The Annals of Applied Probability
1994, Vol. 4, No. 2, 478–493

ON THE ASYMPTOTIC DISTRIBUTION OF THE AREA OUTSIDE A RANDOM CONVEX HULL IN A DISK[1]

BY TAILEN HSING

Texas A & M University

The asymptotic distribution of the area V_n outside the convex hull of n i.i.d. points uniformly distributed on the two-dimensional unit disk is studied. The asymptotic variance of V_n is found to be of the order $n^{-5/3}$, and the asymptotic distribution of V_n is shown to be normal. The results are obtained by carefully analyzing the strength of dependence between sample points at different locations close to the boundary of the unit disk.

1. Introduction. The stochastic properties of the convex hull formed by randomly distributed points in a disk or polygon have been studied by many. The problem goes back to Rényi and Sulanke (1963), where limiting expressions for the expected area, the perimeter and the number of vertices of the convex hull were derived. Papers that investigate variations of these problems are numerous since then [see Groeneboom (1988) and the references therein]. In this context, Groeneboom (1988) contains the best results so far on the distributional properties of the number of vertices of the convex hull. He considers a process running through the vertices of the boundary of the convex hull. After suitable normalization, the process converges in distribution to a strongly mixing Markov process. Using a strong approximation argument and making use of the strong mixing property of the limit process, the number of vertices of the convex hull is shown to be asymptotically normally distributed.

In this paper, we investigate the asymptotic variance and asymptotic distribution of the area outside a random convex hull in a disk in $\mathbb{R}^2$. Our method can be modified to suit the situations where the two-dimensional disk is replaced by a higher-dimensional disk or a polygon, but investigations of those situations are not included in the present paper. Similar to Groeneboom (1988), our results are based on the intuitively clear notion that groups of sample points close to distinct fixed locations on the unit circle are in some sense asymptotically independent. However, our approach is fundamentally different from that of Groeneboom (1988) and is, at least for the present context, more straightforward. For example, we do not use arguments involving strong approximation, and we analyze specifically for each sample size n how the dependence between sample points at different locations close to the unit circle weakens as separation in distance increases. As a result, we

Received January 1992; revised August 1993.
[1]Research supported by NSF Grant 91-07507 and ONR Grant N00014-92-J-1007.
AMS 1991 *subject classifications*. 60D05, 60F05.
Key words and phrases. Asymptotic distribution, convex hull, point processes.

obtain the asymptotic variance of the area directly instead of the variance of the asymptotic distribution.

To explain the basic ideas of the approach we first introduce some notation. Let the sample (X_i, Y_i), $1 \le i \le n$, be i.i.d. random vectors uniformly distributed on the unit disk. Let E_n be the convex hull of the sample, namely, the smallest convex set containing the sample. The random quantity whose asymptotic distribution we are interested in is

$$V_n := \text{area of}\left(E_n^c \cap \{(x, y): x^2 + y^2 \le 1\}\right).$$

For each finite $t \in [0, 2\pi)$, define

$$D_n(t) := \text{length of}\left(E_n^c \cap \{(r\cos t, r\sin t): 0 \le r \le 1\}\right).$$

By symmetry, $D_n(t)$, $0 \le t < 2\pi$, is a stationary process. For the most part, our investigation of V_n is done through studying $D_n(t)$, $0 \le t < 2\pi$, using the simple fact that

$$V_n = \int_{t=0}^{2\pi}\int_{r=1-D_n(t)}^{1} r\,dr\,dt = \int_{t=0}^{2\pi} D_n(t)\,dt - \frac{1}{2}\int_{t=0}^{2\pi} D_n^2(t)\,dt. \tag{1.1}$$

It will be seen that the term $(1/2)\int_{t=0}^{2\pi} D_n^2(t)\,dt$ in this identity is asymptotically negligible, and hence the investigation of V_n can be made by focusing on the properties of $\int_{t=0}^{2\pi} D_n(t)\,dt$ alone.

It is clear that, for fixed $0 \le t < 2\pi$, $D_n(0)$ and $D_n(t)$ are asymptotically independent as $n \to \infty$. The time scale in which the local correlation of D_n can be adequately reflected turns out to be $n^{-1/3}\cdot$. Thus the first step in the variance calculation of V_n is to find the asymptotic covariance of $D_n(0)$ and $D_n(n^{-1/3}a)$, for all $a > 0$. To do that we consider, for $a \ge 0$, the joint asymptotic distribution of the sample points (X_i, Y_i) near $(\cos(n^{-1/3}a), \sin(n^{-1/3}a))$. This is done in a point process theoretic framework. For each $a \in [0, 2\pi n^{1/3})$, we define a two-dimensional point process $\xi_n(a)$ whose points are the properly normalized (X_i, Y_i), which reflects the congregation of points in the neighborhood of $(\cos(n^{-1/3}a), \sin(n^{-1/3}a))$. Thus $\xi_n(a)$, $a \in [0, 2\pi n^{1/3})$, is a stochastic process whose realizations are measure valued. Then a very simple representation of the joint asymptotic distribution of $\xi_n(a_i)$, $1 \le i \le k$, can be obtained for all a_i, $1 \le i \le k$, and all k. Applying the continuous mapping theorem, the asymptotic expression of the product moments, including the mean and covariance of $D_n(n^{-1/3}a)$, are obtained. This, together with other preliminary results, is the content of Section 2. To obtain the asymptotic variance and distribution of V_n, the results in Section 2 are not sufficient. For the variance calculation, it requires, for each fixed sample size n, a close analysis of the rate at which the correlation of $D_n(0)$ and $D_n(n^{-1/3}a)$ decays to 0 as a becomes large. Then, by a standard uniform integrability argument, the results in Section 2 readily provide the form of the asymptotic variance. This is done in Section 3. Having obtained the asymptotic variance of V_n, a blocking method is used in Section 4 to show that V_n is asymptotically normally distributed.

2. Preliminaries. The primary goal of this section is to derive the limiting distribution of the random vector $(n^{2/3}D_n(n^{-1/3}a_1), \dots, n^{2/3}D_n(n^{-1/3}a_k))$ for arbitrary nonnegative constants $a_1, \dots, a_k$. This is achieved through a point process convergence argument. A by-product of this is the convergence of the product moments of $n^{2/3}D_n(n^{-1/3}a)$, $a \ge 0$.

For each $0 \le a < 2\pi n^{1/3}$, let $\xi_n(a)$ be the point process on $[0, \infty) \times (-\infty, \infty)$ consisting of the points

$$\left(n^{2/3}\left(1 - \tilde{X}_i\right), n^{1/3}\tilde{Y}_i/\sqrt{2}\right), \qquad 1 \le i \le n,$$

where

$$\begin{aligned}(2.1)\qquad \left(\tilde{X}_i, \tilde{Y}_i\right) = \big(&X_i \cos(n^{-1/3}a) + Y_i \sin(n^{-1/3}a),\\ &-X_i \sin(n^{-1/3}a) + Y_i \cos(n^{-1/3}a)\big);\end{aligned}$$

namely, $(\tilde{X}_i, \tilde{Y}_i)$ is (X_i, Y_i) rotated through an angle of $-n^{-1/3}a$.

For the theory of point processes, we follow Kallenberg (1983). Note that, since $X_i^2 + Y_i^2 \le 1$, we have $|n^{1/3}Y_i/\sqrt{2}| \le \sqrt{n^{2/3}(1 - X_i)}$. That is, each point (P_i, Q_i) of $\xi_n(0)$, and hence $\xi_n(a)$ for all a by symmetry, satisfies the constraint

$$(2.2)\qquad |Q_i| \le \sqrt{P_i}\,.$$

Let $\mathscr{M}([0, \infty) \times (-\infty, \infty))$ be the space of locally finite counting measures on $[0, \infty) \times (-\infty, \infty)$ having the vague topology and the Borel σ-field generated by that topology. Let $\mathscr{M}$ be the subspace of counting measures whose points satisfy the restriction (2.2). Note that by including the left boundary $x = 0$ in the state space, each counting measure in $\mathscr{M}$ has finitely many points in the set $[0, x] \times (-\infty, \infty)$, for all $0 < x < \infty$. It is not difficult to see that $\mathscr{M}$ is vaguely closed in $\mathscr{M}([0, \infty) \times (-\infty, \infty))$. We regard $\xi_n(a)$ as a random element in $\mathscr{M}$.

THEOREM 2.1. *For arbitrary nonnegative constants $a_1, \dots, a_k$, the point processes $(\xi_n(a_j), 1 \le j \le k)$ converge jointly in distribution to $(\xi(a_j), 1 \le j \le k)$. Here, simultaneously for all $a \ge 0$, $\xi(a)$ has the points*

$$\left(U_i(1 - Z_i^2) + \left(U_i^{1/2}Z_i - \frac{a}{\sqrt{2}}\right)^2, U_i^{1/2}Z_i - \frac{a}{\sqrt{2}}\right), \qquad i \ge 1,$$

where U_i, $i \ge 1$, form a Poisson process on $[0, \infty)$ with intensity measure μ satisfying

$$\mu[0, x] = \frac{4\sqrt{2}}{3}x^{3/2}, \qquad x \ge 0,$$

and the Z_i are i.i.d. uniform$(-1, 1)$, *also independent of the U_i.*

PROOF. We first consider $\xi_n(0)$. The X_i are i.i.d. random variables with

$$P[X_i > 1 - x] = 2\int_{1-x}^{1} \sqrt{1-u^2}\,du = 2\int_0^x \sqrt{2u - u^2}\,du.$$

Hence, for $x_n \to 0$,

$$P[X_i > 1 - x_n] \sim \frac{4\sqrt{2}}{3} x_n^{3/2}.$$

Consequently it is straightforward to show that the point process on $[0, \infty)$ with points $n^{2/3}(1 - X_i)$, $1 \le i \le n$, converges in distribution to the Poisson process with intensity measure μ and points U_i, $i \ge 1$, defined by the theorem. Also it is clear that (X_i, Y_i), $1 \le i \le n$, has the same distribution as $(X_i, \sqrt{1 - X_i^2}\,Z_i)$, $1 \le i \le n$, where the Z_i are i.i.d. uniform $(-1, 1)$ and independent of the X_i. Hence an application of the continuous mapping theorem [cf. Kallenberg (1983)] shows that $\xi_n(0)$ converges in distribution to the point process with points $(U_i, U_i^{1/2} Z_i)$, $i \ge 1$. Now fix $a \ge 0$, and let $(\tilde{X}_i, \tilde{Y}_i)$ be given by (2.1). Then

$$\begin{aligned} n^{2/3}\big(1 - \tilde{X}_i\big) &= n^{2/3}\big(1 - X_i \cos(n^{-1/3}a) - Y_i \sin(n^{-1/3}a)\big) \\ &= n^{2/3}(1 - X_i) + \frac{a^2 X_i}{2}(1 + o(1)) - a n^{1/3} Y_i (1 + o(1)) \end{aligned}$$

and

$$\begin{aligned} \frac{n^{1/3}\tilde{Y}_i}{\sqrt{2}} &= \frac{n^{1/3}}{\sqrt{2}}\big(-X_i \sin(n^{-1/3}a) + Y_i \cos(n^{-1/3}a)\big) \\ &= -\frac{a X_i}{\sqrt{2}}(1 + o(1)) + \frac{n^{1/3} Y_i}{\sqrt{2}}(1 + o(1)). \end{aligned}$$

Thus, by the continuous mapping theorem, in terms of the representation of the distributional limit of $\xi_n(0)$, it is seen that $\xi_n(a)$ converges in distribution to the point process consisting of points

$$\big(U_i + a^2/2 - a(2U_i)^{1/2} Z_i,\, U_i^{1/2} Z_i - a/\sqrt{2}\big), \qquad i \ge 1,$$

which are precisely the points of $\xi(a)$ stated in the theorem. Clearly this representation holds simultaneously for any set of finitely many a's. This concludes the proof. □

COROLLARY 2.2. *The process $\xi(0)$ [and hence $\xi(a)$, for all $a \ge 0$] is a Poisson process on $[0, \infty) \times (-\infty, \infty)$ with intensity measure λ determined by*

$$\lambda(E) = \sqrt{2}\int_E I(|y| < x^{1/2})\,dx\,dy, \qquad E \subset [0, \infty) \times (-\infty, \infty).$$

PROOF. By a standard conditioning argument, for any nonnegative measurable function g on $[0,\infty) \times (-\infty,\infty)$ satisfying $g(x, y) = 0,\ x > B$ for some $B > 0$,

$$
\begin{aligned}
E \exp&\Big[-\sum_i g\big(U_i, \sqrt{U_i} Z_i\big)\Big] \\
&= \sum_{k=0}^{\infty} \exp(-\mu[0,B]) \frac{(\mu[0,B])^k}{k!} \\
&\quad \times \left(\int_{x=0}^{\infty} \int_{y=-1}^{1} \exp[-g(x, \sqrt{x}\, y)] \frac{dy}{2} \frac{\mu(dx)}{\mu[0,B]} \right)^k \\
&= \exp\left[-\int_{x=0}^{\infty} \int_{y=-\sqrt{x}}^{\sqrt{x}} (1 - \exp[-g(x,y)])\, dy \frac{\mu(dx)}{2\sqrt{x}} \right] \\
&= \exp\left[-\int_{x=0}^{\infty} \int_{y=-\infty}^{\infty} (1 - \exp[-g(x,y)]) \lambda(dx\, dy) \right].
\end{aligned}
$$

The last expression is the Laplace transform of the Poisson process with intensity measure λ [cf. Kallenberg (1983)]. Since B is arbitrary, the proof is complete. □

Define a mapping $\Xi\colon \mathscr{M} \to [0,\infty]$ by

$$\Xi\colon \eta \to \inf\{v \ge 0\colon \eta(B_{m,v}) > 0 \text{ for all } m \in (-\infty,\infty)\},$$

where $B_{m,v} = \{(x, y) \in [0,\infty) \times (-\infty,\infty)\colon\ x + my \le v\}$. To illustrate the definition, in Figure 1, if dots represent the points of η, then $\Xi\eta$ is given by the distance between the origin and the intersection of the dashed line and the x axis. Clearly $D_n(a)$ is a function of $\xi_n(a)$. In fact,

$$\Xi \xi_n(a) = n^{2/3} D_n(n^{-1/3} a), \qquad a \ge 0.$$

Observe that Ξ is finite and continuous at each $\eta \in \mathscr{S}$ where

$$\mathscr{S} = \{\eta \in \mathscr{M}\colon \eta \text{ has at least two points whose } y \text{ coordinates have opposite signs}\}.$$

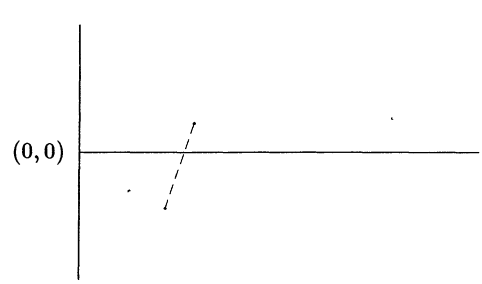

FIG. 1.

By Corollary 2.2, $P[\xi(0) \in \mathscr{S}^c] = 0$. Thus the following result is a consequence of the continuous mapping theorem.

THEOREM 2.3. *For arbitrary nonnegative constants* $a_1, \dots, a_k$,

$$\left(n^{2/3}D_n\left(n^{-1/3}a_j\right), 1 \le j \le k\right) \xrightarrow{d} \left(\Xi\xi(a_j), 1 \le j \le k\right).$$

The distribution of $\Xi\xi(0)$ can be routinely computed. Unfortunately we have not found a simple representation of that distribution.

THEOREM 2.4. *For arbitrary nonnegative constants* $a_1, \dots, a_k$ *and* $s_1, \dots, s_k$,

$$\lim_{n\to\infty} E\left(\prod_{j=1}^{k}\left\{n^{2/3}D_n\left(n^{-1/3}a_j\right)\right\}^{s_j}\right) = E\left(\prod_{j=1}^{k}\left\{\Xi\xi(a_j)\right\}^{s_j}\right) < \infty.$$

PROOF. We will only show that, for any $s > 0$,

$$\lim_{n\to\infty} E\left(\left\{n^{2/3}D_n(0)\right\}^s\right) = E\left(\left\{\Xi\xi(0)\right\}^s\right) < \infty, \tag{2.3}$$

and the same principle applies to the general case in an obvious way. Statement (2.3) follows readily from Theorem 2.3, provided we show that

$$\lim_{\Delta\to\infty}\limsup_{n\to\infty} E\left(\left\{n^{2/3}D_n(0)\right\}^s I\left(n^{2/3}D_n(0) > \Delta\right)\right) = 0. \tag{2.4}$$

A simple observation shows that

$$D_n(0) \le M_n := \left[\bigwedge_{i=1}^{n}(1 - X_i : Y_i \ge 0)\right] \vee \left[\bigwedge_{i=1}^{n}(1 - X_i : Y_i < 0)\right]. \tag{2.5}$$

By the Schwarz inequality,

$$E\left(\left\{n^{2/3}M_n\right\}^s I\left(n^{2/3}M_n > \Delta\right)\right) \le E^{1/2}\left(\left\{n^{2/3}M_n\right\}^{2s}\right)P^{1/2}\left[n^{2/3}M_n > \Delta\right]. \tag{2.6}$$

It is easy to see that $P[n^{2/3}M_n > x]$ decreases exponentially as $x \to \infty$, uniformly in n. Thus

$$\limsup_{\Delta\to\infty}\lim_{n\to\infty} E^{1/2}\left(\left\{n^{2/3}M_n\right\}^{2s}\right)P^{1/2}\left[n^{2/3}M_n > \Delta\right] = 0$$

and (2.3) follows from (2.4)–(2.6). □

In view of the simple identity

$$\begin{aligned} &E(\text{number of vertices of } E_n) \\ &\quad = nEV_{n-1} = 2\pi nED_{n-1}(0) - \pi nED_{n-1}^2(0) \end{aligned} \tag{2.7}$$

Theorem 2.4 provides yet another way of getting a classical result of Rényi and Sulanke (1963). From Theorem 2.4 we also obtain the useful fact

$$\begin{aligned}(2.8)\qquad &\lim_{n\to\infty} \operatorname{cov}\big(n^{2k/3}D_n^k(n^{-1/3}a_1), n^{2k/3}D_n^k(n^{-1/3}a_2)\big)\\ &\quad = \operatorname{cov}\big((\Xi\xi(a_1))^k, (\Xi\xi(a_2))^k\big), \qquad a_1, a_2, k \ge 0.\end{aligned}$$

The following lemma is included for future use.

LEMMA 2.5. *Let $0 < \varepsilon < 1$ be a fixed constant, and let $0 < p_{n1}, p_{n2} < 1$ be constants satisfying $np_{n1}p_{n2} \to 0$. Then there exists a positive constant C such that*

$$\left|\frac{\binom{n}{i}p_{n1}^i(1-p_{n1})^{n-i}\binom{n}{j}p_{n2}^j(1-p_{n2})^{n-j}}{\binom{n}{i,j}p_{n1}^i p_{n2}^j(1-p_{n1}-p_{n2})^{n-i-j}} - 1\right| \le Cnp_{n1}p_{n2},$$

for all n and all integers i, j satisfying $|i/(np_{n1}) - 1| \vee |j/(np_{n2}) - 1| \le \varepsilon$.

PROOF. Straightforward computations give

$$\begin{aligned}&\frac{\binom{n}{i}p_{n1}^i(1-p_{n1})^{n-i}\binom{n}{j}p_{n2}^j(1-p_{n2})^{n-j}}{\binom{n}{i,j}p_{n1}^i p_{n2}^j(1-p_{n1}-p_{n2})^{n-i-j}}\\ &\quad = \left\{\prod_{s=1}^{j}\left(1 + \frac{i}{n}\frac{n}{n-i-s+1}\right)\right\}\frac{(1-p_{n1}-p_{n2})^{i+j}}{(1-p_{n1})^i(1-p_{n2})^j}\\ &\qquad \times\left(\frac{(1-p_{n1})(1-p_{n2})}{1-p_{n1}-p_{n2}}\right)^n.\end{aligned}$$

It is simply seen that, for i, j in the specified range, each of the three factors in the last product differs from 1 by no more than $Cnp_{n1}p_{n2}$ for some constant $C > 0$. This concludes the proof. □

3. Asymptotic variance of V_n. The main result of this section is the following.

THEOREM 3.1. *Let Ξ and $\xi(a)$ be defined as in Section 2. Then*

$$\lim_{n\to\infty} n^{5/3}\operatorname{var}(V_n) = 4\pi\int_{a=0}^{\infty}\operatorname{cov}(\Xi\xi(0), \Xi\xi(a))\,da < \infty.$$

PROOF. For any $k \geq 0$, by stationarity and the fact that $D_n(s) = D_n(2\pi - s)$, we obtain

$$\begin{aligned}\operatorname{var}\left(\int_{t=0}^{2\pi} D_n^k(t)\,dt\right) &= \int_{s=0}^{2\pi}\int_{t=0}^{2\pi} \operatorname{cov}\big(D_n^k(s), D_n^k(t)\big)\,ds\,dt \\ &= 2\int_{s=0}^{2\pi}(2\pi - s)\operatorname{cov}\big(D_n^k(0), D_n^k(s)\big)\,ds \\ &= 2\int_{s=0}^{\pi}(2\pi - s)\operatorname{cov}\big(D_n^k(0), D_n^k(s)\big)\,ds \\ &\quad + 2\int_{s=0}^{\pi} s\operatorname{cov}\big(D_n^k(0), D_n^k(2\pi - s)\big)\,ds \\ &= 4\pi\int_{s=0}^{\pi}\operatorname{cov}\big(D_n^k(0), D_n^k(s)\big)\,ds.\end{aligned}$$

Taking the normalizing constants into account, the result follows readily from Lemma 3.5, (1.1) and (2.8). □

It is interesting to note that since $E(n^{2/3}V_n) \to$ a constant (cf. Section 2), Theorem 3.1 implies that $n^{2/3}V_n$ converges in L_2 to that constant. Suppose one considers the one-dimensional case, namely, where the unit disk is replaced by the unit interval. Then the role of V_n is taken by

$$W_n = \text{sample minimum} + (1 - \text{sample maximum}).$$

There, it is easy to see that the distribution of nW_n converges to a nondegenerate distribution. The difference in degeneracy of the limits for the two cases can be explained by the fact that in the two-dimensional case, V_n is an infinite average of weakly correlated random quantities.

The rest of the section is dedicated to proving Lemma 3.5. Throughout let $\beta \in (0, \frac{2}{3})$ be a fixed constant. First, for $t_1 < t_2 < t_1 + \pi$, define the set

$$R(t_1, t_2) = \left\{(r\cos\theta, r\sin\theta): t_1 < \theta < t_2,\right.$$
$$\left.\cos\left(\frac{t_2 - t_1}{2}\right)\Big/\cos\left(\theta - \frac{t_2 + t_2}{2}\right) \leq r \leq 1\right\}.$$

The set $R(t_1, t_2)$ is nothing but the intersection of the unit disk and the half-plane which does not include the origin and whose boundary is the line going through $(\cos t_1, \sin t_1)$ and $(\cos t_2, \sin t_2)$. It is easily seen that, for $0 < t_n - t_n' \to 0$ as $n \to \infty$,

$$\text{(3.1)} \qquad \text{area of } R(t_n', t_n) = \frac{(t_n - t_n')^3}{12} + O\big((t_n - t_n')^4\big).$$

For $n = 1, 2, 3, \ldots,$ $1 \leq a \leq \pi n^{1/3}$ and $t = 0, a$, define the following. Let $N_n(t; a)$ be the number of points in the set $R(n^{-1/3}(t - a^{\beta/2}), n^{-1/3}(t +$

$a^{\beta/2})$), and let $I_n(a)$ be the set $\{i: |i/(np_n(a)) - 1| \le 1/2\}$ of positive integers where $p_n(a) =$ area of $R(-n^{-1/3}a^{\beta/2}, n^{-1/3}a^{\beta/2})$. Define the events

$$A_n(t;a) = \left(\{(X_i,Y_i)\}_{i=1}^n \cap R\left(n^{-1/3}(t - a^{\beta/2}), n^{-1/3}t\right)\right) \cap \left(\{(X_i,Y_i)\}_{i=1}^n \cap R\left(n^{-1/3}t, n^{-1/3}(t + a^{\beta/2})\right)\right)$$

and

$$B_n(t;a) = \left(N_n(t;a) \in I_n(a)\right).$$

LEMMA 3.2. *For some positive constant C,*

$$\sup_{n\ge 1}\ \sup_{1\le a\le \pi n^{1/3}} \exp\left(\frac{a^{3\beta/2}}{12}\right) P\left[A_n^c(0;a)\right] \le C.$$

PROOF. By (3.1),

$$\text{area of } R(-n^{-1/3}a^{\beta/2}, 0) = \text{area of } R(0, n^{-1/3}a^{\beta/2}) \sim \frac{a^{3\beta/2}}{12n},$$

where the rate of convergence (of the ratio of the two sides to 1) is clearly uniform in $a \in [1, \pi n^{1/3}]$. Thus the probability that none of the (X_i, Y_i) are in $R(-n^{-1/3}a^{\beta/2}, 0)$, which is equal to the probability that none of the (X_i, Y_i) are in $R(0, n^{-1/3}a^{\beta/2})$, is

$$\left(1 - \text{area of } R(-n^{-1/3}a^{\beta/2}, 0)\right)^n \sim \left(1 - \frac{a^{3\beta/2}}{12n}\right)^n \le \exp\left(-\frac{a^{3\beta/2}}{12}\right).$$

The result follows readily from this. □

LEMMA 3.3. *For some positive constant C,*

$$\sup_{n\ge 1}\ \sup_{1\le a\le \pi n^{1/3}} \exp(Ca^{3\beta/2}) P\left[B_n^c(0;a)\right] \le 2.$$

PROOF. By a well-known inequality [cf. Bennett (1962) and Pollard (1984), page 192], for a binomial random variable Z with parameters n and p,

$$\begin{aligned} &P[|Z - np| \ge \lambda] \\ &\quad \le 2\exp\left(-np(1-p)\int_0^{\lambda/np(1-p)} \log(1+x)\,dx\right), \qquad \lambda > 0. \end{aligned} \tag{3.2}$$

However, $N_n(0;a)$ has a binomial distribution with parameters n and $p = p_n(a) \sim (2/3)n^{-1}a^{3\beta/2}$ uniformly in a by (3.1). Thus the conclusion of the result follows in an obvious manner, using (3.2). □

LEMMA 3.4. *For some positive constant C, with $\Delta = 2^{2/(2-\beta)}$,*

$$\sup_{n\ge 1}\ \sup_{\Delta\le a\le \pi n^{1/3}}\ \sup_{i,j\in I_n(a)} \frac{n}{a^{3\beta}}\left|\frac{P[N_n(0;a) = i]P[N_n(a;a) = j]}{P[N_n(0;a) = i,\ N_n(a;a) = j]} - 1\right| \le C.$$

PROOF. Note that, for $\Delta \le a \le \pi n^{1/3}$,

$$\begin{aligned}(3.3)\qquad & R(-n^{-1/3}a^{\beta/2}, n^{-1/3}a^{\beta/2}) \\ & \quad \cap R(n^{-1/3}(a - a^{\beta/2}), n^{-1/3}(a + a^{\beta/2})) = \varnothing.\end{aligned}$$

Thus the result easily follows from Lemma 2.5 with $p_{n1} = p_{n2} = p_n(a)$, where, as in the proof of Lemma 3.3, $p_n(a) \sim (2/3)n^{-1}a^{3\beta/2}$ uniformly for $a \in [1, \pi n^{1/3}]$. □

LEMMA 3.5. *The following holds for any $k \ge 0$,:*

$$\lim_{\Delta\to\infty} \limsup_{n\to\infty} \int_{a=\Delta}^{\pi n^{1/3}} \left|\operatorname{cov}\left(n^{2k/3}D_n^k(0), n^{2k/3}D_n^k(n^{-1/3}a)\right)\right| da = 0.$$

PROOF. Define

$$\tilde{D}_n(t) = n^{2k/3}D_n^k(n^{-1/3}t),$$

and write

$$\begin{aligned}\operatorname{cov}\left(n^{2k/3}D_n^k(0), n^{2k/3}D_n^k(n^{-1/3}a)\right) &= E\tilde{D}_n(0)\tilde{D}_n(a) - \left(E\tilde{D}_n(0)\right)^2 \\ &= \sum_{i=1}^{3} S_{ni}(a),\end{aligned}$$

where

$$\begin{aligned}S_{n1}(a) &= E\tilde{D}_n(0)\tilde{D}_n(a) - E\tilde{D}_n(0)\tilde{D}_n(a)I_{A_n(0;a)\cap B_n(0;a)\cap A_n(a;a)\cap B_n(a;a)} \\ S_{n2}(a) &= E\tilde{D}_n(0)\tilde{D}_n(a)I_{A_n(0;a)\cap B_n(0;a)\cap A_n(a;a)\cap B_n(a;a)} \\ &\quad -\left(E\tilde{D}_n(0)I_{A_n(0;a)\cap B_n(0;a)}\right)^2, \\ S_{n3}(a) &= \left(E\tilde{D}_n(0)I_{A_n(0;a)\cap B_n(0;a)}\right)^2 - \left(E\tilde{D}_n(0)\right)^2.\end{aligned}$$

First,

$$\begin{aligned}|S_{n1}(a)| &= \left|E\tilde{D}_n(0)\tilde{D}_n(a)I_{(A_n(0;a)\cap B_n(0;a)\cap A_n(a;a)\cap B_n(a;a))^c}\right| \\ &\le E\left|\tilde{D}_n(0)\tilde{D}_n(a)\right|\left(I_{A_n^c(0;a)} + I_{B_n^c(0;a)} + I_{A_n^c(a;a)} + I_{B_n^c(a;a)}\right),\end{aligned}$$

which, by the Schwarz inequality, is bounded by

$$2\left(E\tilde{D}_n^4(0)\right)^{1/2}\left\{P^{1/2}[A_n^c(0;a)] + P^{1/2}[B_n^c(0;a)]\right\}.$$

It follows from Theorem 2.4 and Lemmas 3.2 and 3.3 that there exist finite positive constants C_1 and C_2 such that

$$(3.4)\qquad \sup_{n\ge 1} \sup_{\Delta\le a\le \pi n^{1/3}} \exp\left(C_1 a^{3\beta/2}\right)|S_{n1}(a)| \le C_2.$$

Next consider $S_{n2}(a)$. Take $\Delta \ge 2^{2/(2-\beta)}$. Since the (X_i, Y_i) are i.i.d., in view of (3.3) the distributions of the points in $R(-n^{-1/3}a^{\beta/2}, n^{-1/3}a^{\beta/2})$ and those in $R(n^{-1/3}(a - a^{\beta/2}), n^{-1/3}(a + a^{\beta/2}))$ are conditionally independent given that $N_n(0;a)=i$, $N_n(a;a)=j$, for any $i,j\ge 1$. Also observe the crucial fact that on the event $A_n(0;a)\cap A_n(a;a)$, $\tilde{D}_n(0)$ and $\tilde{D}_n(a)$ are determined by the points in $R(-n^{-1/3}a^{\beta/2}, n^{-1/3}a^{\beta/2})$ and $R(n^{-1/3}(a-a^{\beta/2}), n^{-1/3}(a+a^{\beta/2}))$, respectively. Thus, by stationarity,

$$E\big(\tilde{D}_n(0)\tilde{D}_n(a)I_{A_n(0;a)\cap A_n(a;a)}|N_n(0;a)=i, N_n(a;a)=j\big)$$
$$=E\big(\tilde{D}_n(0)I_{A_n(0;a)}|N_n(0;a)=i\big)E\big(\tilde{D}_n(0)I_{A_n(0;a)}|N_n(0;a)=j\big),$$
for all i,j.

We then clearly have

$$|S_{n2}(a)|\le\big(E|\tilde{D}_n(0)|\big)^2\sup_{i,j\in I_n(a)}\left|\frac{P[N_n(0;a)=i, N_n(a;a)=j]}{P[N_n(0;a)=i]P[N_n(a;a)=j]}-1\right|.$$

This together with Lemma 3.4 and Theorem 2.4 give

$$\sup_{n\ge 1}\ \sup_{\Delta\le a\le \pi n^{1/3}} na^{-3\beta}|S_{n2}(a)|\le C_3, \tag{3.5}$$

for some positive constant C_3. Finally, $S_{n3}(a)$ is handled the same way as $S_{n1}(a)$, giving

$$\sup_{n\ge 1}\ \sup_{\Delta\le a\le \pi n^{1/3}} \exp\big(C_4 a^{3\beta/2}\big)|S_{n3}(a)|\le C_5, \tag{3.6}$$

for some finite positive constants C_4 and C_5. Thus, by (3.4)–(3.6) and the fact that $0<\beta<\frac{2}{3}$,

$$\lim_{\Delta\to\infty}\limsup_{n\to\infty}\int_{a=\Delta}^{\pi n^{1/3}}\big|\operatorname{cov}\big(n^{2/3}D_n(0), n^{2/3}D_n(n^{-1/3}a)\big)\big|\,da$$
$$\le\lim_{\Delta\to\infty}\limsup_{n\to\infty}\int_{a=\Delta}^{\pi n^{1/3}}\big(|S_{n1}(a)|+|S_{n2}(a)|+|S_{n3}(a)|\big)\,da=0.$$

This concludes the proof. □

4. A central limit theorem. The main result of this section is the following theorem.

THEOREM 4.1. *As $n\to\infty$, the distribution of $n^{5/6}(V_n-EV_n)$ converges to* Normal$(0,\sigma^2)$, *where $\sigma^2=4\pi\int_{a=0}^{\infty}\operatorname{cov}(\Xi\xi(0),\Xi\xi(a))\,da<\infty$, where Ξ and $\xi(a)$ are as defined in Section* 2.

To prove this result we use a blocking method. Different blocking methods now provide standard tools for proving limit theorems for dependent random quantities. See the collection edited by Eberlein and Taqqu (1986).

Throughout let $\frac{1}{4}<\beta<\frac{1}{3}$ and $0<\varepsilon<1$ be fixed constants. Define

$$t_{n,j}=2\pi j/[n^\beta],$$
$$t'_{n,j}=2\pi(j-\varepsilon)/[n^\beta],\qquad j=0,\dots,[n^\beta]$$

where $[n^\beta]$ is the integer part of n^β. Let $N_{n,j}$ be the number of (X_i, Y_i), $1 \le i \le n$, in $R((t'_{n,j-1} + t_{n,j-1})/2, (t'_{n,j} + t_{n,j})/2)$. Also let

$$I_n = \{i: |i/(np_n) - 1| \le 1/2\},$$

where p_n = area of $R((t'_{n,j-1} + t_{n,j-1})/2, (t'_{n,j} + t_{n,j})/2)$. Define the following events:

$$\begin{aligned} A_{n,j} &= \left(\{(X_i, Y_i)\}_{i=1}^n \cap R\left(\frac{t'_{n,j-1} + t_{n,j-1}}{2}, t_{n,j-1}\right)\right) \\ &\quad \cap \left(\{(X_i, Y_i)\}_{i=1}^n \cap R\left(t'_{n,j}, \frac{t'_{n,j} + t_{n,j}}{2}\right)\right), \\ B_{n,j} &= (N_{n,j} \in I_n), \\ A_n &= \bigcap_{j=1}^{[n^\beta]} A_{n,j}, \qquad B_n = \bigcap_{j=1}^{[n^\beta]} B_{n,j}. \end{aligned}$$

LEMMA 4.2. $\lim_{n\to\infty}\{P[A_n^c] + P[B_n^c]\} = 0$.

PROOF. By the triangle inequality,

$$P[A_n^c] + P[B_n^c] \le [n^\beta](P[A_{n,1}^c] + P[B_{n,1}^c]).$$

The rest of the proof parallels the proofs of Lemmas 3.2 and 3.3 and is omitted. □

LEMMA 4.3.

$$\lim_{n\to\infty} \sup_{i_j \in I_n,\, 1 \le j \le [n^\beta]} \left| \frac{\prod_{j=1}^{[n^\beta]} P[N_{n,j} = i_j]}{P[N_{n,j} = i_j, 1 \le j \le [n^\beta]]} - 1 \right| = 0.$$

PROOF. Clearly, $R((t'_{n,j-1} + t_{n,j-1})/2, (t'_{n,j} + t_{n,j})/2)$, $1 \le j \le [n^\beta]$, are mutually disjoint. By the choice of β it follows from the proof of Lemma 2.5 that there exists a positive constant C such that, for all $n \ge 1$ and $2 \le l \le [n^\beta]$,

$$\text{(4.1)} \qquad \begin{aligned} &\sup_{i_j \in I_n,\, 1 \le j \le l} \left| \frac{P[N_{n,j} = i_j, 1 \le j \le l-1]P[N_{n,l} = i_l]}{P[N_{n,j} = i_j, 1 \le j \le l]} - 1 \right| \\ &\quad \le C(l-1)np_n^2. \end{aligned}$$

Since $n^{2\beta+1}p_n^2 \to 0$, so long as n is greater than or equal to some n_0 we have $Cn^{2\beta+1}p_n^2 < 1$. We will show that, for $n \ge n_0$,

$$\text{(4.2)} \qquad \sup_{i_j \in I_n,\, 1 \le j \le l} \left| \frac{\prod_{j=1}^{k} P[N_{n,j} = i_j]}{P[N_{n,j} = i_j, 1 \le j \le k]} - 1 \right| \le Ck(k-1)np_n^2,$$

for all $2 \le k \le [n^\beta]$, from which the lemma follows readily. We will use an induction argument, as follows: (4.2) holds for $k = 2$ by Lemma 2.5. Suppose now that (4.2) holds for $k \le m - 1$, where $2 \le m - 1 < [n^\beta]$. Observe that, by the triangle inequality,

$$\text{(4.3)} \qquad \begin{aligned} &\left| \frac{\prod_{j=1}^{m} P[N_{n,j} = i_j]}{P[N_{n,j} = i_j, 1 \le j \le m]} - 1 \right| \\ &\quad \le \sum_{l=2}^{m} \frac{\prod_{j=1}^{l-1} P[N_{n,j} = i_j]}{P[N_{n,j} = i_j, 1 \le j \le l-1]} \\ &\qquad \times \left| \frac{P[N_{n,j} = i_j, 1 \le j \le l-1] P[N_{n,l} = i_l]}{P[N_{n,j} = i_j, 1 \le j \le l]} - 1 \right|. \end{aligned}$$

By the induction assumption for $n \ge n_0$ we have

$$\frac{\prod_{j=1}^{l-1} P[N_{n,j} = i_j]}{P[N_{n,j} = i_j, 1 \le j \le l-1]} \le 1 + C(l-1)(l-2)np_n^2 < 2, \qquad 2 \le l \le m.$$

This, together with (4.1) and (4.3), implies

$$\left| \frac{P[N_{n,j} = i_j, 1 \le j \le m]}{\prod_{j=1}^{m} P[N_{n,j} = i_j]} - 1 \right| \le \sum_{l=2}^{m} 2C(l-1)np_{2n} = Cm(m-1)np_n^2,$$

for $n \ge n_0$, proving (4.2) for $k = m$. The proof is complete. □

PROOF OF THEOREM 4.1. By (1.1),

$$\begin{aligned} n^{5/6}(V_n - EV_n) &= n^{5/6} \int_0^{2\pi} (D_n(t) - ED_n(t))\, dt \\ &\quad - \frac{n^{5/6}}{2} \int_0^{2\pi} (D_n^2(t) - ED_n^2(t))\, dt \end{aligned}$$

But, as the proof of Theorem 3.1 showed, the variance of the second term on the right tends to 0, so that the asymptotic distribution of $n^{5/6}(V_n - EV_n)$ is the same as that of $n^{5/6}\int_0^{2\pi}(D_n(t) - ED_n(t))\,dt$. Now write

$$n^{5/6}\int_0^{2\pi}(D_n(t) - ED_n(t))\,dt = \sum_{j=1}^{[n^\beta]} Z_{n,j} + \sum_{j=1}^{[n^\beta]} Z'_{n,j}, \tag{4.4}$$

where

$$Z_{n,j} = n^{5/6}\int_{t_{n,j-1}}^{t'_{n,j}}(D_n(t) - ED_n(t))\,dt,$$

$$Z'_{n,j} = n^{5/6}\int_{t'_{n,j}}^{t_{n,j}}(D_n(t) - ED_n(t))\,dt.$$

The proof below shows that $\sum_{j=1}^{[n^\beta]} Z_{n,j}$ has a limiting distribution Normal(0, $(1-\varepsilon)\sigma^2$), and the same argument proves that $\sum_{j=1}^{[n^\beta]} Z'_{n,j}$ has a limiting distribution Normal(0, $\varepsilon\sigma^2$). In view of (4.4) the result then follows by letting $\varepsilon \to 0$.

We first examine the characteristic function of $\sum_{j=1}^{[n^\beta]} Z_{n,j}$. Write

$$E\exp\left(i\theta\sum_{j=1}^{[n^\beta]} Z_{n,j}\right) = \prod_{j=1}^{[n^\beta]} E\exp(i\theta Z_{n,j}) + S_{n1} + S_{n2} + S_{n3}, \tag{4.5}$$

where

$$S_{n1} = E\exp\left(i\theta\sum_{j=1}^{[n^\beta]} Z_{n,j}\right) - E\exp\left(i\theta\sum_{j=1}^{[n^\beta]} Z_{n,j}\right)I_{A_n\cap B_n}$$

$$S_{n2} = E\exp\left(i\theta\sum_{j=1}^{[n^\beta]} Z_{n,j}\right)I_{A_n\cap B_n} - \prod_{j=1}^{[n^\beta]} E\exp(i\theta Z_{n,j})I_{A_{n,j}\cap B_{n,j}},$$

$$S_{n3} = \prod_{j=1}^{[n^\beta]} E\exp(i\theta Z_{n,j})I_{A_{n,j}\cap B_{n,j}} - \prod_{j=1}^{[n^\beta]} E\exp(i\theta Z_{n,j}).$$

First it follows from Lemma 4.2 that

$$\lim_{n\to\infty}(|S_{n1}| + |S_{n3}|) = 0. \tag{4.6}$$

Next we consider S_{n2}. On the event A_n, it is clear that $Z_{n,j}$ depends only on the (X_i, Y_i) in $R((t'_{n,j-1} + t_{n,j-1})/2, (t'_{n,j} + t_{n,j})/2)$, for each $1 \le j \le [n^\beta]$. To see graphically why this is the case, observe in Figure 2 that if both of the darkened sets contain sample points, then $Z_{n,j}$ depends strictly on sample points that lie on the right of the dashed line.

Notice that $R((t'_{n,j-1} + t_{n,j-1})/2, (t'_{n,j} + t_{n,j})/2)$, $1 \le j \le [n^\beta]$, are mutually disjoint. Thus, using Lemma 4.3, the argument for handling the quantity

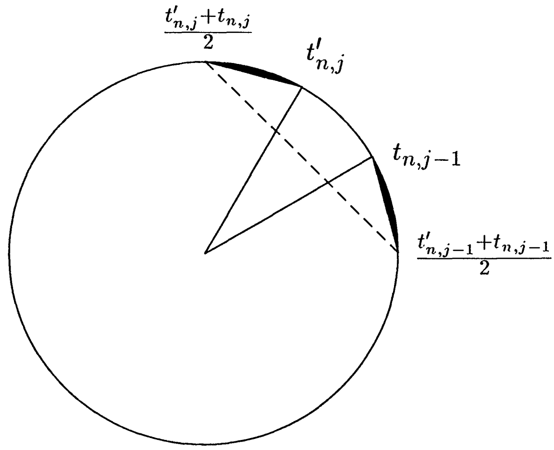

FIG. 2.

$S_{n2}(a)$ in the proof of Lemma 3.5 is readily applicable for showing that $\lim_{n\to\infty}|S_{n2}| = 0$. This, (4.5) and (4.6) imply that

$$E\exp\left(i\theta\sum_{j=1}^{[n^\beta]} Z_{n,j}\right) = \prod_{j=1}^{[n^\beta]} E\exp\bigl(i\theta Z_{n,j}\bigr) + o(1),$$

where the product on the right-hand side is simply the characteristic function of $[n^\beta]$ i.i.d. random variables. Specifically, let $\hat{Z}_{n,j}$, $1 \le j \le [n^\beta]$, be i.i.d. copies of $Z_{n,1}$. It suffices to study the asymptotic distribution of $\sum_{j=1}^{[n^\beta]}\hat{Z}_{n,j}$. For that we pick an arbitrary constant α from $(0, \beta - \frac{1}{6})$ and define additional i.i.d. random variables $\tilde{Z}_{n,j}$, $1 \le j \le [n^\beta]$, such that the joint distribution of $\hat{Z}_{n,j}$ and $\tilde{Z}_{n,j}$ is the same as that of

$$n^{-1/6}\int_0^{2\pi(1-\varepsilon)n^{1/3}/[n^\beta]}\bigl(\tilde{D}_n(u) - E\tilde{D}_n(u)\bigr)\,du$$

and

$$n^{-1/6}\int_0^{2\pi(1-\varepsilon)n^{1/3}/[n^\beta]}\Bigl(\tilde{D}_n(u)I\bigl(\bigl|\tilde{D}_n(u)\bigr| \le n^\alpha\bigr) - E\Bigl(\tilde{D}_n(u)I\bigl(\bigl|\tilde{D}_n(u)\bigr| \le n^\alpha\bigr)\Bigr)\Bigr)\,du$$

where the process $(\tilde{D}_n(u),\ 0 \le u \le 2\pi(1-\varepsilon)n^{1/3}/[n^\beta])$ is distributed the same as $(n^{2/3}D_n(n^{-1/3}u)\colon 0 \le u \le 2\pi(1-\varepsilon)\,n^{1/3}/[n^\beta])$. Since there exist positive constants C_1 and C_2 such that $P[|\tilde{D}_n(0)| > x] \le C_1\exp(-C_2x)$, for all $x > 0$ and $n \ge 1$ (cf. the proof of Theorem 2.4), it follows from the Schwarz inequality and Theorem 2.4 that

$$\lim_{n\to\infty}\sum_{j=1}^{[n^\beta]} E|\hat{Z}_{n,j} - \tilde{Z}_{n,j}| = 0.$$

Thus $\sum_{j=1}^{[n^\beta]}\hat{Z}_{n,j}$ and $\sum_{j=1}^{[n^\beta]}\tilde{Z}_{n,j}$ have the same limiting distribution. Since

$$\operatorname{var}(\tilde{Z}_{n,1}) = 2n^{-1/3}\int_{a=0}^{2\pi(1-\varepsilon)n^{1/3}/[n^\beta]}((1-\varepsilon)2\pi n^{1/3-\beta}-a) \times \operatorname{cov}\Big(\tilde{D}_n(0)I\big(|\tilde{D}_n(0)|\le n^\alpha\big), \tilde{D}_n(a)I\big(|\tilde{D}_n(a)|\le n^\alpha\big)\Big)\,da,$$

the proof of Lemma 3.5 is readily adapted to show that

$$\operatorname{var}(\tilde{Z}_{n,1}) \sim (1-\varepsilon)4\pi n^{-\beta}\int_{a=0}^{\infty}\operatorname{cov}(\Xi\xi(0),\Xi\xi(a))\,da = (1-\varepsilon)\sigma^2 n^{-\beta}. \tag{4.7}$$

By (4.7) and the choice of α,

$$\sum_{j=1}^{[n^\beta]} E|\tilde{Z}_{n,j}|^3 \le [n^\beta]E|\tilde{Z}_{n,1}|^2 n^{-1/6}(2\pi(1-\varepsilon)n^{1/3}/[n^\beta])n^\alpha \to 0 \quad \text{as } n\to\infty.$$

Again making use of (4.7), the Liapounov central limit theorem [cf. Chung (1974), Theorem 7.1.2] implies that $\sum_{j=1}^{[n^\beta]}\tilde{Z}_{n,j}$, and hence $\sum_{j=1}^{[n^\beta]}\hat{Z}_{n,j}$ and $\sum_{j=1}^{[n^\beta]}Z_{n,j}$, all have limiting distribution Normal$(0,(1-\varepsilon)\sigma^2)$. This concludes the proof. □

Acknowledgments. The author sincerely thanks Professor Shaw-Hwa Lo for introducing to him the random convex hull problems and for many enlightening discussions. The author also thanks two referees for their comments and suggestions, which resulted in a better presentation of the paper. The research in this paper was initiated while the author was visiting the Institute of Statistical Science at the Academia Sinica, Taipei, during the summer of 1991. The hospitality and financial support provided by the institute are gratefully acknowledged.

REFERENCES

BENNETT, G. (1962). Probability inequalities for sums of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.

CHUNG, K. L. (1974). *A Course in Probability Theory*. Academic, New York.

EBERLEIN, E. and TAQQU, M. S. (1986). *Dependence in Probability and Statistics*. Birkhäuser, Boston.

GROENEBOOM, P. (1988). Limit theorems for convex hulls. *Probab. Theory Related Fields* **79** 329–368.

KALLENBERG, O. (1983). *Random Measures*. Academic, New York.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

RÉNYI, A. and SULANKE, R. (1963). Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahrsch. Verw. Gebiete* **2** 75–84.

DEPARTMENT OF STATISTICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3143