Covering a circle with random arcs

TECHNISCHE HOGESCHOOL EINDHOVEN

NEDERLAND

ONDERAFDELING DER WISKUNDE

TECHNOLOGICAL UNIVERSITY EINDHOVEN

THE NETHERLANDS

DEPARTMENT OF MATHEMATICS

Covering a circle with random arcs

by

R. Doornbos and J.H. van Lint

T.H. Report 70-WSK-01

January 1970

Covering a circle with random arcs

by

R. Doornbos and J.H. van Lint

1. Introduction. The problem discussed in this note and some of the ideas used in proofs were suggested by a problem posed by F. Göbel in Nieuw Archief voor Wiskunde (Vol. 18, 1970). Consider random points on the unit circle drawn independently from a uniform distribution. Let the circumference of the circle be 1 and suppose each point is "first endpoint" (in a fixed orientation) of an arc of length α (α fixed). What is the expected value of the number of arcs necessary to cover the circle?
If we call the number of arcs $\underline{n}$ then

$$(1.1)\qquad E(\underline{n}) = 1 + \sum_{k=1}^{[\alpha^{-1}]} (-1)^{k-1}(1-k\alpha)^{k-1}(k\alpha)^{-k-1} .$$

Several results in the literature are sufficient to demonstrate (1.1) e.g. S.S. Wilks, Mathematical Statistics, 8.43 or F.W. Steutel, Random division of an interval, Stat. Neerl. 21 (1967), 231 - 244 (cf. last formula on p. 236). In the latter it is proved that

$$(1.2)\qquad E(\underline{n}) = \frac{1}{\alpha}\left\{\log\frac{1}{\alpha} + \log\log\frac{1}{\alpha} + \gamma + o(1)\right\} \qquad (\alpha \to 0) .$$

In § 2 we give a simple proof of (1.1). In § 3 we give a second proof by considering a more difficult problem namely the expected value of the number of arcs necessary to cover a certain part of the circle. Finally in § 4 we consider the problem of covering the circle with arcs of random length.

2. Proof of (1.1). Suppose $(\ell+1)^{-1} \leq \alpha < \ell^{-1}$, ℓ an integer. Instead of speaking of the circle we consider the real numbers mod 1 and random variables $\underline{x}_1, \underline{x}_2, \ldots$ chosen from a uniform distribution on $[0,1]$. We consider the event $\underline{n} > n$. For $i = 1,2,\ldots,n$ we define the event E_i as $\underline{x}_k \notin [\underline{x}_i, \underline{x}_i + \alpha]$ for $k \neq i$. Then by the principle of inclusion and exclusion

$$P(\underline{n} > n) = \sum_{i=1}^{n} P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \ldots$$

where $P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k}) = 0$ if $k > \ell$.

For $k \leqslant \ell$ we have

$$P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_k}) = (1 - k\alpha)^{n-1} .$$

Hence

$$P(\underline{n} > n) = \sum_{k=1}^{\ell} (-1)^{k-1}\binom{n}{k}(1 - k\alpha)^{n-1} \quad \text{for } n = 1,2,\ldots .$$

Trivially

$P(\underline{n} > 0) = 1$. We now find

$$E(\underline{n}) = \sum_{n=1}^{\infty} n \{P(\underline{n} > n-1) - P(\underline{n} > n)\} =$$

$$= 1 + \sum_{n=1}^{\infty} \sum_{k=1}^{\ell} (-1)^{k-1} \binom{n}{k}(1 - k\alpha)^{n-1}$$

$$= 1 + \sum_{k=1}^{\ell} (-1)^{k-1} (1 - k\alpha)^{k-1} (k\alpha)^{-k-1} .$$

For ℓ we can write $[\alpha^{-1}]$ since if $\alpha = (\ell+1)^{-1}$ the term with $k = \ell+1$ is 0 .

3. Covering part of the circle. We now consider the problem of covering $[0,x]$ where $0 < x \leqslant 1$ and define if $\underline{n}$ is the number of intervals necessary for the covering :

(3.1) $F(x) := E(\underline{n})$.

We also consider the case of two intervals. Suppose $\alpha \leqslant 1/2$ and let x and y both be less than α. Furthermore let p be a number with

$$x + \alpha \leqslant p \leqslant 1 - y - \alpha.$$

Let $\underline{m}$ be the number of random intervals necessary to cover $[0,x]$ and $[p,p+y]$.

We define

$$(3.2) \qquad G(x,y) := E(\underline{m}) .$$

The function G does not depend on p .

We now consider the functional equations for F and G .

For $x \leq \alpha$ we have

$$(3.3) \qquad F(x) = \int_0^x \{1+ F(t)\}\, dt + (1-\alpha-x) \{1+ F(x)\} + $$

$$+ \int_{1-\alpha}^{1-\alpha+x} \{1+ F(1-\alpha+x-t)\}\, dt + (\alpha-x) ,$$

i.e.

$$(3.4) \qquad (\alpha+x)\ F(x) = 1 + 2 \int_0^x F(t)\, dt .$$

Hence

$$(3.5) \qquad F(x) = \alpha^{-2} (x+\alpha) \qquad \text{for } 0 \leq x \leq \alpha .$$

If $\alpha \geq 1/2$ we can substitute $x = 1-\alpha$. This gives us

$$(3.6) \qquad F(1) = 1 + F(1-\alpha) = 1 + \alpha^{-2} ,$$

i.e. (1.1) for $1/2 \leq \alpha \leq 1$.

Now assume $\alpha \leq x \leq 2\alpha$. Then we have

$$(3.7) \qquad F(x) = \int_0^{x-\alpha} \{1+ G(t,x-\alpha-t)\}\, dt + 2 \int_0^{\alpha} \{1 + F(x-t)\}\, dt +$$

$$+ (1-\alpha-x) \{1 + F(x)\} ,$$

i.e. by (3.5)

$$(3.8)\qquad (\alpha+x)\ F(x) = 5 - \alpha^{-2}x^2 + 2\int_{\alpha}^{x} F(t)\,dt + \int_{0}^{x-\alpha} G(t,x-\alpha-t)\,dt\ .$$

We use the same method to compute $G(x,y)$ for small values of x and y .

As in (3.3) we find :

$$(3.9)\qquad G(x,y) = (\alpha-x)\ \{1+F(y)\} + (\alpha-y)\ \{1+F(x)\} + 2\int_{0}^{y}\{1+G(x,t)\}\,dt +$$

$$+ 2\int_{0}^{x}\{1+G(t,y)\}\,dt + (1-2\alpha-x-y)\ \{1+G(x,y)\}\ ,$$

i.e.

$$(3.10)\qquad (2\alpha+x+y)\ G(x,y) = 3 - 2\alpha^{-2}xy + 2\int_{0}^{y} G(x,t)\,dt + 2\int_{0}^{x} G(t,y)\,dt.$$

The functional equation (3.10) can be solved by standard methods (differentiation etc.) yielding

$$(3.11)\qquad G(x,y) = -\tfrac{1}{4}\alpha^{-3}xy + \tfrac{3}{4}\alpha^{-2}(x+y) + \tfrac{3}{2}\alpha^{-1}\ \text{for}\ 0 \leqslant x \leqslant \alpha\ ,$$

$$0 \leqslant y \leqslant \alpha\ .$$

Substitution of (3.11) in (3.8) gives us

$$(3.12)\qquad (\alpha+x)\ F(x) = \frac{103}{24} - \frac{1}{8}\alpha^{-1}x - \frac{1}{8}\alpha^{-2}x^2 - \frac{1}{24}\alpha^{-3}x^3 + 2\int_{\alpha}^{x} F(t)\,dt,$$

fron which we find F in the interval $[\alpha, 2\alpha]$:

$$(3.13)\qquad F(x) = \tfrac{9}{8}\alpha^{-1} + \alpha^{-2}x - \tfrac{1}{8}\alpha^{-3}x^2 \qquad (\alpha \leqslant x \leqslant 2\alpha)\ .$$

The treatment of $2\alpha \leqslant x \leqslant 3\alpha$ etc. is essentially the same .

Suppose $\frac{1}{3} \leqslant \alpha \leqslant 1/2$. By (3.13) we have

$$(3.14) \qquad F(1) = 1 + F(1-\alpha) = $$
$$= 1 + \frac{5}{4}\alpha^{-2} - \frac{1}{8}\alpha^{-3} ,$$

i.e. (1.1) for $\frac{1}{3} \leq \alpha \leq 1/2$.

Concerning (3.11) we make the following remark. Suppose random arcs of length $\alpha < 1/2$ are chosen on the circle of circumference 1 and two opposite points on the same diameter are to be covered. Clearly the expected value of the number of intervals necessary for the covering is

$$2 \sum_{k=0}^{\infty} (1-\alpha)^k - \sum_{k=0}^{\infty} (1-2\alpha)^k = \frac{3}{2}\alpha^{-1} = G(0,0) .$$

If we replace the two points by small intervals with length x and y then $G(x,y)$ is the expected value. The other extreme treated by G is the case where two opposite arcs of length $\frac{1}{4}$ are to be covered by random arcs which als have length $\frac{1}{4}$. Then we have

$$x = y = \alpha = \frac{1}{4} , \; p = \frac{1}{2} \text{ and the expected value is } 11 .$$

4. Covering the circle by arcs of random length

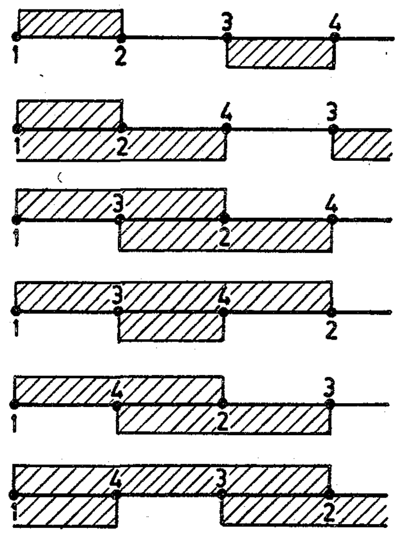
We consider again random points on the circle, but now arcs are formed between pairs of points :

$$[\underline{x}_{2i-1}, \underline{x}_{2i}] \quad (i = 1,2,..)$$

We are interested in the expected number of arcs necessary to cover the circle.
In order to make the principle clear we consider first the case of 2 arcs [1]. Instead of the circle we take again the interval $[0,1]$ and we identify $\underline{x}_1$ with 0 (and consequently also with 1) .

[1] We will still use the name arc when the arcs are straightened in the interval $[0,1]$.

Then all permutations of the remaining 3 points have the same probability. Thus there are 6 possible cases, shown in the figure, where x_i is denoted by i .



In 1 out of the 6 cases the "circle" is covered, viz. in the last case. In 4 cases one endpoint is still "open", that means that it is not lying inside an other interval, in 1 case two endpoints are open .
In the general case we denote by $P_{k,n}$ the probability that k specified endpoints are open when n arcs have been drawn. Thus for instance

$$P_{1,2} = \frac{3}{6} \text{ and } P_{2,2} = \frac{1}{6} . \qquad \text{Then we have :}$$

$$(4.1) \qquad P(\underline{n} > n) = \binom{n}{1} P_{1,n} - \binom{n}{2} P_{2,n} + \ldots + (-1)^{n-1} \binom{n}{n} P_{n,n} .$$

For n = 2 this becomes $P(\underline{n} > 2) = 2 . \frac{3}{6} - 1 . \frac{1}{6} = \frac{5}{6}$.

Next we determine $P_{k,n}$. In the first place the k endpoints must be open when the other (n-k) arcs are left out. This can be arrived at in (k-1)! different ways.
We may namely consider the arcs ending in $\underline{x}_2 , \underline{x}_4 , \ldots, \underline{x}_{2k}$.
Their configuration is determined by the sequence of the (k-1) points $\underline{x}_4 , \ldots , \underline{x}_{2k}$.

The probability that the endpoints $\underline{x}_2, \ldots \underline{x}_{2k}$ remain open does not depend on their order and we assume therefore that

$$0 < \underline{x}_2 < \underline{x}_4 < \ldots < \underline{x}_{2k} < 1 .$$

Each of the remaining (n-k) arcs must lie inside one of the intervals

$$[\underline{x}_2, \underline{x}_4], [\underline{x}_4, \underline{x}_6], \ldots [\underline{x}_{2k}, \underline{x}_2].$$

We denote the number of arcs within the i-th of these intervals by n_i.

Thus $\quad n_i \geq 0 , \sum_1^k n_i = n-k .$

The number of ways in which n_i arcs can be fitted in such an interval is

$$(2n_i + 1)!/2^{n_i} .$$

This can be seen in the following way. The number of permutations of the begin and endpoints of the n_i arcs and the beginpoint $\underline{x}_{2i+1}$ which lies already in the interval $[\underline{x}_{2i}, \underline{x}_{2i+2}]$ is $(2n_i + 1)!$.
But only those permutations are permitted that leave the order of begin- and endpoint of each arc unaltered. This gives the factor 2^{-n_i}.
The total number of favourable possibilities is therefore :

$$(4.2) \qquad \frac{(k-1)!(n-k)!}{2^{n-k}} \sum_{\Sigma n_i = n-k} \frac{(2n_1 + 1)! \ldots (2n_k + 1)!}{n_1! \ldots n_k!}$$

In order to get $P_{k,n}$ this has to be divided by the total number of possibilities which is $(2n - 1)!$ ($\underline{x}_1$ being fixed at 0).
Substitution of this result into (4.1) gives

$$(4.3) \qquad P(\underline{n} \geq n) = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}(k-1)!(n-k)!}{2^{n-k}\,(2n-1)!} \sum_{\Sigma n_i = n-k} \frac{(2n_1+1)!\ldots(2n_k+1)!}{n_1! \ldots n_k!} =$$

$$= \frac{n!}{2^n(2n-1)!} \sum_{k=1}^{n} \frac{2^k(-1)^{k-1}}{k} \sum_{\Sigma n_i = n-k} \prod_{i=1}^{k} \frac{(2n_i+1)!}{n_i!} \,.$$

Now

$$(4.4) \qquad E(\underline{n}) = \sum_{n=0}^{\infty} P(n > n) =$$

$$= 1 + \sum_{n=1}^{\infty} \frac{n!}{2^n(2n-1)!} \sum_{k=1}^{n} \frac{2^k(-1)^{k-1}}{k} \sum_{\Sigma n_i = n-k} \prod_{i=1}^{k} \frac{(2n_i+1)!}{n_i\,!} \,.$$

A numerical calculation up to $n = 15$ showed that

$$(4.5) \qquad E(\underline{n}) = 4.535\ldots$$

Finally (4.4) can also be written as

$$(4.6) \qquad E(\underline{n}) = 1 + \sum_{n=1}^{\infty} \sum_{k=1}^{n} (-1)^{k-1} t_{k,n} \quad , \text{ say } .$$

Rearranging terms we can write alternatively :

$$(4.7) \qquad E(\underline{n}) = 1 + \sum_{k=1}^{\infty} (-1)^{k-1} \sum_{n=k}^{\infty} t_{k,n} \,.$$

Now $\displaystyle\sum_{n=1}^{\infty} t_{1,n} = \sum_{n=1}^{\infty} \frac{n}{2^{n-1}} = 4$.

This gives as a first approximation for $E(\underline{n})$ the value 5, which is the value found for $E(\underline{n})$ in (1.1) for $\alpha = -1/2$, the average arc length in our case .