

RANDOM SPACE FILLING AND MOMENTS OF  
COVERAGE IN GEOMETRICAL PROBABILITY

BY

A. SIEGEL

TECHNICAL REPORT NO. 4

APRIL 4, 1977

PREPARED UNDER GRANT

DAAG29-77-G-0031

FOR THE U.S. ARMY RESEARCH OFFICE

Reproduction in Whole or in Part is Permitted  
for any purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA



RANDOM SPACE FILLING AND MOMENTS OF  
COVERAGE IN GEOMETRICAL PROBABILITY

by

A. Siegel

TECHNICAL REPORT NO. 4

APRIL 4, 1977

Prepared under Grant DAAG29-77-G-0031

For the U. S. Army Research Office

Herbert Solomon, Project Director

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

Partially supported under Office of Naval Research  
Contract N00014-76-C-0475 (NR-042-267) and issued as Technical Report No. 244

THE FINDINGS IN THIS REPORT ARE NOT TO BE  
CONSTRUED AS AN OFFICIAL DEPARTMENT OF  
THE ARMY POSITION, UNLESS SO DESIGNATED  
BY OTHER AUTHORIZED DOCUMENTS.

## INTRODUCTION

Random space filling problems have been discussed in the literature in recent years, (for example [1], [3], [4], [5], and [7]). This is an interesting and usually difficult problem in geometrical probability which in some cases is motivated by specific problems in physical chemistry, biology, and other disciplines. Usually one is concerned with completely covering a fixed region by random regions. However, the specific method by which the random regions are placed is of central importance in the resolution of such problems. Because of the complexity in considering this in any general fashion, an investigator usually chooses special cases to examine. In this paper we consider some situations considered previously by other authors, expand on their analyses, and initiate some new areas for which some solution is found.

In section 1 of this paper, it is shown that the moments of the random variable "proportion of the fixed region covered by random regions" form a monotone decreasing sequence that converges to the probability of coverage. This is shown to hold in great generality and exhibits a link between the method of Robbins [5] for calculating these moments and the random space filling problem.

In Section 2, formulas are developed for the moments of the proportion of a set that is covered by  $N$  independent and identically distributed sets. In Section 3, this is specialized to a particular case of coverage of a circle by random arcs of random size.

Numerical results are reported in Section 4 for probabilities of coverage of the circle by  $N$  uniformly placed random arcs of random sizes. Five arc length distributions are considered and treated by Monte Carlo

and exact methods.

The ordering of coverage probability by decreasing peakedness of arc length distribution, holding expectation constant, is explored in Section 5. This conjectured ordering is suggested by the results of Section 4 and is shown to hold for a large class of arc length distributions for which exact calculations are feasible.

# 1. CONVERGENCE OF THE MOMENTS $E A^k$ TO PROBABILITY OF COVERAGE

Theorem 1 is presented here in full mathematical generality.

For ordinary applications, the reader may wish to think of  $\mathcal{X}$  as a Euclidean space and  $\mu$  as Lebesgue measure on  $\mathcal{X}$ . It is in this context that we apply the theorem in subsequent sections. However, the full generality is presented here because the theorem was developed that way. The proof of theorem 1 is essentially the same in each context.

Let  $(\mathcal{X}, \mathcal{B}, \mu)$  be a measure space. That is,  $\mathcal{B}$  is a  $\sigma$ -algebra of subsets of  $\mathcal{X}$  and  $\mu$  is a countably additive set function mapping  $\mathcal{B}$  to  $[0, \infty]$ . Let  $X$  be a random measurable subset of  $\mathcal{X}$  such that its measure  $\mu(X)$  is a random variable. Let  $K$  be a fixed measurable subset of  $\mathcal{X}$  with  $0 < \mu(K) < \infty$ . Define  $A = \mu(X \cap K) / \mu(K)$ , the proportion of  $K$  that is covered by  $X$ . Note that this framework includes the coverage problem because  $X$  may be, for example, the union of  $N$  independent and identically distributed regions. To avoid measurability problems, we assume  $\mu(X \cap K)$  is a random variable. The event " $X$  covers  $K$ " is defined to mean  $\mu(X \cap K) = \mu(K)$ , that is, if  $X$  covers  $K$  up to a set of measure zero. The main result linking the moments of  $A$  to the probability that  $X$  covers  $K$  is

Theorem 1: Each moment  $E A^m$  ( $m = 1, 2, 3, \dots$ ) is an upper bound on the probability that  $X$  covers  $K$ . The sequence of moments is monotone decreasing,

$$E A \geq E A^2 \geq E A^3 \geq \dots,$$

and converges to the probability that  $X$  covers  $K$  :

$$\lim_{m \rightarrow \infty} E A^m \rightarrow P(X \text{ covers } K) .$$

Moreover, this sequence is strictly decreasing except in the trivial case  $P(0 < A < 1) = 0$  . Thus the probability of coverage is determined in a very simple way from the moments of the proportion of  $K$  covered.

Proof of Theorem 1: Observe that  $A$  is a random variable and thus a measurable function from some probability space  $(\Omega, \mathcal{A}, P)$  to the interval  $[0,1]$  and that  $P(A=1) = P(X \text{ covers } K)$ .  $0 \leq A \leq 1$  implies  $A^m \geq A^{m+1}$ ; hence  $E A^m \geq E A^{m+1}$  and the sequence of moments is monotone decreasing. Next observe that for each  $\omega \in \Omega$  ,  $\lim_{m \rightarrow \infty} A^m(\omega) = I^{(A=1)}(\omega)$  where  $I^{(A=1)}$  is the indicator function of the set  $\{\omega: A(\omega)=1\}$  . Moreover,  $|A(\omega)| \leq 1$  . Thus by the Lebesgue dominated convergence theorem,

$$\lim_{m \rightarrow \infty} E A^m = \lim_{m \rightarrow \infty} \int_{\Omega} A^m(\omega) dP(\omega) = \int_{\Omega} I^{(A=1)}(\omega) dP(\omega) = P(A=1) .$$

Finally, if  $P(0 < A < 1) \neq 0$  then  $P(A^m > A^{m+1}) > 0$  and  $E A^m > E A^{m+1}$  for each positive integer  $m$  . □

As an example to show that measure-theoretic problems can arise, an example is now given in which  $\mu(X \cap K)$  is not measurable, and hence is not a random variable.

Let  $\mathcal{X} = \mathbb{R}$ , the real numbers, let  $K = [0,1]$ , and let  $\mu$  be Lebesgue measure on  $\mathcal{X}$ . Let  $E$  be a nonmeasurable subset of  $[0,1]$ .

For a construction of  $E$  see Royden [6] p. 63. The random set  $X$  is defined in terms of a random point  $\omega$  selected from  $[0,1]$  with uniform measure. We set

$$X = \begin{cases} [0,1] & \text{if } \omega \in E \\ [1,2] & \text{if } \omega \notin E \end{cases}$$

It is clear that the value of  $X$  will always be a measurable subset of  $\mathcal{X}$ , and that  $\mu(X)$  is identically 1 so that  $\mu(X)$  is a (degenerate) random variable. However,  $\mu(X \cap K)$  is not measurable because

$$\mu(X \cap K) = \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{if } \omega \notin E \end{cases}$$

is the indicator function of the nonmeasurable set  $E$ . This is presented here to highlight the measurability aspect referred to earlier and naturally this situation is not expected to occur in applications.

## 2. FORMULAS FOR THE MOMENTS IN THE CASE OF N I.I.D. SETS

Using the method of Robbins [5] the moments of  $A = \mu(X \cap K)/\mu(K)$ , where  $K$  is a fixed set and  $X$  a random set in  $R^n$ ,  $\mu$  is Lebesgue measure, and  $0 < \mu(K) < \infty$ , are given by

$$(2.1) \quad E A^m = [\mu(K)]^{-m} \int_{K^m} P(x_1, \dots, x_m \in X) dx_1 \dots dx_m.$$

To avoid measure-theoretic pathologies, we assume all random quantities we write are random variables. This will be guaranteed, for example, if we work in the framework of coverage spaces. A coverage space, defined by Ailam in [1], is a triple  $(\mathcal{Q}, \mathcal{B}, M)$  formed by two probability spaces  $\mathcal{Q} = (\mathcal{X}, S, P)$  and  $\mathcal{B} = (\mathcal{Y}, T, Q)$  together with a measurable set  $M \subset \mathcal{X} \times \mathcal{Y}$  in their product probability space. A random measurable subset  $M_y$  of  $\mathcal{X}$  is determined by a random point  $y \in \mathcal{Y}$  via the set  $M$  by taking  $M_y = \{x \in \mathcal{X} : (x, y) \in M\}$ .  $\mathcal{Q}$  is called the coverable space, and  $\mathcal{B}$  is called the space of experiments. A concrete illustration that is instructive in viewing this concept of a coverage space is given in the next section.

The integrand  $P(x_1, \dots, x_m \in X)$  is difficult to evaluate for general  $X$ . However, in the case  $X = \bigcup_{i=1}^N X_i$  where  $X_1, \dots, X_N$  are independently and identically distributed sets, the integrand may be expressed in terms of  $x_1, \dots, x_m$  and  $X_1$  alone as follows.

Theorem 2:

$$(2.2) \quad P(x_1, \dots, x_m \in X) = 1 + \sum_{k=1}^m (-1)^k \sum_{1 \leq j_1 < \dots < j_k \leq m} [P(x_{j_1}, \dots, x_{j_k} \notin X_1)]^N$$

where  $X = \bigcup_{i=1}^N X_i$ , and  $X_1, \dots, X_N$  are independent and identically distributed random sets.

From this we immediately obtain a useful formula for the moments.

Corollary:

$$(2.3) \quad E A^m = 1 + \sum_{k=1}^m (-1)^k \binom{m}{k} [\mu(k)]^{-k} \int_{K^k} [P(x_1, \dots, x_k \notin X_1)]^N dx_1 \dots dx_k.$$

Thus, to find the  $m^{\text{th}}$  moment of  $A$ , we need only calculate

$$\int_{K^k} [P(x_1, \dots, x_k \notin X_1)]^N dx_1 \dots dx_k$$

for  $k = 1, 2, \dots, m$ . Moreover, after finding  $E A^m$  in this fashion, we can calculate  $E A^{m+1}$  by doing just one more integral, namely

$$\int_{K^{m+1}} [P(x_1, \dots, x_{m+1} \notin X_1)]^N dx_1 \dots dx_{m+1}.$$

Proof of Theorem 2: Let  $I_Y$  denote the indicator function of the set

$Y$  and let  $Y^c$  denote the complement of  $Y$ . Then

$P(x_1, \dots, x_m \in X) = E I_X(x_1) \dots I_X(x_m)$ . Since  $X = \bigcup_{i=1}^N X_i$ , we have

$I_X(x_j) = 1 - \prod_{i=1}^N I_{X_i^c}(x_j)$ . Thus

$$P(x_1, \dots, x_m \in X) = E \prod_{j=1}^m \left\{ 1 - \prod_{i=1}^N I_{X_i^c}(x_j) \right\}.$$

Expanding the first product, we get

$$E \left\{ 1 + \sum_{k=1}^m (-1)^k \sum_{1 \leq j_1 < \dots < j_k \leq m} \prod_{i=1}^N \left[ I_{X_i^c}(x_{j_1}) \dots I_{X_i^c}(x_{j_k}) \right] \right\} .$$

Taking expectation inside a finite sum and using independence, we get

$$1 + \sum_{k=1}^m (-1)^k \sum_{1 \leq j_1 < \dots < j_k \leq m} \prod_{i=1}^N E \left[ I_{X_i^c}(x_{j_1}) \dots I_{X_i^c}(x_{j_k}) \right] .$$

Theorem 2 now follows from the fact that each  $X_i$  has the same distribution as  $X_1$  and that

$$E I_{X_1^c}(x_{j_1}) \dots I_{X_1^c}(x_{j_k}) = P(x_{j_1}, \dots, x_{j_k} \notin X_1) .$$

□

### 3. SPECIAL CASE: ARCS OF RANDOM SIZES ON A CIRCLE

The circle has held a fascination for a number of authors in connection with either covering the circumference with random arcs, or covering a region with random circles. There is an interesting connection between the two problems which is explored here.

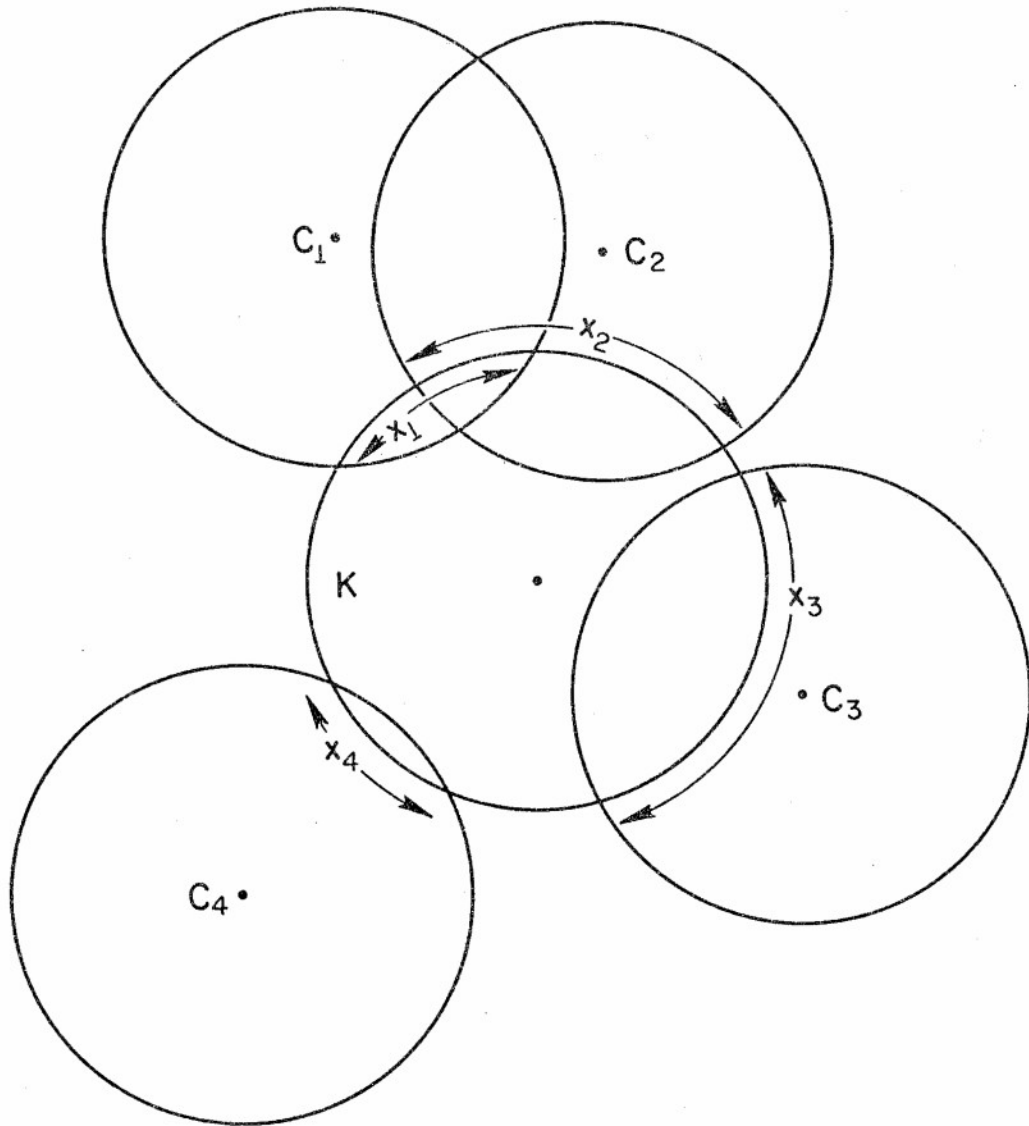
Let  $K$  denote the unit circle. Let  $N$  arcs  $X_1, \dots, X_N$  be placed on  $K$ , with  $X_i$  centered at  $C_i$  radians, and of length  $L_i$ . We assume  $C_1, \dots, C_N$ ,  $L_1, \dots, L_N$  are independent random variables, with each  $C_i$  uniformly distributed on  $[0, 2\pi)$  and each  $L_i$  distributed with a density  $h$  on  $[0, \infty)$ . Let  $A$  denote the proportion of  $K$  covered by  $X = \bigcup_{i=1}^N X_i$ . We wish to find the moments of  $A$ , and hence  $P(X \text{ covers } K)$ .

This problem is motivated by the problem of covering a plane region with random unit discs. Observe that if the region is covered, then each disc whose closure is interior to the region must have its boundary covered by other discs. Let  $K$  denote a fixed unit circle. Place  $N$  unit discs with centers uniformly distributed over the disc of radius 2 concentric with  $K$ . Note each disc determines a random arc of overlap on  $K$  (see Figure 3.1). It can be seen that these arcs have the distribution mentioned in the first paragraph, with  $h(\ell) = \frac{1}{2} \sin \ell \mathbb{I}_{[0, \pi]}(\ell)$ . The probability that the  $N$  discs cover  $K$  is  $P(A=1)$ .

We may place this problem within the framework of coverage spaces as follows. Let  $\mathcal{Y}$  be the surface of the unit sphere in 3 dimensions parametrized by latitude and longitude so that

$\mathcal{Y} = \{(\theta, \phi) : \theta \in [0, \pi], \phi \in [0, 2\pi)\}$ . Let  $M = \{(x, \theta, \phi) \in K \times \mathcal{Y} : x \in [\phi, \theta + \phi] \bmod 2\pi\}$ . Then, as outlined in section 2, a point

Figure 3.1



Each random circle  $C_i$  determines a random arc  $X_i$  on the fixed circle  $K$ .

$(\theta, \phi) \in \mathcal{Y}$  determines a subset  $M_{(\theta, \phi)} \subset K$ . In this case,  $M_{(\theta, \phi)}$  is the arc starting at  $\phi$  and extending counterclockwise with length  $\theta$ . If we place uniform measure on  $\mathcal{Y}$ , a straightforward calculation shows that the resulting distribution of arc length is  $h_0$ . Figure 3.2 may help one visualize the situation.

First we calculate the moments for arbitrary density  $h(\ell)$ . We will use Formula (2.3) as

$$(3.1) \quad E A^m = 1 + \sum_{k=1}^m (-1)^k \binom{m}{k} (2\pi)^{-k} b_k$$

where

$$(3.2) \quad b_k = \int_{K^k} [P(x_1, \dots, x_k \notin X_1)]^N dx_1 \dots dx_k.$$

Note that the integrand above is invariant under permutations of  $(x_1, \dots, x_k)$  and that

$$\int_{K^{k-1}} [P(x_1, \dots, x_{k-1} \notin X_1)]^N dx_1 \dots dx_{k-1}$$

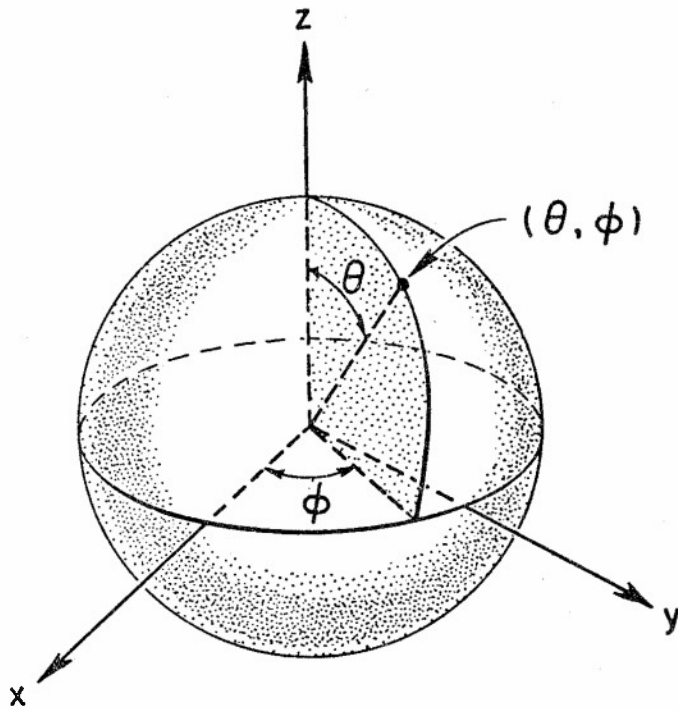
is independent of the value of  $x_k$  by uniformity of  $C_k$ . Thus we may set  $x_k = 0$  and integrate over the ordered variables

$0 \leq x_1 \leq \dots \leq x_{k-1} \leq 2\pi$  to get

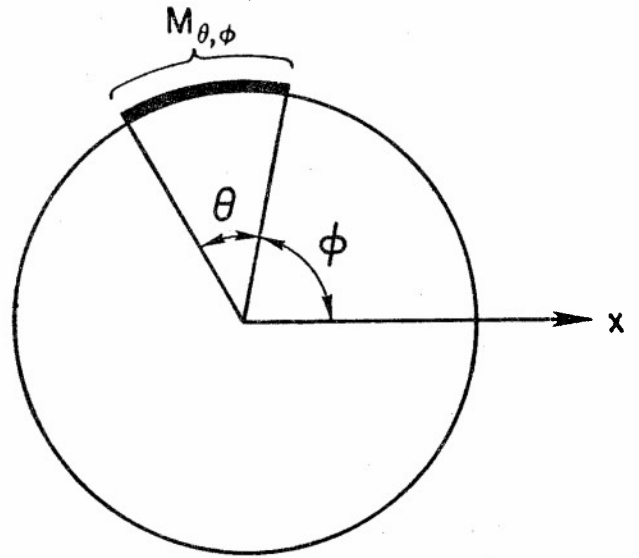
$$(3.3) \quad b_k = (2\pi)(k-1)! \int_{0 \leq x_1 \leq \dots \leq x_{k-1} \leq 2\pi} [P(x_1, \dots, x_{k-1}, 0 \notin X_1)]^N dx_1 \dots dx_{k-1}.$$

Next, transform to  $\alpha_1 = x_1$ ,  $\alpha_2 = x_2 - x_1, \dots, \alpha_{k-1} = x_{k-1} - x_{k-2}$ ,  $\alpha_k = 2\pi - x_{k-1} = 2\pi - \sum_{i=1}^{k-1} \alpha_i$ . Note that

Figure 3.2



$\mathcal{Y}$   
Space of Experiments  
(surface of sphere)



$K$   
Coverable Space  
(circumference of circle)

The random point  $(\theta, \phi) \in \mathcal{Y}$  determines the random arc

$$M_{\theta, \phi} \subset K.$$

$$(3.4) \quad P(x_1, \dots, x_{k-1}, 0 \notin X_1) = \sum_{i=1}^k P(X_1 \text{ is between } 0 \text{ and } \alpha_i),$$

again because  $X_1$  has uniformly distributed center. The Jacobian of the transformation  $(x_1, \dots, x_{k-1}) \rightarrow (\alpha_1, \dots, \alpha_{k-1})$  is 1 so (3.3) becomes

$$(3.5) \quad b_k = (2\pi)(k-1)! \int_R \left[ \sum_{i=1}^k P(X_1 \text{ between } 0 \text{ and } \alpha_i) \right]^N d\alpha_1 \dots d\alpha_{k-1}$$

where  $R = \{(\alpha_1, \dots, \alpha_{k-1}) : \text{all } \alpha_i \geq 0 \text{ and } \sum_{i=1}^{k-1} \alpha_i \leq 2\pi\}$  and  $\alpha_k = 2\pi - \sum_{i=1}^{k-1} \alpha_i$ . Next we calculate  $P(X_1 \text{ between } 0 \text{ and } \alpha)$  for  $0 \leq \alpha \leq 2\pi$ .

$$(3.6) \quad P(X_1 \text{ between } 0 \text{ and } \alpha) = E P(X_1 \text{ between } 0 \text{ and } \alpha | L_1) = E \frac{(\alpha - L_1)_+}{2\pi}$$

where  $(\zeta)_+ = \max(\zeta, 0)$ .  $L_1$  has density  $h$ , and we integrate by parts to obtain

$$(3.7) \quad P(X_1 \text{ between } 0 \text{ and } \alpha) = \frac{1}{2\pi} \int_0^\alpha (\alpha - \ell) h(\ell) d\ell = \frac{1}{2\pi} \int_0^\alpha H(\ell) d\ell$$

where  $H(\ell) = \int_0^\ell h(t) dt$  is the cumulative distribution function of  $L$ .

Using (3.7) in (3.5)

$$(3.8) \quad b_k = (2\pi)(k-1)! \int_R \left[ \frac{1}{2\pi} \sum_{i=1}^k \int_0^{\alpha_i} H(\ell) d\ell \right]^N d\alpha_1 \dots d\alpha_{k-1}$$

For completeness, we put (3.8) into (3.1) to obtain

$$(3.9) \quad E A^m = 1 + \sum_{k=1}^m (-1)^k \binom{m}{k} (k-1)! (2\pi)^{-(k-1)} \int_R \left[ \frac{1}{2\pi} \sum_{i=1}^k \int_0^{\alpha_i} H(\ell) d\ell \right]^N d\alpha_1 \dots d\alpha_{k-1}.$$

Now we specialize to the case  $h_0(\ell) = \frac{\sin \ell}{2} I_{[0,\pi]}(\ell)$  so that

$$H_0(\ell) = \begin{cases} \frac{1 - \cos \ell}{2} & 0 \leq \ell \leq \pi \\ 1 & \ell > \pi \end{cases}$$

and

$$\int_0^\alpha H_0(\ell) d\ell = \begin{cases} \frac{\alpha - \sin \alpha}{4\pi} & \alpha \leq \pi \\ \frac{\alpha}{2\pi} - \frac{1}{4} & \alpha > \pi \end{cases}$$

These different forms on  $[0,\pi]$  and  $(\pi,2\pi]$  complicate matters, but difficulties may be overcome by breaking the region of integration  $R$  up into subregions where some (at most one!)  $\alpha_i$  is greater than  $\pi$ , and where all  $\alpha_i$  are smaller than  $\pi$ . Using this and symmetry considerations, (3.8) becomes

$$(3.10) \quad b_k = (2\pi)(k-1)! \cdot k \int_{R \cap \{\alpha_1 > \pi\}} \left[ \frac{\alpha_1}{2\pi} - \frac{1}{4} + \frac{1}{4\pi} \sum_{i=2}^k (\alpha_i - \sin \alpha_i) \right]^N d\alpha_1 \dots d\alpha_{k-1} \\ + (2\pi)(k-1)! \int_{R \cap \{\text{all } \alpha_j \leq \pi\}} \left[ \frac{1}{4\pi} \sum_{i=1}^k (\alpha_i - \sin \alpha_i) \right]^N d\alpha_1 \dots d\alpha_{k-1}.$$

Recall  $\sum_{i=1}^k \alpha_i = 2\pi$ , so (3.10) simplifies to

$$\begin{aligned}
(3.11) \quad b_k = (2\pi)k! & \int_{R \cap \{\alpha_1 > \pi\}} \left[ \frac{1}{4} + \frac{\alpha_1}{4\pi} - \frac{1}{4\pi} \sum_{i=2}^k \sin \alpha_i \right]^N d\alpha_1 \dots d\alpha_{k-1} \\
& + (2\pi)(k-1)! \int_{R \cap \{\text{all } \alpha_j \leq \pi\}} \left[ \frac{1}{2} - \frac{1}{4\pi} \sum_{i=1}^k \sin \alpha_i \right]^N d\alpha_1 \dots d\alpha_{k-1} .
\end{aligned}$$

For  $k = 1$  , we may use (3.2) directly to obtain

$$(3.12) \quad b_1 = 2\pi \cdot \left(\frac{3}{4}\right)^N .$$

For  $k = 2$  , we use (3.11) and the fact that  $\alpha_1$  and  $\alpha_2 = 2\pi - \alpha_1$  cannot both be less than  $\pi$  to obtain

$$b_2 = 4\pi \int_{\pi}^{2\pi} \left[ \frac{1}{4} + \frac{\alpha_1}{4\pi} - \frac{1}{4\pi} \sin(2\pi - \alpha_1) \right]^N d\alpha_1 .$$

Change variables and simplify to obtain

$$(3.13) \quad b_2 = 4\pi \int_0^{\pi} \left[ \frac{3}{4} - \frac{1}{4\pi} (\alpha_1 + \sin \alpha_1) \right]^N d\alpha_1 .$$

This may easily be evaluated numerically for any particular  $N$  .

The first two moments of  $A$  are thus (using (3.1), (3.12) and (3.13))

$$(3.14) \quad E A = 1 - \left(\frac{3}{4}\right)^N ,$$

$$(3.15) \quad E A^2 = 1 - 2 \cdot \left(\frac{3}{4}\right)^N + \frac{1}{\pi} \int_0^{\pi} \left[ \frac{3}{4} - \frac{1}{4\pi} (\alpha + \sin \alpha) \right]^N d\alpha .$$

The third moment is more of a challenge. From (3.11)

$$(3.16) \quad b_3 = 12\pi \int_{\pi}^{2\pi} \int_0^{2\pi-\alpha_1} \left[ \frac{1}{4} + \frac{\alpha_1}{4\pi} - \frac{1}{4\pi} (\sin \alpha_2 - \sin(\alpha_1 + \alpha_2)) \right]^N d\alpha_2 d\alpha_1 \\ + 4\pi \int_0^{\pi} \int_{\pi-\alpha_1}^{\pi} \left[ \frac{1}{2} - \frac{1}{4\pi} (\sin \alpha_1 + \sin \alpha_2 - \sin(\alpha_1 + \alpha_2)) \right]^N d\alpha_2 d\alpha_1 .$$

For ease of numerical integration, we change variables and combine integrals:

$$(3.17) \quad b_3 = 4\pi \int_0^{\pi} \int_0^{\alpha_1} \left\{ 3 \left[ \frac{3}{4} - \frac{\alpha_1}{4\pi} - \frac{1}{4\pi} (\sin \alpha_2 + \sin(\alpha_1 - \alpha_2)) \right]^N \right. \\ \left. + \left[ \frac{1}{2} - \frac{1}{4\pi} (\sin \alpha_1 + \sin \alpha_2 + \sin(\alpha_1 - \alpha_2)) \right]^N \right\} d\alpha_2 d\alpha_1 .$$

From (3.1)

$$(3.18) \quad E A^3 = 1 - \frac{3}{2\pi} b_1 + \frac{3}{4\pi^2} b_2 - \frac{1}{8\pi^3} b_3 .$$

Some values of  $b_1$ ,  $b_2$ ,  $b_3$ ,  $E A$ ,  $E A^2$ ,  $E A^3$  and variance (A) are tabulated in Table 3.1. Recall from theorem 1 that  $E A$ ,  $E A^2$ , and  $E A^3$  are upper bounds for the probability of coverage.

Table 3.1

| <u>n</u> | <u>b<sub>1</sub></u> | <u>b<sub>2</sub></u> | <u>b<sub>3</sub></u> | <u>E A</u> | <u>E A<sup>2</sup></u> | <u>E A<sup>3</sup></u> | <u>Var (A)</u> |
|----------|----------------------|----------------------|----------------------|------------|------------------------|------------------------|----------------|
| 1        | 4.71                 | 22.67                | 111.3                | .250       | .074                   | .026                   | .0118          |
| 2        | 3.53                 | 13.25                | 52.1                 | .438       | .211                   | .111                   | .0193          |
| 3        | 2.65                 | 7.89                 | 25.4                 | .578       | .356                   | .232                   | .0218          |
| 4        | 1.99                 | 4.78                 | 12.9                 | .684       | .488                   | .361                   | .0211          |
| 5        | 1.49                 | 2.96                 | 6.8                  | .763       | .600                   | .486                   | .0186          |
| 6        | 1.12                 | 1.86                 | 3.7                  | .822       | .691                   | .592                   | .0154          |
| 7        | .84                  | 1.19                 | 2.1                  | .867       | .763                   | .681                   | .0123          |
| 8        | .63                  | .77                  | 1.2                  | .900       | .819                   | .754                   | .0096          |
| 9        | .47                  | .51                  | .71                  | .925       | .863                   | .811                   | .0073          |
| 10       | .35                  | .34                  | .43                  | .944       | .896                   | .855                   | .0055          |
| 15       | .084                 | .052                 | .045                 | .987       | .975                   | .964                   | .0011          |
| 20       | .020                 | .0092                | .0060                | .9968      | .9939                  | .991                   | .00022         |
| 25       | .0047                | .0017                | .00093               | .99925     | .9985                  | .9979                  | .000044        |

#### 4. COVERAGE PROBABILITIES FOR THE CIRCLE: NUMERICAL RESULTS

Consider again the problem of Section 3. Let  $F$  denote the cumulative distribution function of the arc lengths, and let  $P(n,F)$  denote the probability that  $n$  random arcs, with uniformly distributed centers, cover the unit circle. In this section, numerical values for  $P(n,F)$  are calculated for five distributions on  $[0,\pi]$ , which, like the sine distribution of Section 3, are symmetric about  $\pi/2$ .

The distributions studied are as follows:

- (1)  $F_1$ , concentrated at  $\pi/2$ .
- (2)  $F_2$ , with density  $\frac{\sin \ell}{2} I_{[0,\pi]}(\ell)$ . Recall that upper bounds on  $P(n,F_2)$  were found in Section 3.
- (3)  $F_3$ , the uniform distribution on  $[0,\pi]$ .
- (4)  $F_4(\ell) = \begin{cases} (\frac{\ell}{8\pi})^{1/4}, & 0 \leq \ell \leq \pi/2 \\ 1 - (\frac{\pi-\ell}{8\pi})^{1/4}, & \pi/2 < \ell \leq \pi. \end{cases}$

This was chosen as an intermediary between  $F_3$  and  $F_5$ .

- (5)  $F_5$ , placing mass  $1/2$  at  $0$  and at  $\pi$ .
- $F_1$  through  $F_5$  are graphed in Figure 4.1.

Exact values of  $P(n,F)$  were obtained for  $F_1$  and  $F_5$ . Computer simulations gave estimates for  $F_2$ ,  $F_3$  and  $F_4$ .

$P(n,F_1)$  is found using Stevens' formula [7] which states that the probability of covering a circle with  $n$  randomly placed arcs of fixed size  $a$  is

$$(4.1) \quad \sum_{j=0}^{[2\pi/a]} (-1)^j \binom{n}{j} \left(1 - \frac{ja}{2\pi}\right)^{n-1},$$

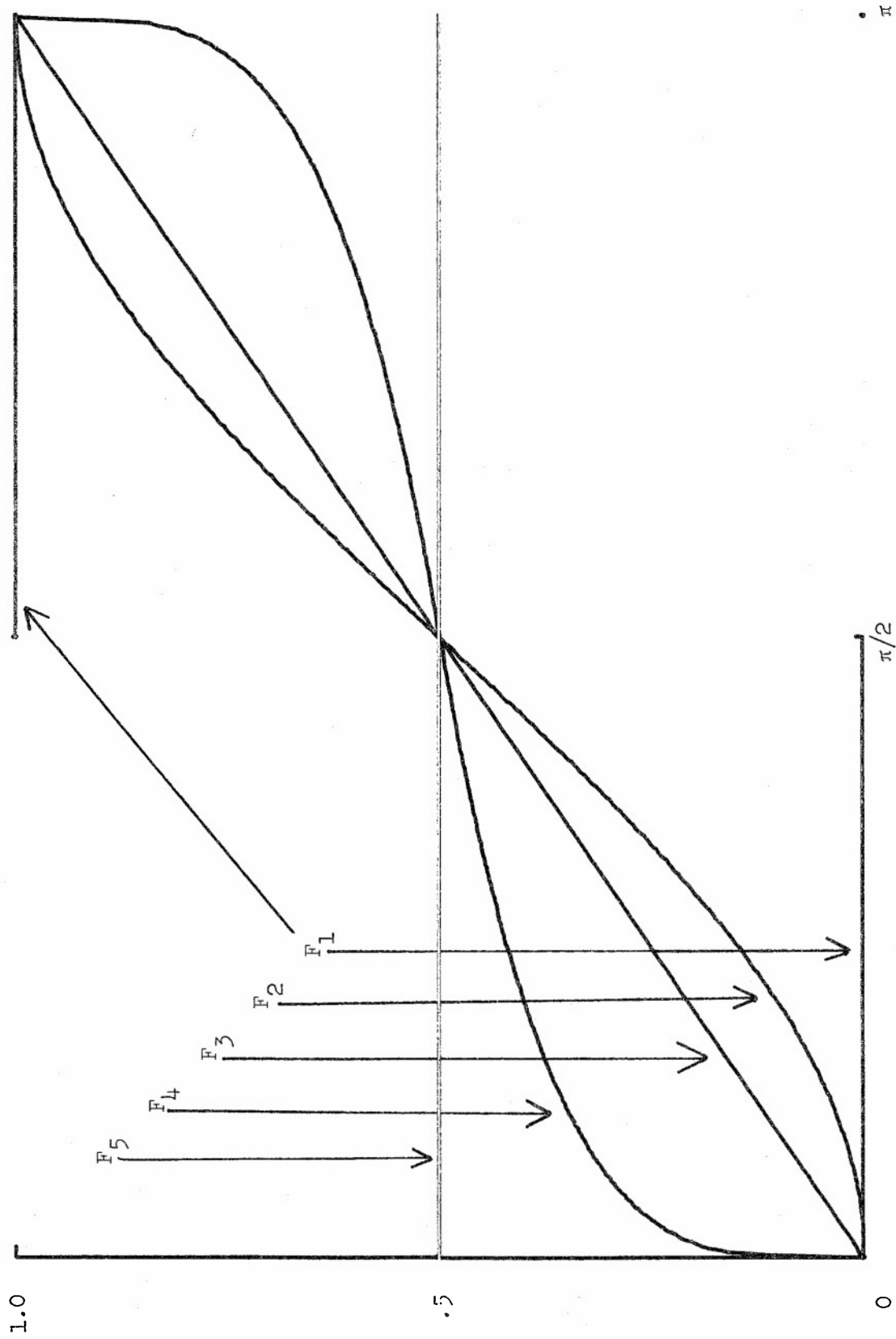


Figure 4.1. Cumulative distribution functions of arc length.

where  $[t]$  denotes the greatest integer contained in  $t$ . Thus

$$(4.2) \quad P(n, F_1) = 1 - n\left(\frac{3}{4}\right)^{n-1} + \binom{n}{2}\left(\frac{1}{2}\right)^{n-1} - \binom{n}{3}\left(\frac{1}{4}\right)^{n-1}.$$

$P(n, F_5)$  is found by conditioning on  $K$ , the random number of arcs of size  $\pi$  actually placed, and using (4.1). Since  $K$  has the binomial distribution  $\mathcal{B}(n, \frac{1}{2})$  we have

$$P(n, F_5) = E P(n, F_5 | K) = \sum_{k=1}^n \binom{n}{k} 2^{-n} (1 - k\left(\frac{1}{2}\right)^{k-1}).$$

To simplify this, note that

$$\sum_{k=1}^n k \binom{n}{k} \left(\frac{1}{2}\right)^{k-1} = 2\left(\frac{3}{2}\right)^n \sum_{k=0}^n k \binom{n}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k} = \frac{2n}{3} \left(\frac{3}{2}\right)^n,$$

using the formula for binomial expectation. Thus

$$(4.3) \quad P(n, F_5) = 1 - 2^{-n} - \frac{2n}{3} \left(\frac{3}{4}\right)^n.$$

Monte Carlo results were done on Stanford's IBM 370/168 computer. For each distribution, 10,000 simulations were performed in which random arcs were placed sequentially until either the circle was covered or 50 arcs were placed. The usual unbiased estimate for probability of success in Bernouilli trials was used. Numerical values of  $P(n, F_i)$  are tabulated in Table 4.1 and graphed in Figure 4.2. Note that the coverage probabilities vary considerably among our distributions, even holding number of arcs and expected arc length constant. For example, with 6 arcs, probability of coverage ranges from .03 to .27; and with 10 arcs it ranges from .34 to .62.

Table 4.1. Coverage probabilities for  $n$  random arcs on the circle.

| $n$ | $P(n, F_1)^*$ | $P(n, F_2)^\dagger$ | $P(n, F_3)^\dagger$ | $P(n, F_4)^\dagger$ | $P(n, F_5)^*$ |
|-----|---------------|---------------------|---------------------|---------------------|---------------|
| 1   | 0.00          | 0.00                | 0.00                | 0.00                | 0.00          |
| 2   | 0.00          | 0.00                | 0.00                | 0.00                | 0.00          |
| 3   | 0.00          | 0.001               | 0.004               | 0.02                | 0.03          |
| 4   | 0.00          | 0.01                | 0.02                | 0.06                | 0.09          |
| 5   | 0.004         | 0.03                | 0.05                | 0.12                | 0.18          |
| 6   | 0.03          | 0.07                | 0.11                | 0.21                | 0.27          |
| 7   | 0.07          | 0.13                | 0.18                | 0.29                | 0.37          |
| 8   | 0.15          | 0.21                | 0.26                | 0.38                | 0.46          |
| 9   | 0.24          | 0.30                | 0.35                | 0.46                | 0.55          |
| 10  | 0.34          | 0.39                | 0.43                | 0.54                | 0.62          |
| 11  | 0.43          | 0.47                | 0.52                | 0.61                | 0.69          |
| 12  | 0.53          | 0.56                | 0.59                | 0.68                | 0.75          |
| 13  | 0.61          | 0.63                | 0.66                | 0.73                | 0.79          |
| 14  | 0.68          | 0.70                | 0.72                | 0.78                | 0.83          |
| 15  | 0.74          | 0.75                | 0.77                | 0.82                | 0.87          |
| 16  | 0.79          | 0.80                | 0.81                | 0.86                | 0.89          |
| 17  | 0.83          | 0.84                | 0.85                | 0.88                | 0.91          |
| 18  | 0.87          | 0.87                | 0.88                | 0.91                | 0.93          |
| 19  | 0.89          | 0.90                | 0.90                | 0.93                | 0.95          |
| 20  | 0.92          | 0.92                | 0.92                | 0.94                | 0.96          |
| 21  | 0.93          | 0.94                | 0.94                | 0.95                | 0.97          |
| 22  | 0.95          | 0.95                | 0.95                | 0.96                | 0.97          |
| 23  | 0.96          | 0.96                | 0.96                | 0.97                | 0.98          |
| 24  | 0.97          | 0.97                | 0.97                | 0.97                | 0.98          |
| 25  | 0.97          | 0.98                | 0.98                | 0.98                | 0.99          |

\* exact

† simulated

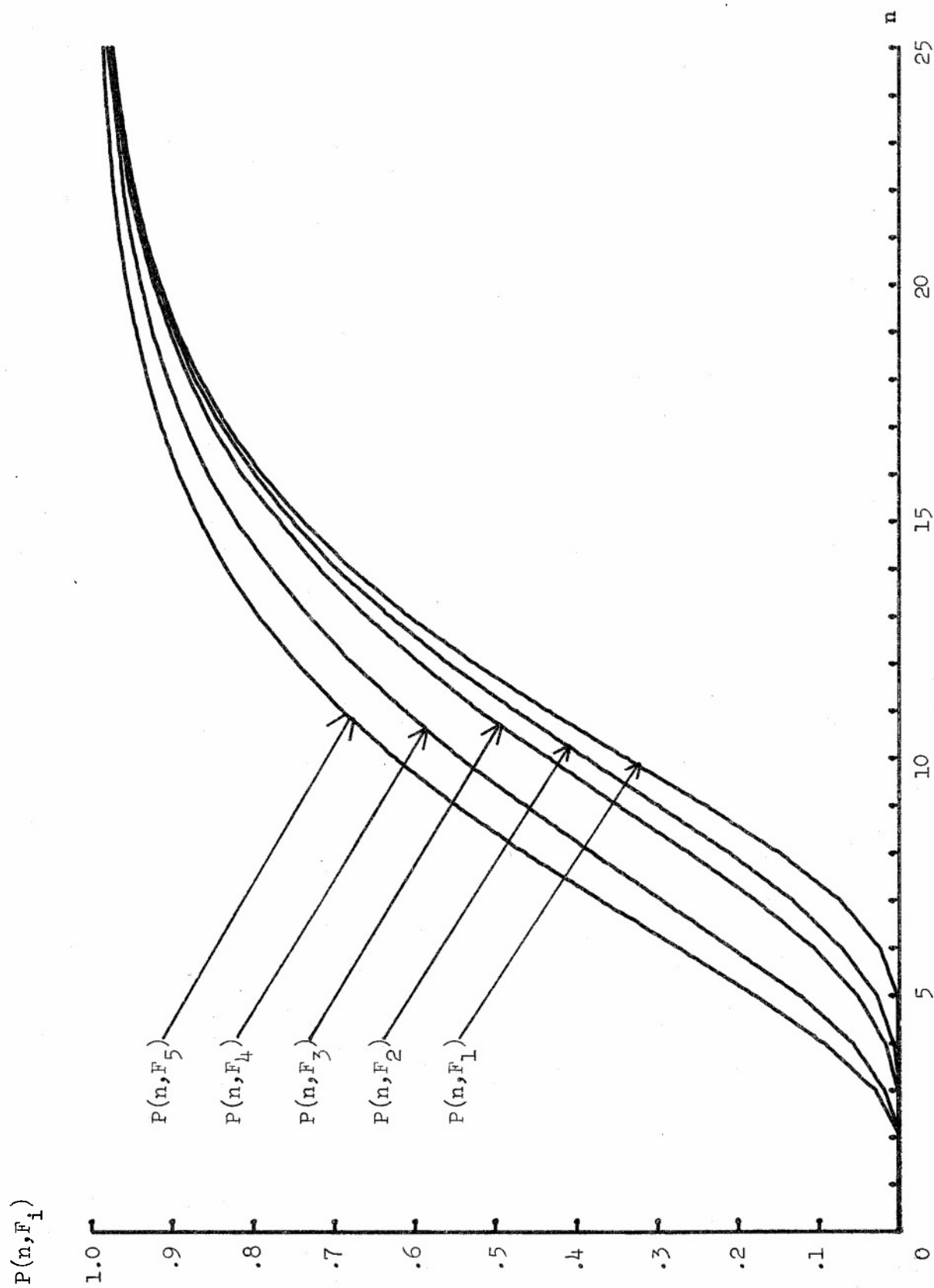


Figure 4.2. Probability that  $n$  random arcs cover the circle.

## 5. COVERAGE PROBABILITIES FOR THE CIRCLE AND PEAKEDNESS OF ARC LENGTH DISTRIBUTION

The values in Table 4.1 and Figure 4.2 suggest that  $P(n, F_1) \leq P(n, F_2) \leq \dots \leq P(n, F_5)$  for each  $n$ . From Figure 4.1 we can see that the  $F_i$  are in order of decreasing peakedness at their common expectation  $\pi/2$ . We say  $F$  is more peaked about  $\alpha$  than  $G$  is about  $\beta$  if  $Y \sim F$  and  $Z \sim G$  imply  $P(|Y - \alpha| \geq t) \leq P(|Z - \beta| \geq t)$  for every  $t \geq 0$ . This definition of peakedness is due to Birnbaum [2].

This suggests a conjecture, namely that among cumulative distribution functions on  $[0, 2\pi]$  having fixed expectation  $e$ , the less peaked  $F$  is about  $e$ , the greater  $P(n, F)$  will be for each  $n$ . That is, one achieves greater probability of coverage if one chooses a distribution of arc length that is spread out rather than more concentrated near its expectation.

So far the conjecture is supported by numerical evidence for five distributions, each with expectation  $\pi/2$ . Next we study a class of distributions that lends further support to the conjecture and enables us to prove that  $P(n, F_1) < P(n, F_5)$ .

Let  $F_{e,a}$  denote the distribution function of arc length  $L$ , where

$$L = \begin{cases} a & \text{probability } \frac{e}{a} \\ 0 & \text{probability } 1 - \frac{e}{a} \end{cases}$$

so that  $e = E L$ . For this class of distributions,  $P(n, F)$  may be calculated exactly as the following lemma shows.

Lemma:

$$(5.1) \quad P(n, F_{e,a}) = 1 - (1 - \frac{e}{a})^n + \sum_{j=1}^{[2\pi/a]} (-1)^j \binom{n}{j} [1 - \frac{ej}{2\pi}]^{n-j} \frac{e}{a} [\frac{e}{a} - \frac{ej}{2\pi}]^{j-1}.$$

Proof: Conditioning on  $K$ , the number of arcs of size  $a$  placed, noting that  $K \sim \mathcal{B}(n, \frac{e}{a})$ , and using Stevens' result, we have

$$\begin{aligned} P(n, F_{e,a}) &= E P(n, F_{e,a} | K) \\ &= \sum_{k=1}^n \binom{n}{k} (\frac{e}{a})^k (1 - \frac{e}{a})^{n-k} \sum_{j=0}^{[2\pi/a]} (-1)^j \binom{k}{j} (1 - \frac{ja}{2\pi})^{k-1}. \end{aligned}$$

Separating off the  $j = 0$  term and rearranging summations, we obtain

$$P(n, F_{e,a}) = 1 - (1 - \frac{e}{a})^n + \sum_{j=1}^{[2\pi/a]} (-1)^j (1 - \frac{ja}{2\pi})^{-1} \sum_{k=1}^n \binom{k}{j} \binom{n}{k} [\frac{e}{a} (1 - \frac{ja}{2\pi})]^k (1 - \frac{e}{a})^{n-k}.$$

Observing that the inner sum is

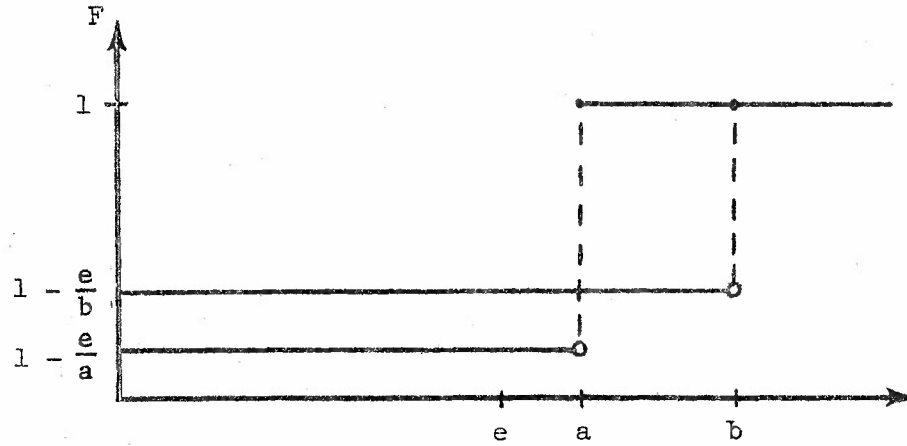
$$[1 - \frac{ej}{2\pi}]^n E \binom{X}{j} \quad \text{where} \quad X \sim \mathcal{B} \left( n, \frac{\frac{e}{a} (1 - \frac{ja}{2\pi})}{1 - \frac{ej}{2\pi}} \right)$$

and using the fact that  $X \sim \mathcal{B}(n, p)$  implies  $E \binom{X}{j} = p^j \binom{n}{j}$  completes the proof.  $\square$

We will be interested in subclasses of distributions with expectation  $e$  that are totally ordered by peakedness at  $e$ . The following proposition exhibits some of these.

Proposition: Let  $e \in (0, 2\pi)$ . Then the class  $\mathcal{F}_e = \{F_{e,a} : e \leq a \leq 2e\}$  is totally ordered by peakedness at  $e$ . That is,  $e \leq a < b \leq 2e$  implies  $F_{e,a}$  is more peaked at  $e$  than is  $F_{e,b}$ .

Proof: Let  $Y \sim F_{e,a}$  and  $Z \sim F_{e,b}$  where  $e \leq a < b \leq 2e$ . A picture will help one visualize the situation.



We must show  $P(|Y-e| \geq t) \leq P(|Z-e| \geq t)$  for all  $t \geq 0$ . Consider three cases:

- i) if  $t \leq b - e$ , then  $P(|Y-e| \geq t) \leq 1 = P(|Z-e| \geq t)$ ,
- ii) if  $b - e < t \leq e$ , then  $P(|Y-e| \geq t) = 1 - \frac{e}{a} \leq 1 - \frac{e}{b} \leq P(|Z-e| \geq t)$ ,
- iii) if  $t > e$ , then  $P(|Y-e| \geq t) = 0 = P(|Z-e| \geq t)$ . □

Theorem 3: For each  $n$ ,  $P(n, F_{\pi/2,a})$  is monotone increasing for  $a \in [\pi/2, \pi]$ . Thus the conjecture holds in  $\mathcal{F}_{\pi/2}$ .

Proof of Theorem 3:  $P(n, F_{\pi/2,a})$  assumes different forms depending upon whether  $a \in [\pi/2, 2\pi/3]$  or  $a \in [2\pi/3, \pi]$ . The proof given here is for  $a \in [2\pi/3, \pi]$  and is analogous (but more complicated) in the other case.

We may assume  $n \geq 3$ , for if  $n < 3$  then  $P(n, F_{\pi/2,a}) = 0$  whenever  $a < \pi$ . Using the lemma,

$$(5.2) \quad P(n, F_{\pi/2,a}) = 1 - \left(1 - \frac{\pi}{2a}\right)^n - \frac{\pi n}{2a} \left(\frac{3}{4}\right)^{n-1} + \frac{\pi n(n-1)}{2^{n+1}a} \left(\frac{\pi}{a} - 1\right)$$

for  $a \in [2\pi/3, \pi]$ . If we set  $t = \pi/2a$ , we need only show that

$$f(t) = 1 - (1-t)^n - nt\left(\frac{3}{4}\right)^{n-1} + \frac{n(n-1)}{2^n} t(2t-1)$$

is monotone decreasing for  $t \in [\frac{1}{2}, \frac{3}{4}]$ . To see this, first observe that

$$f''(t) = n(n-1)\left[\left(\frac{1}{2}\right)^{n-2} - (1-t)^{n-2}\right] \geq 0$$

in this interval. Hence  $f'(t)$  is monotone increasing there.

I will now show that  $\max_{t \in [\frac{1}{2}, \frac{3}{4}]} f'(t) = f'(\frac{3}{4}) \leq 0$ . Observe that

$$f'(\frac{3}{4}) = \frac{n}{2^{n-1}} \left\{ \left(\frac{1}{2}\right)^{n-1} + n - 1 - \left(\frac{3}{2}\right)^{n-1} \right\} = \frac{n}{2^{n-1}} g(n+1)$$

where we define

$$g(x) = \left(\frac{1}{2}\right)^x + x - \left(\frac{3}{2}\right)^x.$$

Observe  $g(2) = 0$ ,  $g(3) = -\frac{1}{4}$ . For  $x > 3$ ,

$$g'(x) = \left(\frac{1}{2}\right)^x \log \frac{1}{2} + 1 - \left(\frac{3}{2}\right)^x \log \frac{3}{2} \leq 1 - \left(\frac{3}{2}\right)^3 \log \frac{3}{2} \approx -.37 < 0.$$

This implies  $f'(t) \leq 0$ , forcing  $f(t)$  to be monotone decreasing for  $t \in [\frac{1}{2}, \frac{3}{4}]$ .

The proof for  $a \in [\pi/2, 2\pi/3]$  uses

$$P(n, F_{\pi/2, a}) = 1 - \left(1 - \frac{\pi}{2a}\right)^n - \frac{\pi n}{2a} \left(\frac{3}{4}\right)^{n-1} + \frac{\pi}{2^n a} \binom{n}{2} \left(\frac{\pi}{a} - 1\right) - \frac{2\pi}{4^n a} \binom{n}{3} \left(\frac{2\pi}{a} - 3\right)^2$$

and starts by observing that the corresponding third derivative is always non-positive. There is no trouble at the boundary  $a = 2\pi/3$  because  $P(n, F_{\pi/2, a})$  is continuous there.  $\square$

$P(n, F_{\pi/2, a})$  is plotted in Figures 5.1 and 5.2, clearly showing its monotone property. In figure 5.2 we see an interesting slope discontinuity at arc size  $\pi$ . This is due to the fact that two arcs of size less than  $\pi$  cannot possibly cover the circle, whereas two arcs of size greater than  $\pi$  can cover the circle with positive probability.

Corollary:  $P(n, F_1) < P(n, F_5)$  whenever  $n \geq 3$ .

Proof: Observe that  $F_1 = F_{\pi/2, \pi/2}$  and  $F_5 = F_{\pi/2, \pi}$ . Hence  $P(n, F_1) \leq P(n, F_5)$  follows immediately from Theorem 3. Strict inequality when  $n \geq 3$  follows from  $f''(t) > 0$  for  $t \in (\frac{1}{2}, \frac{3}{4}]$  in the proof of Theorem 3.  $\square$

Theorem 4: Fix  $e \in (0, \pi)$ . Then  $P(n, F_{e, a})$  is monotone increasing for  $a \in [\max(e, \pi), 2\pi]$  for each  $n$ .

Proof of Theorem 4:  $a \geq \pi$ , so by the lemma,

$$(5.3) \quad P(n, F_{e, a}) = 1 - (1 - \frac{e}{a})^n - \frac{ne}{a} (1 - \frac{e}{2\pi})^{n-1}.$$

Setting  $t = \frac{e}{a}$ , we need only show

$$f(t) = 1 - (1-t)^n - nt(1 - \frac{e}{2\pi})^{n-1}$$

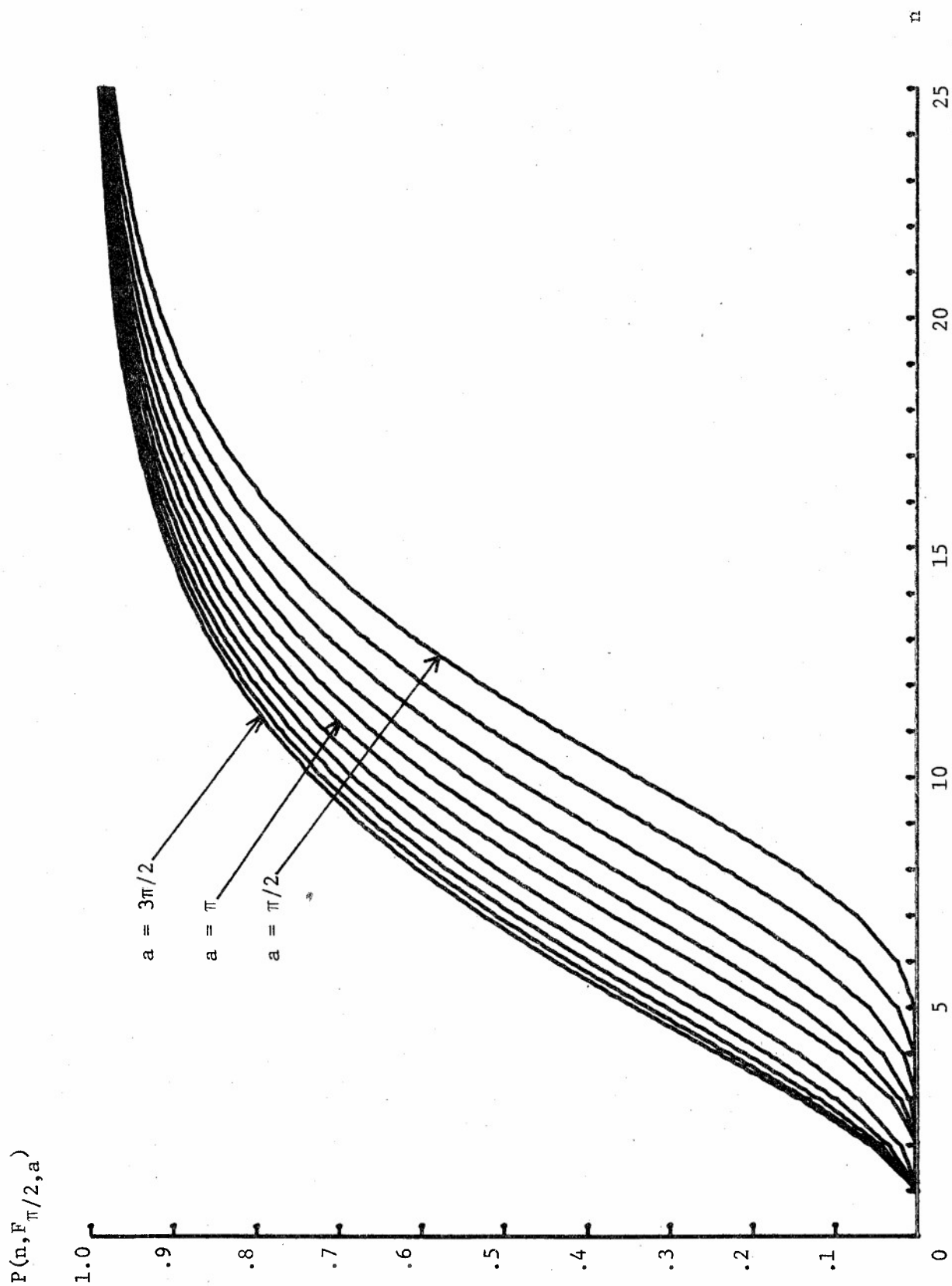


Figure 5.1. Circle coverage probabilities  $P(n, F_{\pi/2, a})$  plotted against  $n$  for arc sizes  $a = \pi/2$  to  $3\pi/2$  in steps of  $\pi/10$ .

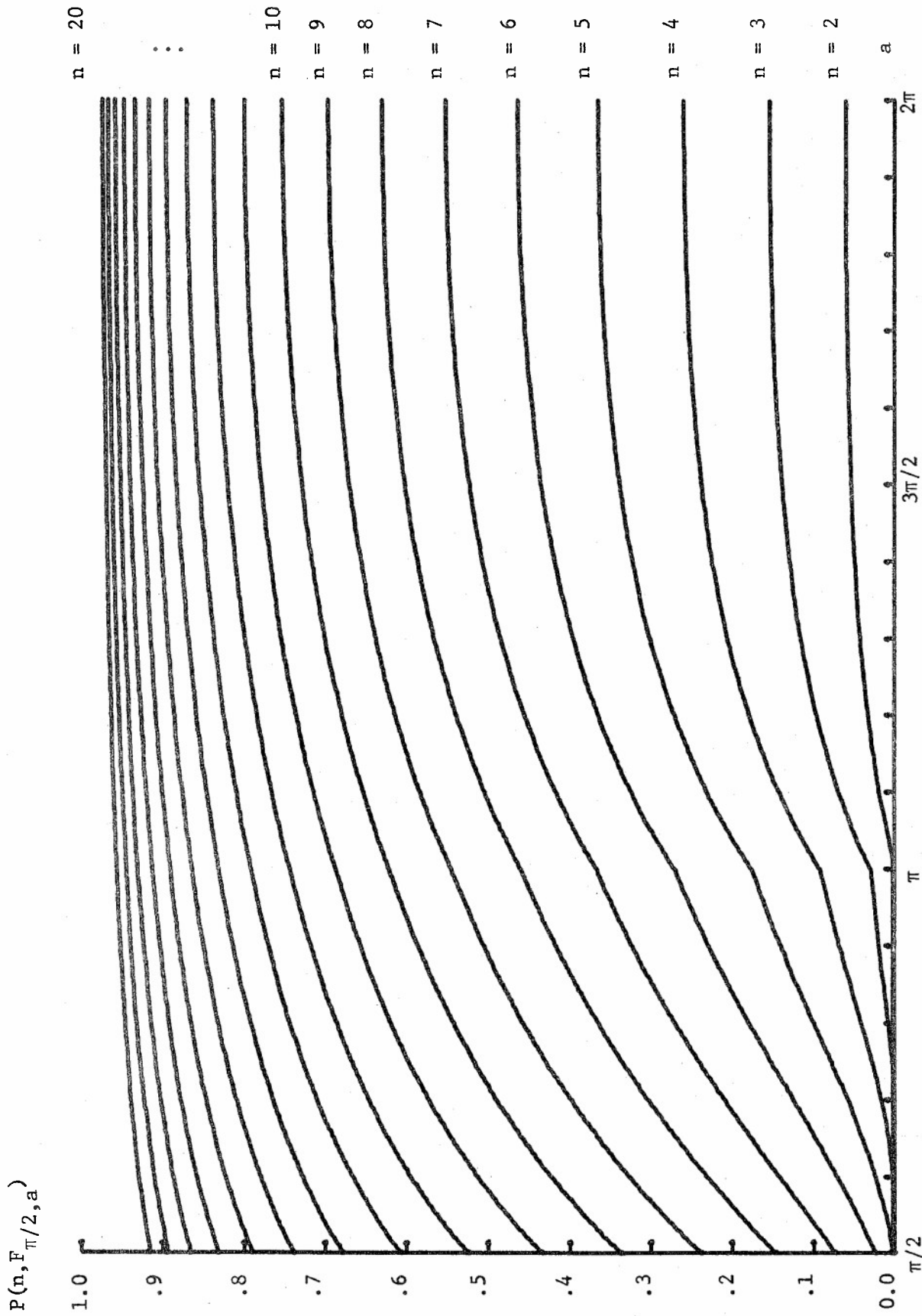


Figure 5.2. Circle coverage probabilities  $P(n, F_{\pi/2, a})$  plotted against arc size  $a$  for some values of  $n$ .

is monotone decreasing for  $t \in [\frac{e}{2\pi}, \min(\frac{e}{\pi}, 1)]$ . But

$$f'(t) = n \left[ (1-t)^{n-1} - \left(1 - \frac{e}{2\pi}\right)^{n-1} \right] \leq 0$$

in that interval, completing the proof. □

Corollary: Let  $\frac{\pi}{2} < e < \pi$ . Then the conjecture holds in the class of distributions

$$\{F_{e,a} : \pi \leq a \leq 2e\}$$

Proof: This is an immediate consequence of theorem 4 and the proposition of this section. □

### References

- [1] Ailam, G. (1966), "Moments of Coverage and Coverage Spaces," J. Appl. Prob. 3, 550-555.
- [2] Birnbaum, Z. W. (1948), "On Random Variables with Comparable Peakedness," Ann. Math. Stat. 19, 76-81.
- [3] Gilbert, E. N. (1965), "The Probability of Covering a Sphere with n Circular Caps," Biometrika 52, 323-330.
- [4] Moran, P. A. P. and Fazekas de St. Groth, S. (1962), "Random Circles on a Sphere," Biometrika 49, 389-396.
- [5] Robbins, H. E. (1944), "On the Measure of a Random Set," Ann. Math. Stat. 15, 70-74.
- [6] Royden, H. L. (1968), "Real Analysis," The Macmillan Co.
- [7] Stevens, W. L. (1939), "Solution to a Geometrical Problem in Probability," Ann. Eugenics, Lond. 9, 315-320.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE                                                                                                                                                                                                                                                                                                       |                       | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM                                 |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|-----------------------------------------------------------------------------|
| 1. REPORT NUMBER<br>4                                                                                                                                                                                                                                                                                                           | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER                                               |
| 4. TITLE (and Subtitle)<br>Random Space Filling and Moments of Coverage<br>in Geometrical Probability                                                                                                                                                                                                                           |                       | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report                      |
|                                                                                                                                                                                                                                                                                                                                 |                       | 6. PERFORMING ORG. REPORT NUMBER                                            |
| 7. AUTHOR(s)<br>A. Siegel                                                                                                                                                                                                                                                                                                       |                       | 8. CONTRACT OR GRANT NUMBER(s)<br>DAAG29-77-G-0031                          |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Department of Statistics<br>Stanford University<br>Stanford, California 94305                                                                                                                                                                                                    |                       | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>P-14435-M |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>U.S. Army Research Office<br>Post Office Box 12211<br>Research Triangle Park, NC 27709                                                                                                                                                                                               |                       | 12. REPORT DATE<br>April 4, 1977                                            |
|                                                                                                                                                                                                                                                                                                                                 |                       | 13. NUMBER OF PAGES<br>31                                                   |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)                                                                                                                                                                                                                                                     |                       | 15. SECURITY CLASS. (of this report)<br>Unclassified                        |
|                                                                                                                                                                                                                                                                                                                                 |                       | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE                               |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release; distribution unlimited.                                                                                                                                                                                                                             |                       |                                                                             |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)                                                                                                                                                                                                                                      |                       |                                                                             |
| 18. SUPPLEMENTARY NOTES<br>The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents. This report partially supported under Office of Naval Research Contract N00014-76-C-0475 (NR-042-267) and issued as Technical Report No. 244 |                       |                                                                             |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>Geometrical probability, coverage spaces, random arcs on a circle, moments of coverage, coverage probability, simulation techniques, peakedness                                                                                           |                       |                                                                             |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>See reverse side                                                                                                                                                                                                                           |                       |                                                                             |

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE  
S/N C102-014-6601

Unclassified.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT: Random Space Filling and Moments of Coverage in  
Geometrical Probability by A. Siegel

The moments of the random proportion of a fixed set that is covered by a random set (moments of coverage) are shown to converge under very general conditions to the probability that the fixed set is almost everywhere covered by the random set. Moments and coverage probabilities are calculated for several cases of random arcs of random sizes on a circle. One tends to increase coverage by choosing an arc length distribution that is less peaked at its expectation when that expectation is constrained at a particular value.