THE ASYMPTOTIC THEORY OF GALTON'S TEST AND A RELATED SIMPLE ESTIMATE OF LOCATION[1]

By P. J. Bickel and J. L. Hodges, Jr.

University of California, Berkeley

1. Introduction and summary. In [8] Hodges and Lehmann showed that robust estimates of location in the one and two sample problems could be obtained by inverting the known robust nonparametric tests and using as estimates the average of the resulting upper and lower 50 per cent confidence bounds. In particular, they proposed the use of the estimate derived from the Wilcoxon test, the median of averages of pairs of observations. Unfortunately, despite some short cuts viz. [8] and [9] computation of this estimate seems to require on the order of $n^2 \log_2 n$ steps, a prohibitive number. In [7] Hodges proposed a simple alternative estimate D_n given by $D_n = \text{median}_i \frac{1}{2}(Z_{in} + Z_{i(n-i+1)})$ where $Z_{1n} < \cdots < Z_{nn}$ are the order statistics of the sample under consideration.

This procedure as was noted in [7] is related in the sense of Hodges and Lehmann to the one sample analogue of one of the oldest of non parametric tests, the Galton rank order test, viz. [6], [4a].

In this paper we derive the asymptotic theory of D_n by employing an invariance principle due to Bickel [2] and thus relating the limiting distribution to that of certain functionals of Brownian motion. The necessary refinements of the stochastic process convergence results of [2], which may be of use in related problems of asymptotic theory, are gathered in Section 7.

Unfortunately, we can only give explicit form to the limiting distribution of D_n in the two cases of rectangular and Laplace parents. Although this limit is *not* normal we conclude that D_n's scatter is quite close to that of the estimate proposed by Hodges and Lehmann.

Using the same techniques we prove in Section 6 the consistency of the Galton test and characterize its power behaviour for alternatives close to the null hypothesis.

Finally, Section 4 gives the small sample distribution of the Galton test statistic and of D_n for a rectangular parent.

The techniques of this paper carry over quite easily to the two sample situations.

Although our evidence is incomplete it would seem that both D_n and the Galton test are robust as well as easily computable non parametric procedures.

2. Asymptotic theory. Let X_i, $1 \leqq i \leqq n$, be a sample from a population with distribution $F(x - \theta)$ and density $f(x - \theta)$, where f is symmetric about 0,

Received 26 July 1965; revised 7 July 1966.

[1] Prepared with the partial support of NSF Grant GP 2593, and ONR-222(43).

and continuous and strictly positive on the convex support of F, $\{x:0< F(x) < 1\}$. Denote by $Z_{1n} < \cdots < Z_{nn}$ the order statistics of the sample. Let $n = 2m$ or $2m - 1$. In either case we define,

$$D_n(X_1, \cdots X_n) = \text{median}_{1 \leqq i \leqq m} \tfrac{1}{2}(Z_{in} + Z_{(n-i+1)n}). \tag{2.1}$$

This estimate is referred to as D in [7].

We observe the usual convention of letting D_n be the mid-point of the interval of medians, if there is any ambiguity. It follows immediately from the definition that D_n is translation invariant and symmetrically distributed about θ. Let,

$$N_n(a) = \textstyle\sum_{i=1}^{m} I([Z_{in} + Z_{(n-i+1)n} > 2a]), \tag{2.2}$$

where $I(A)$ is the indicator of the event A. We denote by $P_\theta[A]$ the probability of A when θ is the true value of the parameter. The following lemma is a ready consequence of the above remarks.

LEMMA 2.1. *Under the above conditions we have,*

$$P_0[N_n(xn^{-\frac{1}{2}}) \leqq \tfrac{1}{2}(m-2)] \leqq P_\theta[n^{\frac{1}{2}}(D_n - \theta) \leqq x] \leqq P_0[N_n(xn^{-\frac{1}{2}}) \leqq m/2]. \tag{2.3}$$

PROOF. $[n^{\frac{1}{2}}(D_n - \theta) \leqq x] = [D_n(X_1 - \theta, \cdots, X_n - \theta) \leqq xn^{-\frac{1}{2}}]$ by translation invariance. The lemma follows upon employing definitions (2.1) and (2.2).

Let $Q(t)$ be a version of standard Brownian motion on $[0, 1]$ with a.s. continuous sample functions, i.e. $Q(t)$ is a Gaussian process with stationary independent increments, $Q(0) = 0$, $\operatorname{Var} Q(1) = 1$. Let $\psi(t) = f(F^{-1}(t))$. Denote by $A(\psi, x)$ the random set $\{t:Q(t) > 2^{\frac{1}{2}}x\psi(t),\ t \in (0, \frac{1}{2})\}$. We now proceed to prove,

LEMMA 2.2. *Under the above conditions,*

$$\lim_n P_0(N_n(xn^{-\frac{1}{2}}) \leqq m\alpha) = P(\lambda(A(\psi, x)) \leqq \alpha/2)$$

where λ *is Lebesgue measure on* $(0, \frac{1}{2})$.

PROOF. Let $Z_n^*(t)$ be the process on $[0, 1]$ defined by,

$$Z_n^*(t) = Z_{kn}^* \quad \text{on} \quad [(k-1)/n, k/n), \qquad 1 \leqq k \leqq n, \tag{2.4}$$

where

$$Z_{kn}^* = Z_{kn} - F^{-1}(k/(n+1)), \qquad Z_{0n}^* = 0, \qquad Z_n^*(1) = Z_{nn}^*.$$

Now, $Z_n^*(t)$ may be considered as a probability measure on $D[\epsilon, 1 - \epsilon]$, the space of all real valued functions on $[\epsilon, 1 - \epsilon]$ which possess right and left hand limits at each point of $(\epsilon, 1 - \epsilon)$, endowed with the Prohorov metric. For details on the properties of this space and the formal definition of the above correspondence the reader is referred to Prohorov (12) and Section 7. Then, if $\theta = 0$, by Theorem 7.1, $n^{\frac{1}{2}}Z_n^*(t)$ converges weakly on $[\epsilon, 1 - \epsilon]$ to $(\psi(t)^{-1}Z(t)$ where $Z(t)$ is a 'brownian bridge', a Gaussian process with a.s. continuous sample functions and cov $(Z(s), Z(t)) = s(1 - t)$, $s \leqq t$.

For various equivalent definitions of weak convergence of measures the reader is referred to (12) pp. 164–165. The definition which we shall employ is embodied in Theorem 1.8 of (12) and states:

DEFINITION 2.5. A sequence of measures μ_n on a complete separable metric space C endowed with the Borel σ-field converges weakly to a measure μ if and only if for every real valued measurable function f on C which is μ almost everywhere continuous and bounded, $F_n(x) = \mu_n(f^{-1}(-\infty, x)) \to G(x) = \mu(f^{-1}(-\infty, x))$ for all $-\infty < x \leqq \infty$ at which G is continuous.

By corollary 7.1, if $\theta = 0$, $\frac{1}{2}n^{\frac{1}{2}}(Z_n^*(t) + Z_n^*(1 - t))$ converges weakly to $\frac{1}{2}(\psi(t))^{-1}(Z(t) + Z(1 - t))$ on $[\epsilon, \frac{1}{2}]$ for every $\epsilon > 0$. Now, $\frac{1}{2}(Z(t) + Z(1 - t))$ is, of course, Brownian motion with $\sigma^2 = \frac{1}{2}$ on $[0, \frac{1}{2}]$. Let $\hat{Z}_n(t)$ be the process defined on $[0, 1]$ by,

$$\hat{Z}_n(t) = Z_{kn} \quad \text{for} \quad t \,\varepsilon\, ((k - 1)/n, k/n) \tag{2.6}$$

where $Z_{0n} = 0$, $\hat{Z}_n(1) = Z_{nn}$. Then,

$$\begin{aligned} &\sup_{(\epsilon, \frac{1}{2})} n^{\frac{1}{2}}|\hat{Z}_n(t) + \hat{Z}_n(1 - t) - (Z_n^*(t) + Z_n^*((1 - t))| \\ &\quad \leqq \sup_{\epsilon n \leqq k \leqq n/2} n^{\frac{1}{2}}|F^{-1}(k/(n + 1)) - F^{-1}(k/n)| \\ &\qquad + |F^{-1}(1 - (k/n)) - F^{-1}(1 - ((k + 1)/(n + 1)))| \\ &\quad \leqq 2/n^{\frac{1}{2}} \sup_{\epsilon \leqq t \leqq \frac{1}{2}} (\psi(t))^{-1}. \end{aligned} \tag{2.7}$$

We may conclude from (2.7) by arguments similar to those used in Theorem 7.4 that $(n^{\frac{1}{2}}/2)(\hat{Z}_n(t) + \hat{Z}_n(1 - t))$ also converges to $(\psi(t))^{-1}/2(Z(t) + Z(1 - t))$ if $\theta = 0$. Let

$$A_n(\epsilon, x) = \{t: \tfrac{1}{2}(\hat{Z}_n(t) + \hat{Z}_n(1 - t)) > xn^{-\frac{1}{2}}, t \,\varepsilon\, [\epsilon, \tfrac{1}{2}]\}$$

and

$$A^*(\epsilon, x) = \{t: \tfrac{1}{2}(Z(t) + Z(1 - t)) > x\psi(t), t \,\varepsilon\, [\epsilon, \tfrac{1}{2}]\}.$$

We remark that $P(\lambda(t: \frac{1}{2}(Z(t) + Z(1 - t)) = x\psi(t)) = 0) = 1$. This follows by an argument similar to that used in Levy (10) p. 31. Now, by Definition 2.4, and Theorem 7.4 we may conclude that,

$$P_0(\lambda(A_n(\epsilon, x)) \leqq \alpha/2) \to P(\lambda(A^*(\epsilon, x) \leqq \alpha/2) \quad \text{for } 0 \leqq \alpha \leqq 1 - 2\epsilon. \tag{2.8}$$

We remark that $|\lambda(A_n(\epsilon, x)) - \lambda(A_n(0, x))| \leqq \epsilon$ and hence that,

$$\limsup_{\epsilon \to 0} \limsup_n P(\lambda(A_n(\epsilon, x)) - \lambda(A(0, x)) \geqq \delta) = 0, \tag{2.9}$$

for every $\delta > 0$.

Moreover $\lambda(A^*(\epsilon, x))$ converges in probability to $\lambda(A^*(0, x))$ as $\epsilon \to 0$. We then find from lemma 4.1 of Bickel (2) that

$$P_0(\lambda(A_n(0, x)) \leqq \alpha/2) \to P(\lambda(A^*(0, x)) \leqq \alpha/2), \qquad 0 \leqq \alpha \leqq 1.$$

But $|\lambda(A_n(0, x)) - N_n(xn^{-\frac{1}{2}})/2m| \leqq 1/n \to 0$ as $n \to \infty$. The lemma is proved.

As an immediate consequence of the lemmas we have,

THEOREM 2.1. *The asymptotic distribution of D_n is symmetric about θ and is* *given by,*

$$\lim_n P_\theta(n^{\frac{1}{2}}(D_n - \theta) \leqq x) = P(\lambda(A(\psi, x)) \leqq \tfrac{1}{4}) \tag{2.10}$$

The distribution of $\lambda(A(\psi, x))$ is, in general, unknown. We are able to give an explicit expression in two special cases which follow in Section 3.

3. Large sample distribution. (Rectangular and Exponential Cases). Let $f(x) = \frac{1}{2}$, $-1 \leqq x \leqq 1$ and 0 otherwise. In this case since $\psi(t) \equiv \frac{1}{2}$, the limiting distribution takes on a particularly simple form, and we have,

$$\lim_n P_\theta(n^{\frac{1}{2}}(D_n - \theta) \leqq x) = P(\lambda(t{:}2^{\frac{1}{2}}Q(t) > x, t \in [0, \tfrac{1}{2}]) \leqq \tfrac{1}{4}). \tag{3.1}$$

But if $Q_1(t)$ is standard Brownian motion on $[0, \frac{1}{2}]$, $R(t) = 2^{\frac{1}{2}}Q(t/2)$ is standard Brownian motion on [0, 1] and we conclude that

$$\lim_n P_\theta(n^{\frac{1}{2}}(D_n - \theta) \leqq x) = P(\lambda(t{:}R(t) > x, t \in (0, 1)) \leqq \tfrac{1}{2}). \tag{3.2}$$

Denote $P(\lambda(t{:}R(t) > x, t \in (0, 1)) \leqq \alpha)$ by $g(x, \alpha)$. Let $f(x, s)$ be the density of T_x, the first time $R(t)$ reaches x for $x > 0$. Then it is well known that (cf. Lévy (10))

$$g(0, \alpha) = 2/\pi \cdot \text{arc sin } \alpha^{\frac{1}{2}}, \tag{3.3}$$

$$f(x, s) = x(2\pi)^{-\frac{1}{2}}s^{-\frac{3}{2}} \exp - x^2/2s \quad \text{for} \quad s > 0 \tag{3.4}$$

and,

$$P(\lambda(t{:}R(t) > 0, t \in (0, a)) \leqq \alpha) = g(0, \alpha/a). \tag{3.5}$$

We now evaluate $g(x, \alpha)$ for $x > 0$. We obtain first,

$$g(x, \alpha) = \int_0^\infty P(\lambda(t{:}R(t) - R(s) > 0, t \in (s, 1)) \leqq \alpha) \, dP(T_x \geqq s). \tag{3.6}$$

From (3.4), (3.5) and the strong Markov property of Brownian motion we conclude that,[2]

$$g(x, \alpha) = \int_0^{1-\alpha} g(0, \alpha/(1 - s))f(x, s) \, ds + \int_{1-\alpha}^\infty f(x, s) \, ds. \tag{3.7}$$

Let us make the change of variable $v = xt^{-\frac{1}{2}}$. We then after some computation obtain,

$$g(x, \alpha) = (2/\pi)^{\frac{3}{2}} \int_{x(1-\alpha)^{-\frac{1}{2}}}^\infty \text{arc sin } (\alpha v^2(v^2 - x^2)^{-1})^{\frac{1}{2}} \cdot \exp - v^2/2 \, dv + (2/\pi)^{\frac{1}{2}} \int_0^{x(1-\alpha)^{-\frac{1}{2}}} \exp - v^2/2 \, dv. \tag{3.8}$$

But now by standard formulae of the calculus we find,

$$\partial g(x, \alpha)/\partial x = \alpha^{\frac{1}{2}}x(2/\pi)^{\frac{3}{2}} \int_{x(1-\alpha)^{-\frac{1}{2}}}^\infty v \exp \tfrac{1}{2}v^2[(v^2 - x^2)((1 - \alpha)v^2 - x^2)^{\frac{1}{2}}]^{-1} \, dv. \tag{3.9}$$

[2] A result essentially equivalent to (3.7) and (3.8) is to be found in (4b).

Changing variables to $z = v^2/2 - (x^2/2)(1 - \alpha)^{-1}$ we obtain,

(3.10) $\partial g(x, \alpha)/\partial x$

$$= 2\pi^{-\frac{1}{2}}x \exp - x^2/2(1 - \alpha) \int_0^\infty (\exp - z)z^{\frac{1}{2}}(2(1 - \alpha)z + \alpha x^2)^{-1} dz.$$

After a final change of variable to $v = (s(1 - \alpha)/\alpha)^{\frac{1}{2}}z/x$ we have,

$$\partial g(x, \alpha)/\partial x = (4/\pi)\phi(x)H(\tfrac{1}{2}\alpha/(1 - \alpha)x^2), \tag{3.11}$$

where ϕ is the standard normal density and

$$H(\lambda) = \int_0^\infty (\exp - \lambda(1 + y^2))(1 + y^2)^{-1} dy.$$

Formula (8b) (p. 314) of (1) yields

$$H(\lambda) = \pi(1 - \Phi((2\lambda)^2)) \tag{3.12}$$

where Φ is the standard normal distribution function. We then obtain

$$\partial g(x, \alpha)/\partial x = 4\phi(x)(1 - \Phi(x(\alpha/(1 - \alpha))^{\frac{1}{2}})). \tag{3.13}$$

Of course for $\alpha = \frac{1}{2}$ the expression simplifies and we conclude that,

THEOREM 3.1. *If f is rectangular on* $(-1, 1)$,

$$\begin{aligned} \lim_n P_\theta(n^{\frac{1}{2}}(D_n - \theta) \leqq x) &= 1 - 2(1 - \Phi(x))^2 \quad \textit{for} \quad x > 0, \\ &= 2(1 - \Phi(-x))^2 \quad \textit{for} \quad x < 0. \end{aligned} \tag{3.14}$$

This distribution will be examined more closely in sections 4 and 5.

Now, let $f(x) = \frac{1}{2} \exp -|x|, -\infty < x < \infty$. In this case $\psi(t) = t$ for $0 < t \leqq \frac{1}{2}$ and we have the only other situation in which we are able to give an explicit form to the asymptotic distribution of D_n. By Theorem 2.1 we may write,

$$\lim_n P_\theta(n^{\frac{1}{2}}(D_n - \theta) \leqq x) = P(\lambda(t:Q(t) > 2^{\frac{1}{2}}xt, t \,\varepsilon\, (0, \tfrac{1}{2})) \leqq \tfrac{1}{4}). \tag{3.15}$$

An explicit formula for the right hand side of (3.15) is given by,

LEMMA 3.1. *If $Q(t)$ is standard Brownian motion on $(0, b)$ then,*

$$P(\lambda(t:Q(t) > at, t \,\varepsilon\, (0, b)) \leqq \alpha)$$

$$\begin{aligned} &= \pi^{\frac{1}{2}} \int_0^\infty \int_{(b-\alpha)/\alpha y^2}^\infty (y^2(\alpha - b)/b + \alpha s/b)s^{-\frac{1}{2}}e^{-s}\phi(y + ab^{\frac{1}{2}})\, ds\, dy \\ &\quad + \int_{-\infty}^0 (1 - \pi^{-\frac{1}{2}} \int_{(\alpha y^2/(b-\alpha)}^\infty (-\alpha y^2/b + (b - \alpha)s/b)s^{-\frac{1}{2}}e^{-s}\, ds) \\ &\qquad \cdot\phi(y + ab^{\frac{1}{2}})\, dy. \end{aligned} \tag{3.16}$$

PROOF: Consider the random walk $S_n = \sum_{i=1}^n T_{in}$ where the T_{in} are independent, and

$$\begin{aligned} P(T_{in} = 1) &= \tfrac{1}{2}(1 - a(b/n)^{\frac{1}{2}}), \\ P(T_{in} = -1) &= \tfrac{1}{2}(1 + a(b/n)^{\frac{1}{2}}). \end{aligned} \tag{3.17}$$

Define a stochastic process on $[0, b]$ by,

$$Y_n(t) = S_k b^{\frac{1}{2}}/n^{\frac{1}{2}} \quad \text{on} \quad [(k-1)b/n, kb/n). \tag{3.18}$$

$1 \leqq k \leqq n$ where as usual $S_0 = 0$, $Y_n(b) = S_n b^{\frac{1}{2}}/n^{\frac{1}{2}}$.

It is well known (see e.g. (12)) that $n^{\frac{1}{2}} Y_n(t)$ converges in $\mathcal{P}(D[0, b])$ to $Q(t) - at$. Let N_n denote the number of positive terms among $S_1, \cdots, S_n$. Then,

$$bN_n/n = \lambda(t: n^{\frac{1}{2}} Y_n(t) > 0). \tag{3.19}$$

We conclude from (3.19) and Theorem 7.3 that,

$$P(\lambda(t: Q(t) > at, t \varepsilon (0, b)) \leqq \alpha) = \lim_n P(N_n/n \leqq \alpha/b). \tag{3.20}$$

Now,

$$P(N_n \leqq \alpha n/b) = \textstyle\sum_{k=0}^{n} P(N_n \leqq \alpha n/b \mid S_n = k) P(S_n = k). \tag{3.21}$$

Let M_n be the measure assigning mass $P(S_n = k)$ to $kn^{-\frac{1}{2}}$.

We also define,

$$p_n(x, y) = P(N_n \leqq xn \mid S_n = yn^{\frac{1}{2}}) \tag{3.22}$$

for $yn^{\frac{1}{2}}$ a natural number between 0 and n, and $p_n(x, (n+1)n^{-\frac{1}{2}}) = 0$. More generally let

$$\begin{aligned} p_n(x, y) = p_n(x, (k-1)n^{-\frac{1}{2}}) + n^{\frac{1}{2}}(y - (k-1)n^{-\frac{1}{2}})(p_n(x, kn^{-\frac{1}{2}}) \\ - p_n(x, (k-1)n^{-\frac{1}{2}})), \end{aligned}$$

for $(k-1)n^{-\frac{1}{2}} \leqq y < kn^{-\frac{1}{2}}$, $1 \leqq k \leqq n+1$. For all other values of y we take $p_n(x, y) = 0$. From (3.21) and (3.22) we obtain,

$$P(N_n \leqq \alpha_n/b) = \textstyle\int_{-\infty}^{\infty} p_n(\alpha/b, y)\, dM_n(y). \tag{3.23}$$

But by the central limit theorem and (3.17) M_n converges completely to a normal distribution with mean $-ab^{\frac{1}{2}}$ and variance 1. On the other hand since the conditional distribution of $S_1, \cdots, S_n$ given $S_n = k$ is independent of a and b we may take $a = 0$ and apply the following theorem of Chung and Feller (3), somewhat modified to suit our purposes.

THEOREM. *For fixed x, $p_n(x, y)$ converges uniformly on compacts and,*

$$\begin{aligned} \lim_n p_n(x, y) &= \pi^{-\frac{1}{2}} \textstyle\int_{(1-x)/xy^2}^{\infty} (y^2(x-1) + xs) s^{-\frac{3}{2}} e^{-s}\, ds \\ &\qquad\qquad \text{for} \quad 0 < x < 1, \quad y > 0 \\ &= 1 - \pi^{-\frac{1}{2}} \textstyle\int_{xy^2/(1-x)}^{\infty} (-xy^2 + (1-x)s)^{-\frac{3}{2}} e^{-s}\, ds \\ &\qquad\qquad \text{for} \quad 0 < x < 1, \quad y < 0 \\ &= x \qquad\qquad \text{for} \quad 0 < x < 1, \quad y = 0. \end{aligned} \tag{3.24}$$

For $x = 0$ $p_n(x, y) \to 0$ and for $x = 1$, $p_n(x, y) = 1$ for all y.

Since $\lim_n p_n(x, y)$ is continuous and uniformly bounded for each x as are the $p_n(x, y)$ themselves, upon employing a slight extension of Helly's theorem and (3.20), (3.23), and (3.24) we obtain (3.16), and the lemma is proved.

We now note that,

$$(3.25) \qquad \int_p^\infty z^{-\frac{3}{2}} e^{-z}\, dz = 2p^{-\frac{1}{2}} e^{-p} - 2\int_p^\infty z^{-\frac{1}{2}} e^{-z}\, dz.$$

Let $\gamma(p) = \pi^{-\frac{1}{2}} \int_p^\infty y^{-\frac{1}{2}} e^{-y}\, dy$, the upper tail of the χ_1^2 distribution. Let $H(x) = \lim_n P_\theta(n^{\frac{1}{2}}(D_n - \theta) \leqq x)$. Using (3.15), (3.16), and (3.25) we obtain after some simplification,

$$(3.26) \qquad H(x) = \Phi(x) + 3^{-\frac{1}{2}}\pi^{-\frac{1}{2}} y \exp(x^2/3) + \int_0^\infty (y^2 + \tfrac{1}{2})\gamma(y^2) \cdot(\phi(x + y) - \phi(x - y))\, dy$$

and,

$$(3.27) \qquad H'(x) = \phi(x) + 3^{-\frac{1}{2}}\pi^{-\frac{1}{2}}(\exp(x^2/3) - \tfrac{2}{3}x^2 \exp(-x^2/3)) + \int_0^\infty (y^2 + \tfrac{1}{2})\gamma(y^2)\, d(\phi(x + y) + \phi(y - x)).$$

We now show that $H(x)$ may be expressed in terms of tabled functions.

Remark first,

$$(3.28) \qquad \gamma(y^2) = 2(1 - \Phi(2^{\frac{1}{2}}y)).$$

By integration by parts we can also establish,

$$(3.29) \qquad \int_0^\infty (1 - \Phi(ay))\, d\phi(y + b) = -(\phi(b)/2 + \phi(ab/(1 + a^2)^{\frac{1}{2}})a(1 + a^2)^{-\frac{1}{2}}(1 - \Phi(b(1 + a^2)^{-\frac{1}{2}})),$$

and

$$(3.30) \qquad \int_0^\infty y^2(1 - \phi(ay))\, d\phi(y + b) = 2b\Phi(-b, ab(1 + a^2)^{-\frac{1}{2}}, -a(1 + a^2)^{-\frac{1}{2}}) - \phi(b)(1 + ab(1 + a^2)^{-2}\phi(0)) + a(1 + a^2)^{-\frac{1}{2}} \cdot(1 - \Phi(b(1 + a^2)^{-\frac{1}{2}})(2 + (1 + a^2)^{-2}(1 + a^2 + b^2)).$$

where $\Phi(x, y, \rho)$ is the distribution function of (Z, Z'), bivariate standard normal deviates with correlation ρ.

Substituting $a = 2^{-\frac{1}{2}}$, $b = \pm x$ in (3.29) and (3.30) and employing (3.27) and (3.28) we can now state our final formula,

$$(3.31) \qquad H'(x) = 4x\Phi(-x, x(\tfrac{2}{3})^{\frac{1}{2}}, -(\tfrac{1}{3})^{\frac{1}{2}}) - \Phi(x, -x(\tfrac{2}{3})^{\frac{1}{2}}, -(\tfrac{1}{3})^{\frac{1}{2}}) - 4\phi(x) + 6(\tfrac{2}{3})^{\frac{1}{2}}\phi((\tfrac{2}{3})^{\frac{1}{2}}x).$$

4. Small sample distribution of D_n for a rectangular parent. In this section we assume that $f(x) = \frac{1}{2}$, $-1 \leqq x \leqq 1$, and 0 otherwise. We require some terminology. Let $S_1, \cdots, S_n, \cdots$ denote the sequence of fortunes in a fair coin tossing game, i.e. $S_n = \sum_{i=1}^n Z_i$ where the Z_i are independent and $P(Z_i = 1) = P(Z_i = -1) = \frac{1}{2}$. Let L_n equal $\sum_{i=1}^n (I((S_i > 0)) + I((S_i = 0, S_{i-1} > 0))$, the gambler's lead time. Let T_k equal the first m such that $S_m = k$. It is well known that (cf. Feller (5)),

$$P(L_{2n} = 2k) = \binom{2k}{k}\binom{2(n-k)}{(n-k)}2^{-(2n)} \tag{4.1}$$

$$P(T_k = m) = k/m\binom{m}{\frac{1}{2}(m-k)}2^{-m}. \tag{4.2}$$

In the case that D_n is uniquely defined and n is even, i.e. $n = 4m + 2$, we can express $P_\theta(D_n - \theta) \leqq -a)$ in terms of the quantities defined in (4.1)–(4.3) which may readily be computed. We, of course, have $P_\theta((D_n - \theta) < a) = 1 - P_\theta((D_n - \theta) \leqq a)$. We can now state,

THEOREM 4.1. *If $n = 4m + 2$ then, for $0 < a < 1$, we have,*

$$P_\theta((D_n - \theta) \leqq -a) = \sum_{k=0}^{m} \sum_{r=0}^{k} \sum_{s=0}^{(k-r)} \binom{n}{r} a^r(1-a)^{n-r} \cdot P(T_r = r + 2s)P(L_{n-2(r+s)} = 2(k - (r+s))). \tag{4.3}$$

For $a > 1$ $P((D_n - \theta) \leqq -a) = 0$.

We require a combinatorial lemma proved in (7). Let $a = 2b$ be an even natural number, $y_1 < \cdots < y_r$ be real numbers with different absolute values. We define N to be the number of sums $y_i + y_{a-i+1}$, $1 \leqq i \leqq b$, which are positive. If we now rearrange $y_1, \cdots, y_a$ in order of decreasing absolute value, say, $t_1, \cdots, t_a$ we let $S(q)$ denote the function $\frac{1}{2}(\operatorname{sgn} q + 1)$.

Let $M^+(q) = \sum_{j=1}^{q-1} S(j)$, $M^-(q) = (q - 1) - M^+(q)$, with the usual convention for an empty sum. We shall say there is a positive lead at q if either $M^+(q) > M^-(q)$ or $M^+(q) = M^-(q)$ and $S(q) = 1$. Let $I(q) = 1$ or 0 as there is or is not a positive lead at q. Let $L = \sum_{j=1}^{a} I(j)$ be the total number of positive leads. We state,

LEMMA 4.1 (Hodges). *Under the assumptions given above, $2N = L$.*

PROOF. See (7) and for an earlier result of the same type (6).

We can now prove (4.3). Remark first that, in this case,

$$P_\theta((D_n - \theta) \leqq -a) = P_a(N_n(0) \leqq m). \tag{4.4}$$

We order the Z_i's in decreasing order of absolute value and in accordance with the terminology employed for lemma 4.1 let L_n^* denote the number of positive leads in the corresponding sequence of signs of the Z_i's. We conclude from Lemma 4.1 that,

$$P_a(N_n(0) \leqq m) = P_a(L_n^* \leqq 2m). \tag{4.5}$$

Let K denote the number of Z_i, falling above $(1 - a)$, where $0 \leqq a \leqq 1$. Then,

$$P_a(K = r) = \binom{n}{r} a^r(1-a)^{n-r}. \tag{4.6}$$

Moreover, given $K = r$, $Z_1, \cdots, Z_{n-r}$ are distributed jointly as the order statistics of a sample of $(n - r)$ from a population which is uniformly distributed on $(a - 1, 1 - a)$. We conclude that if, as before, $S(k)$ equals 1 if the kth member in the sequence of ordered absolute values of the Z's is positive and 0 otherwise, then, given $K = r$, $S(r + 1), \cdots S(n)$ are independently and identically distributed with $P_a(S(j) = 1 \mid K = r) = \frac{1}{2}$, for $j > r$. Let T^* be the first q such that $M^+(q+1) = M^-(q+1)$, where M^+ and M^- are defined as before, if such a q exists, and if not let $T^* = n + 1$. Then,

$$P_a(T^* = 2(r+b) \mid K = r) = P(T_r = r + 2b) \tag{4.7}$$

if $2(r+b) \leqq n$. Also we may readily see that,

$$P_a(L^* = 2(r+k) \mid K = r, T^* = 2r + 2b) = P(L_{n-2(r+b)} = 2(k-b)). \tag{4.8}$$

From (4.5)–(4.8)(4.3) follows and the theorem is proved.

The convergence of the small sample distribution to the limiting distribution would seem to be extremely rapid. Thus, if $G_n(x)$ denotes the cumulative distribution of $n^{\frac{1}{2}}(D_n - \theta)$ and $G(x)$ the limit we derived in (3.14) we find d here denotes the Kolmogorov-Smirnov distance and G_n^* is that member of the scale family $G(x/\sigma)$ which is closest to G_n. We remark that convergence to the 'shape' G is much faster than is convergence to the scale.

As a final remark we note that Theorem 3.1 may be proved from Theorem 4.1 of this section.

5. Efficiency of the estimate D_n. As we have seen in Sections 2 and 3 the estimate D_n is typically *not* asymptotically normal. Thus to compare D_n and $\bar{X}$ and D_n and W_n, the median of averages of pairs, in terms of the ratio of asymptotic variances may not be realistic.

The definition of efficiency which we shall employ is a special case of a more general concept of efficiency recently advanced by Hodges and Lehmann.

DEFINITION 5.1. Let $\{T_n\}\{V_n\}$ be two sequences of estimates of θ such that

$$\mathcal{L}_\theta(n^{\frac{1}{2}}(T_n - \theta)) \to G(x), \qquad \mathcal{L}_\theta(n^{\frac{1}{2}}(V_n - \theta)) \to H(x),$$

where H, G are continuous strictly increasing. Then, there exists a unique σ_0 such that $\inf_\sigma d(G(x), H(x/\sigma)) = d(G(x), H(x/\sigma_0))$ where d is the Kolmogorov-Smirnov distance. We define the asymptotic efficiency of $\{T_n\}$ to $\{V_n\}$ to be σ_0^2.

This definition coincides with the usual one when H, G are members of the normal scale family. With this definition we find,

$$\text{eff}\,(D_n, \bar{X}) = 1.0126, \qquad \text{if the parent population is rectangular,} \tag{5.1}$$

and,

$$\text{eff}\,(D_n, \bar{X}) = 1.6057 \qquad \text{for a Laplacian parent} \tag{5.2}$$

TABLE 4.1

n	$d(G_n, G)$	$d(G_n, G_n^*)$
2	.0197	.0197
6	.0100	.0036
10	.00007	.00001
14	.00005	.00001

If we let W_n denote the median of averages of pairs we have,

$$(5.3) \qquad \text{eff}\,(D_n\,,\,W_n) = 1.0126 \qquad \text{for a rectangular population}$$

and

$$(5.4) \qquad \text{eff}\,(D_n\,,\,W_n) = 1.0632 \qquad \text{for a Laplacian parent.}$$

An intuitively reasonable relationship appears if we compare the variance of the normal 'best fitting' to the asymptotic distribution of $n^{\frac{1}{2}}D_n$ to the asymptotic variance of $n^{\frac{1}{2}}W_n$ and that of $n^{\frac{1}{2}}M_n$ where $M_n = \text{median}_i\, X_i$. We find, both in the rectangular case, and in the double exponential case that the asymptotic 'best fit' variance of D_n falls between that of M_n and W_n, and closer to the latter. This supports our feeling that this estimate is an adequate easily computable replacement for W_n possessing all the latter's advantages in a gross error situation. Further evidence for this view appears in numerical studies carried out by J. L. Hodges in (7) on normal samples of size 18. These yielded an estimated efficiency of $.95 \pm .007$ with respect to $\bar{X}$. This number again falls between the asymptotic efficiency of W_n to $\bar{X}$ and that of M_n to $\bar{X}$ and is of course much closer to the former.

6. The One-sample Galton test. The estimate D_n is the mid point of the interval between the upper and lower 50 per cent confidence bounds obtained from the one sample Galton test proposed by Hodges (7). This is a test of $H:\theta = 0$ vs $K:\theta > 0$ in our model, and is given by,

$$\begin{aligned} \phi(x_1\,, \cdots, x_n) &= 1 \quad \text{if} \quad N_n(0) > k \\ &= 0 \quad \text{if} \quad N_n(0) < k. \end{aligned}$$

Roughly speaking we compare the order statistics of $X_1, \cdots, X_n$ with the corresponding order statistics of $-X_1, \cdots, -X_n$ and reject if too many of the former are larger than their correspondents. As was observed in (7) and may easily be seen by applying the probability integral transformation the Galton test is distribution free under the null hypothesis of a continuous distribution symmetric about 0. If $n = 2m$, Lemma 4.1 and (4.1) immediately yield,

$$(6.1) \qquad \begin{aligned} P_0(N_n(0) = k) &= P(L_n = 2k) \\ &= \binom{2k}{k}\binom{2(m-k)}{(m-k)}2^{-n} \qquad \text{for} \quad 0 < k \leqq m \\ &= 0 \quad \text{otherwise;} \end{aligned}$$

and the cut off points may for instance be determined from tables of individual terms of the hypergeometric distribution. Thus, if $H(N, D, n, j) = \binom{D}{j}\binom{N-D}{n-j}/\binom{N}{n}$,

$$(6.2) \qquad P_0(N_n(0) > k) = 2^{-n} \textstyle\sum_{r=k}^{m} H(2m, 2r, m, r)\binom{2m}{m}.$$

For even moderate samples the excellent large sample approximation given by the arcsine law can be used. Thus if $n = 2m$ or $2m + 1$,

$$P_0(N_n(0) > k) \sim 1 - 2/\pi \arcsin (k/m)^{\frac{1}{2}}, \tag{6.3}$$

and the upper α cut off point can be readily computed. Lemma 2.2 can now be translated into a theorem about the asymptotic power of the Galton test for infinitely close alternatives as follows:

THEOREM 6.1. *Let* $\theta_n = xn^{-\frac{1}{2}}$. *Define* $k(\alpha)$ *by* $1 - \alpha = 2/\pi \arcsin k^{\frac{1}{2}}(\alpha)$. *Finally denote by* $\beta_n(\theta, \alpha, f)$ *the power of the level* α *Galton test, randomized if necessary, against the alternative* $f(x - \theta)$. *Then,*

$$\beta_n(\theta_n, \alpha, f) \to P(\lambda(A(\psi, -x)) \geqq k(\alpha)/2).$$

PROOF. The theorem is an immediate consequence of Lemma 2.2 and the identity,

$$P_\theta(N_n(0) \geqq k) = P_0(N_n(-\theta) \geqq k). \tag{6.1}$$

We may similarly interpret the results of Section 3 as providing information about the behaviour of the power function of the Galton test for rectangular and double exponential parents. Of course Section 4 can be used to construct tables of the power function for any even n for a rectangular parent. We finally remark that Theorem 6.1 and the results of Section 3 indicate that the Pitman efficiency of the Galton test with respect to some of the more common tests such as the t or one sample Wilcoxon tests depends not only on the asymptotic power β but also on the asymptotic level of significance. We are again faced here with the consequences of trying to compare distributions with different shapes.

The methods of section 2 also enable us to prove quite easily the consistency of the Galton test in our model. We have,

THEOREM 6.2. *Under the assumptions of Section* 2, $\beta_n(\theta, \alpha, f) \to 1$ *for every* α, $\theta > 0$.

PROOF. It clearly suffices to show that,

$$P_\theta(N_n(0) \geqq k_n(\alpha)) \to 1 \quad \text{where} \quad k_n(\alpha)/m \to k(\alpha).$$

But by (6.4), $P_\theta(N_n(0) \geqq k_n(\alpha)) = P_0(N_n(-\theta) \geqq k_n(\alpha))$ and hence,

$$\begin{aligned} \liminf_n P_\theta(N_n(0) \geqq k_n(\alpha)) &\geqq \sup_A \liminf_n P_0(N_n(-An^{-\frac{1}{2}}) \\ \geqq k_n(\alpha)) &= \sup_A P(\lambda(\psi, -A) \geqq k(\alpha)/2) = 1 \end{aligned}$$

Q.E.D.

REMARK. (1) Obvious modifications of Lemma 2.1 and the above proof will show that the Galton test is consistent for alternatives in which the parent distribution of the X_i's, say G, is such that $G(x) < 1 - G(-x)$ for all x. It may also be shown that the test is inconsistent for some levels for a parent population G such that $G(x) = 1 - G(x)$ on a set of positive Lebesgue measure. The argument may be roughly sketched as follows.

In the general case, $n^{\frac{1}{2}}/2((\hat{Z}_n(t) + \hat{Z}_n(1 - t)) - (G^{-1}(t) + G^{-1}(1 - t)))$ converges in the sense of Prohorov on $[\epsilon, \frac{1}{2}]$ to $V(t) = Z(t)/\psi^*(t) + Z(1 - t)/$

$\psi^*(1 - t)$ where $\psi^*(t) = g(G^{-1}(t))$, g is the density of G, and $Z(t)$ is the Brownian bridge. Then,

$$\liminf_n P(N_n(0) \geqq k_n(\alpha)) \geqq \sup_\epsilon P(\lambda(t: V(t) \geqq -n^{\frac{1}{2}}/2(G^{-1}(t) + G^{-1}(1 - t)), t \in [\epsilon, \tfrac{1}{2}]) \geqq k(\alpha)/2) = 1,$$

if $G(x) < 1 - G(-x)$ for all x.

On the other hand if

$$A = \{t: G^{-1}(t) = G^{-1}(1 - t) t \in (0, \tfrac{1}{2})\}$$

is the union of closed intervals with a non-empty interior, $\psi^*(t) = \psi^*(1 - t)$ for $t \in A$. We can easily conclude that,

$$\lambda(t: V(t) > -n^{\frac{1}{2}}(G^{-1}(t) + G^{-1}(1 - t)), t \in (0, \tfrac{1}{2})) \leqq \lambda(t: Z(t) + Z(1 - t) > 0, t \in (0, \tfrac{1}{2})) + \lambda(t: G^{-1}(t) + G^{-1}(1 - t) > 0, t \in (0, \tfrac{1}{2})).$$

The stated result follows.

(2) Csaki and Vincze in (4) have announced the consistency of the two sample Galton test.

7. Stochastic process convergence. In this section we develop some results necessary for the proof of Theorem 2.1., which may be of independent interest. We recall first some definitions and theorems of Prohorov (12). Let $[a, b]$ be a fixed closed interval. Define $D[a, b]$ to be the set of all functions f on $[a, b]$ such that f^+ and f^- the right and left limits of f exist for every point of $[a, b)$, $(a, b]$ respectively and $f(t) = f^+(t)$ or $f^-(t)$ for all $t \in [a, b]$. Let,

$$\Gamma_f = \{(x, y): y = f^+(x) \quad \text{or} \quad y = f^-(x), x \in (a, b)\} \cup \{(a, f(a)), (b, f(b))$$

Γ_f will be referred to as the graph of f. We identify all functions possessing the same graph. If $A \subset [a, b]$, let $w_f(A)$ equal the upper bound of the absolute values of the difference between the ordinates of points of Γ_f whose abcissae lie in A. Let us then, following Prohorov, define,

$$w_f(\delta) = \sup_{\{\Delta: |\Delta| \leqq \delta\}} w_f(\Delta) \tag{7.1}$$

where $|\Delta|$ is the diameter of Δ. Also, define,

$$\tilde{w}_f(\delta) = \sup_{\{\Delta: |\Delta| \leqq \delta\}} \sup_{\tau \in \Delta} \min w_f([\tau_1, \tau)), w_f((\tau, \tau_2]) \tag{7.2}$$

where Δ ranges over all intervals $(\tau_1, \tau_2) \subset (a, b)$ whose length is $\leqq \delta$. Prohorov has shown that the space of all graphs as above may be transformed into a complete separable metric space, $D(a, b)$, with the following property, (Theorems 2, 4 of Appendix (1) (12)).

THEOREM 7.1. *$f_n \to f$ (in the Prohorov metric) if and only if,*

(1) $\Gamma f_n \subset Q$ *for all n,*

(2) $\tilde{w}_{f_n}(\delta) \leqq w(\delta), \quad 0 \leqq \delta \leqq b - a$ *for all n,*

(3) $f_n^{\pm}(t) \to f(t)$ *at all points of continuity of* f. Q *here is a fixed rectangle and* $w(\delta)$ *a function on* $(0, b - a)$ *such that* $\lim_{\delta \to 0} w(\delta) = 0$.

Let $\mathcal{P}(D(a, b))$ denote the class of all probability measures on $D(a, b)$ endowed with the Borel σ field. Then, Prohorov has established the following,

THEOREM 7.2. *Let* $R_n(t)$ *be a sequence of processes* on (a, b) *whose sample functions are with probability* 1 *members of* $D(a, b)$ *and let* μ_n *be the corresponding probability measures on* $D(a, b)$. *Then* $R_n \leftrightarrow \mu_n$ *converges weakly to* $R \leftrightarrow \mu$ *if and only if,*

(1) $\mathcal{L}(R_n(t_1), \cdots, R_n(t_k)) \to \mathcal{L}(R(t_1), \cdots, R(t_k))$ *for all* $t_1, \cdots, t_k \in [a, b]$.

(2) $\mu_n(f: |f| \leqq M_\epsilon) \geqq 1 - \epsilon$ *for all* $n, \epsilon > 0$, *some* $M_\epsilon < \infty$.

(3) $\mu_n(f: \tilde{w}_f(\delta) \leqq w(\delta, \epsilon)$ for all $\delta \leqq \delta_\epsilon) \geqq 1 - \epsilon$ *for all* $n, \epsilon > 0$, *for a function* $w(\delta, \epsilon)$ *which* $\downarrow 0$ *as* $\delta \downarrow 0$ *for all* $\epsilon > 0$.

Let Y_{kn} be a double sequence of random variables. Define,

$$Y_n^*(t) = Y_{kn} \quad \text{on} \quad [t_{(k-1)n}, t_{kn}), \qquad 1 \leqq k \leqq n,$$

where $t_{0n} = a$, $t_{nn} = b$, $Y_n^*(b) = Y_{nn}$. Similarly define,

$$Y_n(t) = Y_{(k-1)n} + (t - t_{(k-1)n})(t_{kn} - t_{(k-1)n})^{-1}(Y_{kn} - Y_{(k-1)n})$$

on $[t_{(k-1)n}, t_{kn})$ where again $Y_n(b) = Y_{nn}$, $t_{0n} = a$, $t_{nn} = b$. Then $Y_n^*(t)$, $Y_n(t)$ correspond respectively to members of $\mathcal{P}(D[a, b])$ and $\mathcal{P}(C[a, b])$ the set of all probabilities on the set of all continuous functions with uniform norm. We shall say that $Y_n^*(t)$ converges weakly to $Y(t) \in \mathcal{P}(C[a, b])$ if $Y_n^*(t)$ converges weakly in $\mathcal{P}(D[a, b])$ to a measure yielding the same finite dimensional distributions as $Y(t)$ and putting probability 1 on $C[a, b]$ considered as a closed subset of $D[a, b]$. We then have the following.

LEMMA 7.3. *Suppose that* $\sup_{1 \leqq k \leqq n} (t_{(k+1)n} - t_{kn}) \to 0$. *If* $Y_n(t)$ *converge weakly to* $Y_n(t)$ *in* $\mathcal{P}(C[a, b])$ *then* $Y_n^*(t)$ *converges weakly to* $Y(t)$ *in* $\mathcal{P}(D[a, b])$. *Conversely if* $Y_n^*(t)$ *converges weakly to* $Y(t) \in \mathcal{P}(D[a, b])$ *such that* $P(Y(t) \in C[a, b]) = 1$ *there exists* $Y(t) \in \mathcal{P}(C[a, b])$ *such that* $Y_n(t)$ *converges weakly to* $Y(t)$.

PROOF. Since $Y_n(t)$ converges weakly to Y, $\mathcal{L}(\sup_t |Y_n(t)|) \to \mathcal{L}(\sup_t |Y(t)|)$. But $\sup_t |Y_n(t)| = \sup_t |Y_n^*(t)|$, and property 3 of Theorem 7.2 follows.

To establish property (2) remark first that,

$$\tilde{w}_{Y_n^*}(\delta) \leqq \sup_{\{|t_k - t_b| \leqq \delta\}} |Y_{kn} - Y_{bn}| = w_{Y_n}(\delta). \tag{7.1}$$

Since the Y_n converge, by Lemma 2.1 of Prohorov there exists a $w(\delta, \epsilon)$ such that $w(\delta, \epsilon) \downarrow 0$ as $\delta \downarrow 0$ for which,

$$P(w_{Y_n}(\delta) \leqq w(\delta, \epsilon) \quad \text{for all} \quad \delta) \geqq 1 - \epsilon \tag{7.2}$$

and hence

$$\begin{aligned} P(\tilde{w}_{Y_n^*}(\delta) \leqq w(\delta_\epsilon, \epsilon) \quad \text{for all} \quad \delta \leqq \delta_\epsilon) &\geqq P(w_{Y_n}(\delta) \leqq w(\delta_\epsilon, \epsilon)) \\ &\geqq P(w_{Y_n}(\delta) \leqq w(\delta, \epsilon) \quad \text{for all} \quad \delta) \geqq 1 - \epsilon. \end{aligned} \tag{7.3}$$

(2) follows.

To prove that property (1) holds we remark first that $\mathcal{L}(Y_n(t_1), \cdots, Y_n(t_k)) \to \mathcal{L}(Y(t_1), \cdots, Y(t_k))$ for each set $(t_1, \cdots, t_k)$.

Now, $|Y_n(t_i) - Y_n^*(t_i)| \leqq |Y_n(\bar{t}_n(t_i)) - Y_n(\underline{t}_n(t_i))|$ where $\bar{t}_n(x)$ is the smallest $t_{in} \geqq x$, and $\underline{t}_n(x)$ is similarly defined. But by hypothesis $\bar{t}_n(t_i) - \underline{t}_n(t_i) \to 0$ for all i and hence

$$\begin{aligned} &\limsup_n P(|Y_n(t_i) - Y_n^*(t_i)| \geqq \epsilon) \\ &\qquad \leqq \limsup_n P(\sup_{|s-t| \leqq \delta} |Y_n(s) - Y_n(t)| \geqq \epsilon) \\ &\qquad = P(\sup_{|s-t| \leqq \delta} |Y(s) - Y(t)| \geqq \epsilon) \end{aligned} \tag{7.4}$$

for all δ. Since the last quantity goes to 0 we conclude that property (1) holds.

It remains to be shown that the limiting measure $Y^*(t)$ of the Y_n^* processes concentrates on $C[a, b]$. Let $\nu_\epsilon(f)$ be the number of discontinuities of size $\geqq \epsilon$ of $f \varepsilon \mathfrak{D}[a, b]$. Then it is easy to see that ν_ϵ is a continuous function on $D[a, b]$ and hence that

$$P(\nu_\epsilon(Y_n^*) > 0) \to P(\nu_\epsilon(Y^*) > 0). \tag{7.5}$$

But,

$$\lim_n P(\nu_\epsilon(Y_n^*) > 0) \leqq \limsup_n P(\sup_{|s-t| \leqq \delta} |Y_n(s) - Y_n(t)| \geqq \epsilon)$$

for each $\delta > 0$, since $\sup_k |t_{kn} - t_{(k+1)n}| \to 0$. Hence, $P(\nu_\epsilon(Y^*) > 0) = 0$ for all $\epsilon > 0$ and our first conclusion follows. The converse may be proved similarly by checking the conditions of Lemma 2.1, p. 180 of Prohorov [12].

THEOREM 7.4. *Let $Z_n^*(t)$ be defined as in (2.4), $\theta = 0$. Then, $n^{\frac{1}{2}} Z_n^*(t)$ converges weakly in $\mathcal{P}(D[\epsilon, 1 - \epsilon])$ to $(\psi(t))^{-1}Z(t)$ where $Z(t)$ is a Brownian bridge with* a.s. *continuous sample functions.*

PROOF. Let

$$Z_n(t) = Z^*_{(k-1)n} + (t - t_{(k-1)n})(t_{kn} - t_{(k-1)n})^{-1}(Z^*_{kn} - Z^*_{(k-1)n})$$

on $[(k - 1)/n, k/n)$ where $Z^*_{kn} = Z_{kn} - F^{-1}k/(n + 1)$, $Z^*_{0n} = 0$, and $Z_n(1) = Z^*_{nn}$. Then by Theorem 3.1 of Bickel (2) $n^{\frac{1}{2}}Z_n(t)$ converges weakly in $\mathcal{P}(C[\epsilon, 1 - \epsilon])$ to $Z(t)$ as above. Our conclusion now follows from Lemma 7.3.

Let $\{Y_{kn}\}$, $\{M_{kn}\}$ be two double sequences of random variables, $\{t_{kn}\}$ be the associated partitions for both sequences where $\sup (t_{(k+1)n} - t_{kn}) \to 0$. We define $(Y_n(t), M_n(t))$, $Y_n^*(t)$, $M_n^*(t)$, to be the corresponding processes (measures) on $C[a, b] \times C[a, b]$, $D[a, b] \times D[a, b]$ respectively. The following theorem and remark were suggested by a referee.

THEOREM 7.5. *Let e be a function from $D[a, b] \times D[a, b]$ to $D[a, b]$ such that e is continuous at every point of $C[a, b] \times C[a, b]$. Then, if (Y_n, M_n) converges weakly to $(Y(t), M(t))$ in $\mathcal{P}[C[a, b] \times C[a, b]]$, $e(Y_n^*, M_n^*)$ converges weakly to $e(Y, M)$ in $\mathcal{P}(D[a, b])$.*

PROOF. By arguments similar to those used in Lemma 7.1 it may readily be shown that (Y_n^*, M_n^*) converges weakly to (Y, M) where (Y, M) concentrates on $C[a, b] \times C[a, b]$. The theorem follows by a slight extension of Prohorov's

Theorem 1.8 to mappings from R to any separable metric space rather than just the real line.

REMARK. It follows from Theorem 1 of appendix 1 in [12] that $e(f, g) = \alpha f + \beta g$, $e(f, g) = fg$ etc., satisfy the given hypotheses. Le Cam has indicated in [11] that these mappings are not continuous on the whole of $D[a, b] \times D[a, b]$.

COROLLARY 7.6. *If* $\theta = 0$, $n^{\frac{1}{2}}(Z_n^*(s) + Z_n^*(1 - s))$ *converges weakly in* $\mathcal{P}(D[\epsilon, \frac{1}{2}])$ *to* $(\psi(t))^{-1}V(t)$ *where* V *is Brownian motion on* $[0, \frac{1}{2}]$ *with* $\sigma^2 = \frac{1}{2}$.

PROOF. By employing (7.4) for the process $n^{\frac{1}{2}}Z_n(t)$ we can reduce consideration to even n.

Let $C_\epsilon^* = \{(f, g): f, g \text{ continuous on } [\epsilon, \frac{1}{2}), f(\frac{1}{2}) = g(\frac{1}{2})\}$. Since C_ϵ^* is a closed subset of $C[\epsilon, \frac{1}{2}] \times C[\epsilon, \frac{1}{2}]$ and the process $n^{\frac{1}{2}}(Z_n(s), Z_n(1 - s))$ concentrates on C_ϵ^* for $\epsilon \leqq s \leqq \frac{1}{2}$, to show that $n^{\frac{1}{2}}(Z_n(s), Z_n(1 - s))$ converges in $\mathcal{P}(C[\epsilon, \frac{1}{2}] \times C[\epsilon, \frac{1}{2}])$ it suffices to show convergence in $\mathcal{P}(C_\epsilon^*)$. But, the map

$$
\begin{aligned}
Q:(f, g) \to h \quad \text{where} \quad h(s) &= f(s), \qquad \epsilon \leqq s \leqq \tfrac{1}{2}, \\
&= g(s), \qquad \tfrac{1}{2} \leqq s \leqq 1 - \epsilon,
\end{aligned}
$$

is homeomorphic from C_ϵ^* to $C[\epsilon, 1 - \epsilon]$ and $1 - 1$.

Then if r is continuous from C_ϵ^* to R and we define $\hat{r}: C[\epsilon, 1 - \epsilon] \to R$ by $\hat{r}(h) = r(Q^{-1}(h))$ $\hat{r}$ is continuous and $r(n^2 Z_n(s), n^{\frac{1}{2}}Z_n(1 - s)) = \hat{r}(n^{\frac{1}{2}}Z_n(s))$. The domain of the arguments of r and $\hat{r}$ is $[\epsilon, \frac{1}{2}]$ and $[\epsilon, 1 - \epsilon]$ respectively.

Clearly,

$$
\begin{aligned}
(7.6) \qquad \mathcal{L}(\hat{r}(n^{\frac{1}{2}}Z_n(s))) &\to \mathcal{L}(\hat{r}((\psi(s))^{-1}Z(s))) \\
&= \mathcal{L}(r((\psi(s))^{-1}Z(s), (\psi(1 - s))^{-1}Z(1 - s))).
\end{aligned}
$$

By Definition 2.1 this suffices to show that $n^{\frac{1}{2}}(Z_n(s), Z_n(1 - s))$ converges to $((\psi(s))^{-1}Z(s), (\psi(1 - s))^{-1}Z(1 - s))$. The corollary now follows from Theorem 7.5 upon remarking that $\psi(s) = \psi(1 - s)$.

The final theorem of this section deals with the functional we are interested in considering in Section 2.

THEOREM 7.6. *Let* H *be the subset of* $C[a, b]$ *defined by* $H = \{f: \lambda(t: f(t) = 0) = 0\}$. *Then* H *is a measurable subset of* $D[a, b]$ *and the map* $f \to \lambda(t: f(t) > 0)$ *is continuous on* H.

PROOF. Let $\underline{T}(x)$ be the indicator of the set 0. Then there exists a sequence of continuous uniformly bounded functions ψ_n such that $\psi_n(x) \to \underline{T}(x)$. The maps $\psi_n: f \to \int_a^b \psi_n(f(t))\, dt$ are continuous on $D[a, b]$ for each fixed n. But $\psi_n(f) \to \lambda(t: f(t) = 0)$ as $n \to \infty$ and we may conclude that H is measurable.

Define,

$$
\begin{aligned}
\bar{\psi}(x, \epsilon) &= 0, & x &< -\epsilon, \\
(7.7) \qquad &= -x/\epsilon + 1, & \epsilon &< x \leqq 0, \\
&= 1, & x &> 0,
\end{aligned}
$$

and,

$$\begin{aligned} \underline{\psi}(x, \epsilon) &= 0, && x < 0 \\ &= x/\epsilon, && 0 < x < \epsilon \qquad (7.8) \\ &= 1, && x > \epsilon. \end{aligned}$$

Then, if $f_m \to f$ in $D[a, b]$, where $f \varepsilon H$, by Theorem 2 of (12) $f_m(x) \to f(x)$ for all x and hence,

$$\int_a^b \bar{\psi}(f_m(t), \epsilon)\, dt \to \int_a^b \bar{\psi}(f(t), \epsilon)\, dt. \tag{7.9}$$

A similar statement holds for $\underline{\psi}$.

Now, by (7.9) and (7.10), for any g,

$$\int_a^b \underline{\psi}(g(t), \epsilon)\, dt \leqq \lambda(t{:}g(t) > 0) \leqq \int_a^b \bar{\psi}(g(t), \epsilon)\, dt. \tag{7.10}$$

Combining (7.9) and (7.10) we find that,

$$\begin{aligned} &|\lim (\inf)(\sup)_m \lambda(t{:}f_m(t) > 0) - \lambda(t{:}f(t) > 0)| \qquad (7.11) \\ &\qquad \leqq \int_a^b (\bar{\psi}(f(t), \epsilon) - \underline{\psi}(f(t), \epsilon))\, dt. \end{aligned}$$

But by (7.7) and (7.8)

$$\int_a^b (\bar{\psi} - \underline{\psi}) \leqq \lambda(t{:}|f(t)| \leqq \epsilon) \to \lambda(t{:}f(t) = 0),$$

as $\epsilon \to 0$. Since $f \varepsilon H$, the theorem follows.

Acknowledgment. We should like to thank the referees of this paper for some helpful corrections.

REFERENCES

[1] Grobner, W. and Hofreiter, N. (1961). *Integraltafeln* (Part II). Springer-Verlag, Vienna.

[2] Bickel, P. J. (1965). Some contributions to the theory of order statistics. (To appear in *Proc. Fifth Berkeley Symp. Prob. Statist.*)

[3] Chung, K. L. and Feller, W. (1949). Fluctuations in coin tossing. *Proc. Nat. Academy of Sciences, U.S.A*: **35** 605–608.

[4] (a) Csaki, E. and Vincze, I. (1961). On some problems connected with the Galton test. *Publ. Math. Inst. Hung. Acad. Sci.* **6** 97–109.

[4] (b) Csaki, E. and Vincze, I. (1963). On some distributions connected with the arcsine law. *Publ. Math. Inst. Hung. Acad. Sci.* **8** 281–291.

[5] Feller, W. (1957). *An Introduction to Probability Theory and its Applications* (Second Edition). Wiley, New York.

[6] Hodges, J. L., Jr. (1955). Galton's rank order test. *Biometrika*, **42** 261–262.

[7] Hodges, J. L., Jr. (1965). Efficiency in normal samples and tolerance of extreme values for some estimates of location. (To appear in *Proc. Fifth Berkeley Symp. Prob. Statist.*).

[8] Hodges, J. L., Jr. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611.

[9] Høyland, A. (1964). Numerical Evaluation of Hodges-Lehmann estimates. *Det. Kongelige Norske Videnskabers Selskabs Forhandlinger* **37** 42–47.

[10] LEVY, P. (1954). *Le Mouvement Brownien.* Gauthier-Villars Paris.
[11] LECAM, L. (1958). Remarques sur les variables aleatoires dans les espaces vectoriels non separables. *Publ. de l'Inst. de Stat. de l'Univ. de Paris* **7** 39–53.
[12] PROHOROV, Y. (1956). Convergence of random processes and limit theorems in probability. *Theor. Prob. Appl.* **1** 157–214.
[13] VINCZE, I. (1965). On some questions connected with two sample tests of Smirnov type. (To appear in *Proc. Fifth Berkeley Symp. Prob. Statist.*).