last inequality is strict. Since the line $\alpha_0 + \beta_0 h(t)$ passes through the function $\phi[h(t)]$ for exactly one value t^* with $\frac{1}{2} < t^* < 1$, we have $\alpha_0 + \beta_0 h(t^*) - \phi[h(t^*)] = 0$ for this t^*. It follows that

$$F(\alpha_1 + \beta_1 h(t^*) - \phi[h(t^*)]) > F(\alpha_0 + \beta_0 h(t^*) - \phi[h(t^*)]) = F(0),$$

because of the fact $\alpha_1 + \beta_1 h(t^*) - \phi[h(t^*)] > 0$ and $F'(0) > 0$. We thus have

$$\frac{1}{4} = \int_{\frac{1}{2}}^{1} F(\alpha_0 + \beta_0 h(t) - \phi[h(t)])\, dt < \int_{\frac{1}{2}}^{1} F(\alpha_1 + \beta_1 h(t) - \phi[h(t)])\, dt = \frac{1}{4}$$

which is a contradiction, and so (β_0, α_0) is the unique solution to $(^*)$.

5.2. *The small sample behavior of R_n under H_0.* In this subsection we shall give some conditions under which the theoretical asymptotic null distribution of R_n yields an adequate approximation to the true null distribution of R_n. We here restrict ourselves to the case where $n + 1 X_i$ values are equally spaced in the interval $[c, d]$, $X_0 = c$, $X_n = d$, the points (X_i, Y_i) such that $X_{k_\delta} \leqq X_i \leqq X_{n-k_\delta}$ being used for line estimation, and with $a(t) \equiv 1$. We then define $\delta = k_\delta/n$. According to our earlier results, under the null hypothesis of linearity R_n will have an asymptotically normal distribution with mean $n\delta$ and with variance $n\delta/2(1 - 2\delta)$. The adequacy of this normal approximation, however, depends upon the magnitude of n and δ.

The two major respects in which the approximation can be inadequate are
(i) distribution of R_n not approximately normal
(ii) Var $R_n \neq n\delta/2(1 - 2\delta)$.

In order to determine the extent to which (i) and (ii) occur, Monte Carlo methods have been used to investigate the null distribution of R_n for a variety of n and δ in the case where the Y_i are independently normally distributed with mean zero and variance one. Now for large values of δ (near $\frac{1}{2}$) it appears that both (i) and (ii) can occur, and make the asymptotic normal approximation inadequate even for n as large as 100. For this reason we shall restrict attention to small values of δ, and more or less arbitrarily take $0 < \delta \leqq 0.20$. (We remark, however, that this is a fairly conservative cut off point, and for values of δ near 0.20 we may expect the approximation to be good.)

Under this condition our investigation then indicates that so long as $n(2\delta) > 10$ the null distribution of R_n is satisfactorily normal for "ordinary" usage in testing the null hypothesis of linearity. Note here that $n(2\delta)$ is the number of points used to count R_n, and hence that the "rule" $n(2\delta) > 10$ is analogous to the "rule" for approximating the binomial distribution with parameters $P = \frac{1}{2}$ and N by a normal distribution when $NP = N/2 > 5$.

Now it appears from the Monte Carlo experiments that $n\delta/2(1 - 2\delta)$ is an overestimate of Var R_n. When $\delta \leqq 0.20$ and $n(2\delta) > 10$ the amount of bias is quite small, however, and the use of $n\delta/2(1 - 2\delta)$ in place of the true Var R_n will only slightly alter the upper tail probabilities involved in the test of linearity. Since $n\delta/2(1 - 2\delta) \geqq$ Var R_n, it follows that when performing the test at a nominal level θ of significance one would in fact be operating at a slightly smaller level.

TABLE I

R_n	Obs.	Exp.	R_n	Obs.	Exp.	R_n	Obs.	Exp.	R_n	Obs.	Exp.
0	0		0	0		0	0		7	1	
1	0	5.879	1	1	5.197	1	0		8	2	8.457
2	3		2	1		2	0	9.143	9	2	
3	8	7.299	3	15	7.085	3	2		10	4	
4	12	11.939	4	9	12.021	4	5		11	4	5.774
5	13	16.036	5	12	16.516	5	7	7.935	12	6	8.009
6	21	17.694	6	21	18.362	6	12	11.332	13	14	10.093
7	16	16.036	7	18	16.516	7	16	14.048	14	10	11.595
8	12	11.939	8	11	12.021	8	15	15.084	15	16	12.144
9	9	7.299	9	6	7.085	9	12	14.048	16	9	11.595
10	5		10	5		10	15	11.332	17	12	10.093
11	1	5.879	11	1	5.197	11	8	7.935	18	3	8.009
12	0		12	0		12	4		19	6	5.774
						13	3		20	6	8.457
						14	1	9.143	21	5	
						15	0				
						16	0				

n	29	33	37	99
δ	0.21	0.18	0.22	0.15
Sample mean	6.19	5.99	8.12	15.00
Sample variance	4.32	4.49	5.80	10.37
Theoretical mean	6.00	6.00	8.00	15.00
Theoretical variance	5.12	4.72	7.04	10.78
χ^2	3.068	13.692	2.621	8.887
d.f.	8	8	8	10

The results of some of the Monte Carlo experiments performed are presented in the table below, which the reader may use to judge the adequacy of the normal approximation. We remark that in approximating the true discrete distribution of R_n by a normal distribution the usual continuity correction of $\frac{1}{2}$ is desirable. Hence we would use, for example.

$$\Pr\{k_1 \leqq R_n \leqq k_2 \mid H_0\}$$

$$= \Pr\left\{\frac{k_1 - n\delta}{\left(\frac{n\delta}{2(1-2\delta)}\right)^{\frac{1}{2}}} \leqq \frac{R_n - n\delta}{\left(\frac{n\delta}{2(1-2\delta)}\right)^{\frac{1}{2}}} \leqq \frac{k_2 - n\delta}{\left(\frac{n\delta}{2(1-2\delta)}\right)^{\frac{1}{2}}} \,\middle|\, H_0\right\}$$

$$\sim \Phi\left\{\frac{k_2 - n\delta + \frac{1}{2}}{\left(\frac{n\delta}{2(1-2\delta)}\right)^{\frac{1}{2}}}\right\} - \Phi\left\{\frac{k_1 - n\delta - \frac{1}{2}}{\left(\frac{n\delta}{2(1-2\delta)}\right)^{\frac{1}{2}}}\right\}.$$

At present Monte Carlo Methods have only been used in the case where the Y_i are normally distributed. Whether the conditions $0 < \delta \leqq 0.20$ and $n(2\delta) > 10$ are appropriate for other distributions has not yet been investigated. The author suspects that they will be appropriate in a wide variety of cases.

In Table 1 we present the observed and expected frequency distributions of 100 random values of R_n, where expected frequencies are calculated from the asymptotic probability distribution of R_n. Four pairs of values of n and δ were investigated: $n = 29$, $\delta = .21$, $n = 33$, $\delta = .18$, $n = 37$, $\delta = .22$, and $n = 99$, $\delta = .15$. Some relevant statistics from the samples are presented, including the value of χ^2 from the χ^2 goodness of fit test.

Acknowledgments. I wish to express my thanks to Professor Lincoln E. Moses for his many insightful suggestions and guidance during the preparation of the doctoral dissertation, upon which the present paper is based.

REFERENCES

[1] Brown, G. W. and Mood, A. M. (1951). On median tests for linear hypotheses. *Second Berkeley Symp. Math. Statist. Prob.* 159–166. Univ. of Calif. Press.

[2] Cramér, Harald (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[3] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

[4] Loève, Michel (1955). *Probability Theory,* 1st ed. Van Nostrand, New York.

[5] Mood, Alexander McFarlane (1950). *Introduction to the Theory of Statistics.* McGraw-Hill, New York.

[6] Scheffé, Henry (1947). A useful convergence theorem for probability distributions. *Ann. Math. Statist.* **18** 434–438.

[7] Scheffé, Henry (1959). *The Analysis of Variance.* Wiley, New York.