The Annals of Statistics
1976, Vol. 4, No. 2, 375-383

CELL SELECTION IN THE CHERNOFF-LEHMANN CHI-SQUARE STATISTIC[1]

BY M. C. SPRUILL

Georgia Institute of Technology

The approximate Bahadur slope of the Chernoff-Lehmann χ^2-test-of-fit to a scale-location family on R^k is computed. The goal is to select cells (whose number is independent of sample size) to maximize this slope. The supremum is found and is shown to be a maximum only in trivial cases. If the sup is finite there is always a best selection for a fixed number of cells. Equally likely cells are shown to be admissible when the alternative is large.

1. Introduction. Let $(\mathscr{X}, \mathscr{A})$ be a measure space and μ and ν be two probability measures on $(\mathscr{X}, \mathscr{A})$. If $\pi: A_1, A_2, \cdots, A_M$ is a partition of $\mathscr{X}$ it is well known that

$$X_n^2(\pi) = n \sum_{j=1}^{M} (\mu_n(A_j) - \mu(A_j))^2/\mu(A_j) \tag{1}$$

converges in law to a χ^2_{M-1} distribution where μ_n is the empiric measure based on the independent $y_1, \cdots, y_n$ sampled from μ. Suppose however that, rather than μ, a nontrivial alternative $(\nu \ll \mu)$ ν holds and define

$$M(\alpha) = \inf\{m: X_n^2(\pi) > \chi^2_{M-1,\alpha} \text{ for } n \geqq m\}.$$

That is, $M(\alpha) - 1$ is the last time the approximate α-level test makes an error. It is known that w.p. 1 (ν)

$$\lim_{\alpha\to 0} -(2 \log \alpha)/M(\alpha) = \sum_{j=1}^{M} (\nu(A_j) - \mu(A_j))^2/\mu(A_j)$$

if the rhs is positive and finite (see Lemma 1). Thus the approximate Bahadur slope of the test based on (1) for testing $H_0: \mu$ against $H_1: \nu$ is $\Psi = \sum_{j=1}^{M} \nu^2(A_j)/\mu(A_j) - 1$ (Bahadur (1967)). It is readily seen that the "best" partition by this measure is the one resulting in the largest Ψ value.

The problem of maximizing the slope $\Psi(\pi)$ as a function of π for μ and ν fixed is considered in Section 2. The main result is Theorem 1, and the implications of this result are explained in comments following that theorem. In Section 3 the parallel result, Theorem 2, is obtained in the case of testing fit to a location-scale family against a location-scale alternative family using the Chernoff-Lehmann statistic with random cells (see Moore and Spruill (1975)). Section 4 shows that an optimality property of equally likely cells in the ordinary χ^2-test pointed out by Mann and Wald (1942) holds in the case of testing fit to a parametric family against a large alternative.

Received August 1974; revised July 1975.

[1] This research was in part sponsored by the Air Force Office of Scientific Research under Grant No. AFOSR 72-2350 at Purdue University.

AMS 1970 *subject classifications*. Primary 62G20; Secondary 62F10.

Key words and phrases. Chi-square test, Bahadur slope.

It should be noted that the results which follow could differ from those based on exact Bahadur slope since Bahadur (1967) has pointed out that there may be large differences between exact and approximate slopes especially at alternatives far from the null hypothesis.

2. Cell selection for simple hypotheses. Let $(\mathscr{X}, \mathscr{A})$ be a measure space and $\nu \ll \mu$ be probability measures on $(\mathscr{X}, \mathscr{A})$. Any finite collection of sets $\pi : A_1, \cdots, A_M$ satisfying

$$A_j \in \mathscr{A} \qquad j = 1, \cdots, M,$$

$$\mu(A_i \cap A_j) = 0 \qquad i \neq j,$$

and

$$\mu(\textstyle\bigcup_1^M A_j) = 1$$

is a partition of $\mathscr{X}$.

DEFINITION 1. (a) If $\pi : A_1, \cdots, A_M$ is a partition and $\mu(A_j) > 0$ for $j = 1, \cdots, M$ the partition is said to be admissible.

(b) The notation $\pi \leqq \pi'$ means that each set of π is contained except for a set of μ-measure zero in a set of π'.

(c) If Π is any collection of partitions $\mathscr{C}(\Pi) = \{A : A \in \pi \text{ for some } \pi \in \Pi\}$.

(d) The collection Π is complete if each $\pi \in \Pi$ is admissible and given $C \in \mathscr{C}(\Pi)$ and $\pi^0 \in \Pi$, $\pi^0 : B_1, \cdots, B_M$, there is a partition $\pi \in \Pi$ such that $\pi \leqq \pi^0$ and $C \cap B_j \in \pi$ unless $\mu(C \cap B_j) = 0$.

A set $D \subset R^k$ is said to be a rectangle with sides parallel to the coordinate axes if there are numbers $-\infty \leqq a_j \leqq b_j \leqq +\infty$, $j = 1, \cdots, k$, such that $D = \{x \in R^k : a_j \leqq x^{(j)} \leqq b_j, j = 1, \cdots, k\}$, $x = (x^{(1)}, x^{(2)}, \cdots, x^{(k)})$. The collection of all admissible M-partitions whose M component sets $A_1, \cdots, A_M$ are each rectangles with sides parallel to the coordinate axes will be denoted by Π_M. If for each sequence of sets $\{C_n\}$ such that $C_n \in \mathscr{C}(\bigcup_M \Pi_M)$, $\mu(C_n) \to 0$ implies $\nu^2(C_n)/\mu(C_n) \to 0$, one writes $\nu^2 = o(\mu)$.

THEOREM 1. *Assume* $\nu \ll \mu$.

(a) *If* $\pi : A_1, \cdots, A_M$ *is admissible and* $\nu(A_j) \neq \mu(A_j)$ *for some* j *the test of* $H_0 : \mu$ *against* $H_1 : \nu$ *which rejects for large values of* $X_n^2(\pi)$ *has approximate slope*

$$\Psi(\pi) = \textstyle\sum_{j=1}^M \nu^2(A_j)/\mu(A_j) - 1 .$$

(b) *If* Π *is complete and* $\pi^0 \in \Pi$, *a necessary and sufficient condition for* $\pi^0 : A_1^0, A_2^0, \cdots, A_{M_0}^0$ *to satisfy* $\psi(\pi^0) = \sup\{\psi(\pi) : \pi \in \Pi\}$ *is that there exist constants* $0 \leqq a_i < \infty$, $j = 1, \cdots, M_0$ *such that for all* $C \in \mathscr{C}(\Pi)$,

$$\nu(C) = \textstyle\sum_{i=1}^{M_0} a_i \mu(A_i^0 \cap C) .$$

(c) *If* $\mathscr{X} = R^k$, $\nu^2 = o(\mu)$, *and* $\mu \ll m$ *(Lebesgue measure on* R^k*) then for each integer* $M > 1$ *there is a* $\pi^M \in \Pi_M$ *such that* $\Psi(\pi^M) = \sup\{\Psi(\pi) : \pi \in \Pi_M\}$.

(d) *If* $\mathscr{X} = R^k$ *and* $\mu \ll m$ *then*

$$\sup\{\Psi(\pi) : \pi \in \textstyle\bigcup_{M=1}^\infty \Pi_M\} = \int g^2/f \, dm - 1$$

where $g = d\nu/dm$ *and* $f = d\mu/dm$.

COROLLARY 1. *Let* $\mathscr{X} = R^k$ *and* $\nu \ll \mu \ll m$.

(a) *If* $\pi^0 \in \bigcup_M \Pi_M$, *then* $\Psi(\pi^0) = \int g^2/f - 1$ iff g/f *is constant* a.e. m *on* A_i^0 *for* $i = 1, \cdots, M(\pi^0)$.

(b) *If* $\int g^2/f\, dm < \infty$ *then for each integer* $M > 1$ *there is a* $\pi^M \in \Pi_M$ *such that*

$$\Psi(\pi^M) = \sup \{\Psi(\pi) : \pi \in \Pi_M\} .$$

A few comments are in order. The theorem and corollary show that there is a "best" partition only in trivial cases. This implies that except in those trivial cases the partition may be bettered by subdividing an existing cell.

They may also be applied in the following way. If we are testing fit to μ against ν, and a competing test (not necessarily a χ^2-test) has slope Ψ^*, then if $\int g^2/f > \Psi^* + 1$ we can do better with a χ^2 test; if $\int g^2/f < \Psi^* + 1$ we can not do as well; and if $\int g^2/f = \Psi^* + 1$ the conclusion depends upon the conclusion of (a) of Corollary 1.

EXAMPLE 1. Consider the problem of testing fit to the continuous df F on R against the continuous df G. The Kolmogorov–Smirnov approximate level α test rejects H_0: F if $n^{\frac{1}{2}}||F_n - F||$ exceeds the upper α cutoff of

$$K(x) = \sum_{-\infty}^{\infty} (-1)^j \exp[-2j^2x^2] , \qquad 0 < x < \infty .$$

Here $||\ \ ||$ is the sup norm and F_n is the empiric df. Abrahamson (1965) shows that the approximate slope of this test is $\Psi^* = 4||F - G||^2$. Since F and G are continuous there is a point t_0 such that $||F - G|| = |F(t_0) - G(t_0)|$. Hence

$$\begin{aligned} \int g^2/f - 1 &\geqq \frac{(F(t_0) - G(t_0))^2}{F(t_0)} + \frac{(F(t_0) - G(t_0))^2}{1 - F(t_0)} \\ &= \frac{\Psi^*}{4}\left(\frac{1}{F(t_0)} + \frac{1}{1 - F(t_0)}\right) \geqq \Psi^* \end{aligned}$$

and the χ^2 test using $\pi: (-\infty, t_0), (t_0, +\infty)$ has slope at least as great as the $K - S$ test. In fact the only case in which no partition can be found to make the slope greater than Ψ^* is when

$$\frac{g(x)}{f(x)} = c_0 I(x)_{(-\infty, F^{-1}(\frac{1}{2}))} + c_1 I(x)_{(F^{-1}(\frac{1}{2}), +\infty)} ,$$

$c_j \geqq 0$, $c_0 + c_1 = 2$.

The proofs of the theorem and corollary are given below.

PROOF OF COROLLARY 1.

(a) It may be verified that $\bigcup_{M=1}^{\infty} \Pi_M$ is complete and that $\mathscr{C}(\bigcup_M \Pi_M)$ contains all rectangles of nonzero μ-measure. Using (b) and (d) Theorem 1 one proves (a).

(b) Let $\mu(A) > 0$. By Schwarz's inequality.

$$[\textstyle\int_A (g/f^{\frac{1}{2}})(f^{\frac{1}{2}})]^2 \leqq [\int_A g^2/f][\int_A f]$$

so that

$$\nu^2(A)/\mu(A) \leqq \textstyle\int_A g^2/f . \tag{2}$$

Now (b) follows from Theorem 1 (c) and the dominated convergence theorem.

The proof of Theorem 1 is carried out below. Let $\chi_1^2, \chi_2^2, \cdots, \chi_m^2$ be independent χ^2 random variables with a(j) degrees of freedom and let $C(x)$ be the df of the random variable $\sum_{j=1}^M \lambda_j \chi_j^2$, $1 = \lambda_1 > \lambda_j \geqq 0$, $j = 2, \cdots, M$, λ_j fixed. The first lemma is an easy consequence of the arguments given on page 312 of Bahadur (1967) and in Theorem 7 of Gleser (1966).

LEMMA 1. *Let Z_n be nonnegative random variables with* df's *$H_{n\theta}$ for $\theta \in \Theta$, and suppose for $\theta \in \Theta_0 \subset \Theta$, $||H_{n\theta} - C|| \to 0$ as $n \to \infty$. If for some $\theta' \in \Theta - \Theta_0$, $Z_n/n \to a$* a.s., *then the approximate slope $\Psi(\theta')$ of the test which rejects $H_0: \theta \in \Theta_0$ if $Z_n > C^{-1}(\alpha)$ exists and equals a at θ'.*

Part (a) of Theorem 1 is an immediate consequence of Lemma 1.

LEMMA 2. *Let $a_i \geqq 0$, $b_i > 0$, $\sum_{i=1}^M a_i = a$, and $\sum_{i=1}^M b_i = b$. Then*

$$\sum_{i=1}^M \frac{a_i^2}{b_i} - \frac{a^2}{b} = \frac{1}{b} \sum_{i=1}^{M-1} \sum_{j>i} b_i b_j \left(\frac{a_i}{b_i} - \frac{a_j}{b_j}\right)^2 \geqq 0 .$$

PROOF OF (b) OF THEOREM 1. Assume there are constants $0 \leqq a_i < \infty$, $i = 1, \cdots, M$, and a partition $\pi^0 : A_1^0, \cdots, A_M^0$ such that for all $C \in \mathscr{C}(\Pi)$

$$\nu(C) = \sum_{j=1}^M a_j \mu(C \cap A_j^0) .$$

Let $\pi' \in \Pi$ be arbitrary. Since Π is complete there is a $\pi \in \Pi$ such that $\pi \leqq \pi'$ and $\pi \leqq \pi^0$. By Lemma 2 $\Psi(\pi) \geqq \Psi(\pi')$. Computation shows that $\Psi(\pi) = \Psi(\pi^0)$. Sufficiency of the condition has been shown. To see the necessity let A_i^0 be an arbitrary member of π^0 where

$$\Psi(\pi^0) = \sup\{\Psi(\pi) : \pi \in \Pi\} .$$

Without loss of generality we may assume it to be A_1^0. Suppose there is a $C \in \mathscr{C}(\Pi)$ such that $\mu(A_1^0 \cap C) = \mu(C)$, and $\mu(C^c \cap A_1^0) > 0$. By the completeness of Π there is a $\pi \in \Pi$, $\pi : B_1, B_2, \cdots, B_n$, such that $B_1 = A_1^0 \cap C$. Thus by Lemma 2

$$\Psi(\pi) \geqq \nu^2(A_1^0 \cap C)/\mu(A_1^0 \cap C) + \nu^2(A_1^0 \cap C^c)/\mu(A_1^0 \cap C^c) + \sum_{j=2}^M \nu^2(A_j^0)/\mu(A_j^0) - 1 .$$

It follows from Lemma 2 that if

$$\frac{\nu(A_1^0 \cap C)}{\mu(A_1^0 \cap C)} \neq \frac{\nu(A_1^0 \cap C^c)}{\mu(A_1^0 \cap C^c)} \tag{3}$$

then

$$\frac{\nu^2(A_1^0 \cap C)}{\mu(A_1^0 \cap C)} + \frac{\nu^2(A_1^0 \cap C^c)}{\mu(A_1^0 \cap C^c)} > \frac{\nu^2(A_1^0)}{\mu(A_1^0)} .$$

This however implies that $\Psi(\pi) > \Psi(\pi^0)$ which is a contradiction. Thus equality holds in (3). Manipulation shows that this implies that $\nu(A_1^0 \cap C) = (\nu(A_1^0))/(\mu(A_1^0))\mu(A_1^0 \cap C)$. The conclusion for arbitrary $C \in \mathscr{C}(\Pi)$ follows by using this argument and the completeness of Π.

Let $\{\pi_n\}$ be a sequence of partitions in Π_M with $\{x_{1n}, x_{2n}, \cdots, x_{Kn}\}$, $K \leqq 2^k M$, the set of vertices of the component rectangles; $x_{jn} \in [-\infty, +\infty]^k$. Since $[-\infty, +\infty]^k$ is compact and K is bounded, we may in an obvious manner select a subsequence $\{\pi_{n'}\}$ and a partition π_0 so that the vertices of the $\pi_{n'}$ converge to those of π_0. π_0 may fail to be admissible, however it is clear that there is a $\pi^M \in \Pi_M$ such that $\pi^M \leqq \pi_0$ if $\mu \ll m$.

PROOF OF (c) OF THEOREM 1. Let $\pi_n \in \Pi_M$ be such that

$$\Psi(\pi_n) \to \sup\{\Psi(\pi) : \pi \in \Pi_M\} .$$

Let π_0, π^M and $\{\pi_{n'}\}$ be as above. Clearly

$$\Psi(\pi_{n'}) \to \Sigma\nu^2(A_j^0)/\mu(A_j^0) - 1$$

where the sum extends over only those indices j such that $\mu(A_j^0) > 0$. By Lemma 2 $\Psi(\pi^M) \geqq \lim \Psi(\pi_n)$ since $\pi^M \leqq \pi_0$.

Rudin (1966) defines a substantial family S of open sets in R^k to be any family satisfying:

(a) There is a constant $\beta < \infty$ such that each $E \in S$ is contained in a ball B with $m(B) < \beta m(E)$.

(b) To every $y \in R^k$ and $\delta > 0$ there exists an $E \in S$, whose diameter is less than δ, such that $y \in E$.

DEFINITION 2. The sequence of partitions $\{\pi_n\}$ has mesh converging to zero properly (wrt μ) if there is a substantial family S and for almost every (μ) $y \in R^k$ there is a sequence of sets $\{A_n(y)\}$, $A_n(y) \in \pi_n$, and an integer $N(y) < \infty$ such that $y \in \operatorname{int}(A_n(y)) \in S$ for $n \geqq N(y)$, and $\operatorname{diam}[A_n(y)] \to 0$ as $n \to \infty$. For definiteness we take S to be the set of all rectangles A which are contained in a ball B for which $m(B) < k^k m(A)$. This collection is nonvoid and in particular contains all k-cubes.

PROOF OF (d) OF THEOREM 1. It is first proved that if $\{\pi_n\}$ has mesh converging to zero properly then

$$\lim_{n\to\infty} \Psi(\pi_n) = \int g^2/f\, dm - 1 . \tag{4}$$

The proof then follows from the fact that if

$$\Psi(\pi_n^*) \to \sup\{\Psi(\pi) : \pi \in \bigcup \Pi_M\}$$

there is a sequence $\{\pi_n\}$ with mesh converging to zero properly and $\pi_n \leqq \pi_n^*$ for all n. To prove (4) let $\pi_n : A_{1n}, \cdots, A_{M(n)n}$ and define the functions

$$h_n(y) = \sum_{A_{jn} \in \pi_n} \nu(A_{jn})/\mu(A_{jn}) I_{A_{jn}}(y) .$$

Since π_n converges to zero properly (wrt μ) and $\nu \ll \mu$, for almost every (ν) y eventually

$$h_n(y) = \frac{\nu(A_n(y))}{\mu(A_n(y))}, \qquad A_n(y) \in \pi_n, \qquad \text{and} \qquad A_n(y) \in S .$$

Writing

$$h_n(y) = \frac{\nu(A_n(y))/m(A_n(y))}{\mu(A_n(y))/m(A_n(y))}$$

it follows from Rudin (1966), page 154, that $\lim h_n(y) = g(y)/f(y)$. Thus $h_n \to g/f$ a.s. ν. By Fatou's lemma

(5) $$\int g^2/f\, dm - 1 = \int \liminf h_n\, d\nu - 1 \leqq \liminf \int h_n\, d\nu - 1 .$$

Since $\int h_n\, d\nu - 1 = \Psi(\pi_n)$ it follows from (2) and (5) that

$$\int g^2 f\, dm - 1 \leqq \liminf \Psi(\pi_n) \leqq \limsup \Psi(\pi_n) \leqq \int g^2/f\, dm - 1 .$$

3. Cell selection in a composite hypotheses case. In this section Theorem 1 is extended to the case of testing fit to a location-scale family against another location-scale family as alternative. Let $\{F_\theta\}$ be a scale-location family on R^k generated by the absolutely continuous df F,

$$F_\theta(y) = F\left(\frac{y^{(1)} - \theta^{(1)}}{\theta^{(2)}}, \cdots, \frac{y^{(k)} - \theta^{(2k-1)}}{\theta^{(2k)}}\right),$$

$$-\infty < y^{(j)} < +\infty ,$$

$$\theta \in \Theta = \{\theta \in R^{2k} : \theta^{(2j)} > 0,\ -\infty < \theta^{(2j-1)} < +\infty\} .$$

If $\pi_0 \in \Pi_M(F)$ (the admissibility is determined by F) and the vertices of the rectangles of π_0 are $\{x_{01}, x_{02}, \cdots, x_{0m}\}$, $x_{0j} \in R^k$, denote by $\pi(\theta, \pi_0)$ for each $\theta \in \Theta$ the partition with sets corresponding to π_0 and vertices $\{x_1(\theta), \cdots, x_m(\theta)\}$ given by $x_i^{(j)}(\theta) = \theta^{(2j)} x_{0i}^{(j)} + \theta^{(2j-1)}, j = 1, \cdots, k; i = 1, \cdots, m$. The Chernoff–Lehmann statistic T_{2n} (see Moore and Spruill (1974)) is given by

$$T_{2n} = \sum_{j=1}^{M} (N_{nj}(\hat\theta_n) - np_j(\hat\theta_n))^2/np_j(\hat\theta_n)$$

where $N_{nj}(\hat\theta_n)$ is the number of observations $y_1, \cdots, y_n$ falling in the jth cell of $\pi(\hat\theta_n, \pi_0)$ and $p_j(\hat\theta_n)$ is the probability which $F_{\hat\theta n}$ assigns to it. The MLE $\hat\theta_n$ maximizes $\sum_{j=1}^{n} \log f(y_j \mid \theta)$, where $f(\cdot \mid \theta)$ is the density of F_θ.

The following assumptions are made.

(B1) T_{2n} has as its limiting distribution the distribution of $\chi^2_{M-2k-1} + \sum_{j=M-2k}^{M-1} \lambda_j \chi_j^2$ under any $F_\theta \in \{F_\theta : \theta \in \Theta\}$.

(B2) Under an alternative df G such that $G \ll F_{\theta_0}$, $\hat\theta_n \to \theta_0$ w.p. 1 where $\theta_0 \in \Theta$ maximizes $\int \log f(x \mid \theta)\, dG(x)$.

(B3) $M > 2k + 1$.

Sufficient conditions for (B1) may be found in Moore and Spruill (1975). See Perlman (1972) for conditions relating to (B2).

Under these assumptions it follows from Lemma 1 that the slope of the test of fit to the family $\{F_\theta\}$ is

$$\Psi(\pi(\theta_0, \pi_0)) = \sum_{j=1}^{M} (G(A_j) - F_{\theta_0}(A_j))^2/F_{\theta_0}(A_j)$$

if $\pi(\theta_0, \pi_0) : A_1, \cdots, A_M$ is admissible. Here as below the set function determined by a distribution function L is denoted by the same letter. Thus $L(A) = \int_A dL(x)$ when A is a measurable set. Elementary arguments (Spruill (1973), Moore and Spruill (1975)) show that if $G_{\theta'} \in \{G_\theta : \theta \in \Theta\}$ the slope does not depend upon the particular element $G_{\theta'}$.

THEOREM 2. *Under the assumptions* B:

(a) *If* $\pi_0 \in \Pi_M(F)$ *and* $G(A_j) \neq F_{\theta_0}(A_j)$ *for some* j, *the slope of the test of* $H_0: \{F_\theta\}$ *against* $H_A: \{G_\theta\}$ *based on* T_{2n} *is*

$$\Psi(\pi(\theta_0, \pi_0)) = \sum_{j=1}^{M} (G(A_j) - F_{\theta_0}(A_j))^2 / F_{\theta_0}(A_j)$$

for every $G_\theta \in \{G_\theta\}$.

(b) *If* $\pi_0 \in \bigcup_M \Pi_M(F)$ *then*

$$\Psi(\pi(\theta_0, \pi_0)) = \int g^2/f_{\theta_0} - 1 \quad \text{iff}$$

g/f_{θ_0} *is constant* a.e. m *on the cells of* $\pi(\theta_0, \pi_0)$.

(c) *If* $G^2 = o(F_{\theta_0})$ *then for each* $M > 2k + 1$ *there is a partition* $\pi^M \in \Pi_M(F)$ *such that*

$$\Psi(\pi(\theta_0, \pi^M)) = \sup \{\Psi(\pi(\theta_0, \pi)) : \pi \in \Pi_M(F)\} .$$

(d) $\sup \{\Psi(\pi(\theta_0, \pi)) : \pi \in \bigcup_M \Pi_M(F)\} = \int g^2/f_{\theta_0} \, dm - 1.$

EXAMPLE 2. Consider the scale-location families generated by the following densities:

(1) $g(x) = e^{-x}$, $x \geqq 0$.

(2) $g(x) = e^{-|x|/2}$, $-\infty < x < +\infty$.

(3) $g(x) = e^x(1 + e^x)^{-2}$, $-\infty < x < +\infty$.

(4) $g_\beta(x) = e^{-x}x^{\beta-1}/\Gamma(\beta)$, $x > 0$, $\beta > 2$ known.

(5) $g_\beta(x) = \{\exp[-\frac{1}{2}|x|^{2/1+\beta}]\}/2^{(3+\beta)/2}\Gamma((3 + \beta)/2)$, $-\infty < x < +\infty$, $\beta \in (-1, 1] - \{0\}$ known.

If $G(x)$ is the corresponding df and $F(x)$ is the standard normal density it can be checked that B1 and B2 are satisfied (see Dahiya and Gurland (1971)). It is easily seen that in each case except (5) for $\beta \in (-1, 0)$, $\int g^2/f_{\theta_0} = +\infty$. In fact it is easily checked that $(1 - G(x))^2/(1 - F_{\theta_0}(x)) \to \infty$ as $x \to \infty$ so G^2 is not $o(F_{\theta_0})$. Hence with that exception given any specific test a χ^2 test can be found with larger slope using no more than 4 cells. In the exceptional case a best M-partition exists for every $M > 3$.

4. Equally likely cells for large alternatives. In Section 3 certain results were given concerning cell selection in the case of testing $H_0: \{F_\theta\}$ against $H_A: \{G_\theta\}$. In this section it is shown that for much larger alternatives $\mathscr{G}$ the use of equally likely cells in T_{2n} for testing $H_0: \{\Phi_\theta\}$ against $H_A: \mathscr{G}$, Φ the standard normal, is desirable. The reasons are similar to those given in Mann and Wald (1942) for testing $H_0: F$ against $H_A: \mathscr{G}$.

Let $\mathscr{G} = \{G : G \ll m, G \text{ a df}, \int |x| \, dG(x) < \infty\}$ and $\{\Phi_\theta\}$ be the unit normal location family. An M-partition $\pi: (-\infty, a_1), (a_1, a_2), \cdots, (a_{M-1}, +\infty)$ is said to be equally likely if $\Phi(a_j) - \Phi(a_{j-1}) = 1/M$, $j = 1, \cdots, M$, $a_0 = -\infty$, $a_M = +\infty$.

THEOREM 3. (a) *For testing* H_0: $\{\Phi(x - \theta)\}$ *against* H_A : $\mathscr{G}$ *the test* T_{2n} *using cells* $(a_{j-1} + \bar{x}, a_j + \bar{x})$, $j = 1, \cdots, M$, *has slope*

$$\Psi^{\cdot} = \sum_{j=1}^{M} (G(a_j + \mu) - G(a_{j-1} + \mu))^2/(\Phi(a_j) - \Phi(a_{j-1})) - 1$$

at $G \in \mathscr{G}$ *where* $\mu = \int x\, dG(x)$ *if* $G(a_j + \mu) \neq \Phi(a_j)$ *some* j.

(b) *Let* π_e *be the equally likely M-partition and* π *be any other M-partition for which* $\Psi(\pi_e, G) \leqq \Psi(\pi, G)$ *for all* $G \in \mathscr{G}$. *Then* π *is equally likely.*

(c) *If* π_e *is equally likely there is a* $G \in \mathscr{G}$ *and a partition* π *for which* $0 < \Psi(\pi_e, G) < \Psi(\pi, G)$.

Part (a) of Theorem 3 is just a special case of Theorem 2(a). Theorem 3(c) is easily proved by using Theorem 2(b) and simply constructing a df $G \in \mathscr{G}$ whose density is a constant multiple of the normal on nonequally likely cells. It remains to prove Theorem 3(b). Let π : $(-\infty, a_1), (a_1, a_2), \cdots, (a_{M-1}, +\infty)$ and π' : $(-\infty, b_1), \cdots, (b_{M-1}, +\infty)$ be two partitions. Define $I_i = (a_{i-1}, a_i)$, $J_i = (b_{i-1}, b_i)$, $f_1(G) = \sum_{j=1}^{M} G^2(I_j)/\Phi_{\theta_0}(I_j)$, and $f_2(G) = \sum_{j=1}^{M} G^2(J_j)/\Phi_{\theta_0}(J_j)$.

LEMMA 3. *If* $G^* \in \mathscr{G}$ *is such that* $\int x\, dG^*(x) = \theta_0$ *and* $f_1(G^*) > f_2(G^*) = 1$, *then there is a* $G_0 \in \mathscr{G}$ *such that* $\int x\, dG_0(x) = \theta_0$ *and* $f_1(G_0) > f_2(G_0) > 1$.

PROOF OF (b) OF THEOREM 3. It is shown that if π : $(-\infty, b_1), \cdots, (b_{M-1}, \infty)$ is not equally likely, there is a $G \in \mathscr{G}$ such that $0 < \Psi(\pi, G) < \Psi(\pi_e, G)$. Since

$$\sup\{||\Phi_{\theta_0} - G|| : \int x\, dG(x) = \theta_0, G(b_j) = \Phi(b_j - \theta_0), j = 1, \cdots, M, G \in \mathscr{G}\}$$
$$= \max\{\Phi(b_j - \theta_0) - \Phi(b_{j-1} - \theta_0)\},$$

there is a $G^* \in \mathscr{G}$ such $||\Phi_{\theta_0} - G^*|| > 1/M$, $\Phi(b_j - \theta_0) = G^*(b_j)$, $j = 1, \cdots, M$, and $\int x\, dG^*(x) = \theta_0$. For the same reason it must be that if π_e : $(-\infty, a_1), \cdots, (a_{M-1}, \infty)$ is equally likely then $\Phi(a_j - \theta_0) - \Phi(a_{j-1} - \theta_0) \neq G^*(a_j) - G^*(a_{j-1})$ for some index j. This shows that $f_1(G^*) > f_2(G^*) = 1$. The proof of (b) now follows by taking G to be the G_0 of Lemma 3.

Acknowledgment. The research reported here was part of the author's doctoral thesis written at Purdue University. I thank Professor D. S. Moore for his guidance and helpful comments.

REFERENCES

[1] ABRAHAMSON, I. G. (1965). On the stochastic comparison of tests of hypotheses. Ph. D. thesis, Univ. of Chicago.

[2] BAHADUR, R. R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Statist.* **38** 303–325.

[3] DAHIYA, R. C. and GURLAND, J. (1971). Pearson chi-square test of fit with random intervals II. Nonnull case. Technical Summary Report #1046, Mathematics Research Center, Univ. of Wisconsin.

[4] GLESER, L. J. (1966). The comparison of multivariate tests of hypotheses by means of Bahadur efficiency. *Sankhyā Ser. A* **28** 151–174.

[5] KAC, M., KIEFER, J. and WOLFOWITZ, J. (1955). On tests of normality and other tests of goodness of fit based on distance methods. *Ann. Math. Statist.* **26** 189-211.

[6] MANN, H. B. and WALD, A. (1942). On the choice of the number of class intervals in the application of the chi-square test. *Ann. Math. Statist.* **13** 306-317.

[7] MOORE, D. S. and SPRUILL, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist.* **3** 599-616.

[8] PERLMAN, M. D. (1972). On the strong consistency of approximate maximum likelihood estimators. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 263-281, Univ. of California Press.

[9] RUDIN, W. (1966). *Real and Complex Analysis.* McGraw-Hill, New York.

[10] SPRUILL, M. C. (1973). Large sample properties of chi-square goodness-of-fit tests. Ph. D. thesis, Purdue Univ.

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332