J. Appl. Prob. **17**, 134–144 (1980)
Printed in Israel
0021–9002/80/010134–11 $01.35

ON THE MINIMUM OF GAPS GENERATED BY ONE-DIMENSIONAL RANDOM PACKING

YOSHIAKI ITOH,* *The Institute of Statistical Mathematics*

Abstract

Let $L(t)$ be the random variable which represents the minimum of length of gaps generated by random packing of unit intervals into $[0,t]$. We have

$$P(L(x+1)\geqq h)=\frac{1}{x}\int_0^x P(L(y)\geqq h)P(L(x-y)\geqq h)dy$$

with

$$P(L(x)\geqq h)=\begin{cases}0 & \text{for} \quad 0\leqq x<h\\ 1 & \text{for} \quad h\leqq x<1\\ 0 & \text{for} \quad x=1.\end{cases}$$

Using this equation the asymptotic behaviour of $P(L(x)\geqq h)$ is discussed.

ONE-DIMENSIONAL RANDOM PACKING; MINIMUM OF GAPS; DELAY NON-LINEAR INTEGRAL EQUATION; ASYMPTOTIC BEHAVIOUR; NUMERICAL CALCULATION

1. Introduction

Random packing in an interval $[0,x]$ by unit intervals has been studied by several authors. Rényi (1958) proved that the expectation of the random packing density approaches a constant

$$\int_0^\infty \exp\left\{-2\int_0^t \frac{1-e^{-u}}{u}du\right\}dt \doteqdot 0.748$$

as x goes to infinity. Ney (1962) studied higher moments for the problem. Dvoretzky and Robbins (1964) proved the corresponding central limit theorem. Bankövi (1962) obtained the limiting distribution of a gap chosen at random. In this paper we introduce a non-linear delay integral equation to study the asymptotic behaviour of the minimum of lengths of gaps.

As in Rényi (1958), let us place at random a unit interval I_1 in the interval $[0,x]$. We assume that the initial point ξ of the interval I_1 is a random variable

Received 28 December 1978; revision received 26 February 1979.

*Postal address: The Institute of Statistical Mathematics, 4–6–7 Minami-Azabu Minato-Ku, Tokyo, Japan.

uniformly distributed in the interval $[0, x-1]$. If the intervals $I_1, I_2, \cdots, I_k$ have already been chosen, the next randomly chosen interval will be kept only if it does not intersect any of the intervals $I_1, I_2, \cdots, I_k$. In this case this interval will be denoted by I_{k+1}. If it does intersect any of the intervals $I_1, I_2, \cdots, I_k$, we neglect it and choose a new interval. The procedure is continued until none of the lengths of gaps generated by the intervals placed in $[0, x]$ is greater than 1. We denote by $N(x)$ the random number of intervals packed in $[0, x]$ by this procedure and write $M(x) \equiv E(N(x))$.

If a unit interval is already placed in the interval $[0, x+1]$ with initial point at y $(0 \leqq y \leqq x)$, then the average number of the intervals at the left is $M(y)$ and that of the intervals at the right is $M(x-y)$. We have $M(x+1 \mid y) = M(y) + M(x-y) + 1$, denoting by $|y$ conditioning on the event that the first interval is $[y, y+1]$. Since y is uniformly distributed in $[0, x]$, we have

$$M(x+1) = \frac{1}{x}\int_0^x (M(y) + M(x-y) + 1)dy. \tag{1}$$

Analyses given by Rényi (1958), Dvoretzky and Robbins (1964), and Bankövi (1962) are based on this equation.

Let $L(x)$ be the random variable which represents the minimum of lengths of gaps generated by the above random packing procedure. If a unit interval is already placed on the interval $[0, x+1]$ with initial point at y, then

$$P(L(x+1) \geqq h \mid y) = P(L(y) \geqq h)P(L(x-y) \geqq h).$$

Since y is uniformly distributed in $[0, x]$, we have

$$P(L(x+1) \geqq h) = \frac{1}{x}\int_0^x P(L(y) \geqq h)P(L(x-y) \geqq h)dy \tag{2}$$

with

$$P(L(x) \geqq h) = \begin{cases} 0 & \text{for} \quad 0 \leqq x < h \\ 1 & \text{for} \quad h \leqq x < 1 \\ 0 & \text{for} \quad x = 1. \end{cases}$$

A random packing model for elections is introduced in Itoh and Ueda (1978), (1979), and Itoh (1978), to explain percentage gains among candidates of the Liberal Democratic Party, using numerical values obtained by computer experiments. In Itoh (1978), Equation (2) is introduced as a remark. In this paper we study the asymptotic behaviour of $P(L(x) \geqq h)$ which will give a justification for the values obtained by computer experiments.

2. Results

Put

$$f(x) \equiv P(L(x) \geqq h)$$

then $f(x)$ satisfies

$$f(x+1)=\frac{1}{x}\int_0^x f(x-y)f(y)dy \tag{3}$$

with

$$f(x)=\begin{cases}0 & \text{for } 0\leqq x<h\\ 1 & \text{for } h\leqq x<1\\ 0 & \text{for } x=1.\end{cases}$$

We define the functions

$$f_k(x) \qquad k=1,2,3,\cdots,$$

as follows;

$$f_1(x)=\begin{cases}1 & \text{for } h\leqq x<1\\ 0 & \text{for } x<h \text{ or } 1\leqq x,\end{cases}$$

and for $k\geqq 2$,

$$f_k(x)=\begin{cases}\dfrac{1}{x-1}\sum_{i+j=k} f_i * f_j(x-1) & \text{for } x\neq 1\\ 0 & \text{for } x=1,\end{cases}$$

where

$$f_i * f_j(x)=\int_0^x f_i(x-y)f_j(y)dy.$$

We define the probability densities $P_n^{k}{}^*(x)$ as the k-fold convolution of $P_h(x)$, where

$$P_h(x)=\begin{cases}\dfrac{1}{1-h} & \text{for } h\leqq x\leqq 1\\ 0 & \text{for } x<h \text{ or } 1<x.\end{cases}$$

Lemma 1.

$$P(L(x)\geqq h)=\sum_{i=1}^{\infty} f_i(x).$$

Proof. The existence and uniqueness of the solution of Equation (3) is trivial, since the functional equation gives a unique way of constructing the solution from the $f(x)$ for $x\in[0,1]$.

$$f(x+1)=\frac{1}{x}\Big(\sum_{i=1}^{\infty} f_i\Big)*\Big(\sum_{j=1}^{\infty} f_j\Big)(x)=f_2(x+1)+f_3(x+1)+\cdots.$$

Since $f_1(x+1)=0$ for $x>0$, $f(x)=\Sigma_{i=1}^{\infty} f_i(x)$ satisfies Equation (3) formally.

$$f_k(x)>0 \quad \text{for} \quad k(1+h)-1<x<2k-1, \text{and otherwise } f_k(x)=0.$$

So if $k(1+h)-1<x<(k+1)(1+h)-1$, $f(x)=\Sigma_{i=1}^{k} f_i(x)$. Hence

$$f(x)=\sum_{i=1}^{[(x+1)/(1+h)]} f_i(x)$$

is the solution of Equation (3).

Lemma 2.

$$\frac{(1-h)^k}{2^{k-1}}P_h^{k*}(x-k+1)\leqq f_k(x).$$

Proof. The statement for $k=1$ follows from the definition. Assume that the statement holds for every positive integer k for which $k<k_0$. Then

$$\begin{aligned} f_{k_0+1}(x) &= \frac{1}{x-1}\int_0^{x-1}\sum_{i+j=k_0+1} f_i(x-1-y)f_j(y)dy \\ &= \frac{1}{x-1}k_0\frac{(1-h)^{k_0+1}}{2^{k_0-1}}P_h^{(k_0+1)*}(x-k_0). \end{aligned}$$

Since $P_h^{(k_0+1)*}(x-k_0)=0$ for $2k_0+1\leqq x$, we have

$$f_{k_0+1}(x)\geqq\frac{1}{2k_0}k_0\frac{(1-h)^{k_0+1}}{2^{k_0-1}}P_h^{(k_0+1)*}(x-k_0)\geqq\frac{(1-h)^{k_0+1}}{2^{k_0}}P_h^{(k_0+1)*}(x-k_0).$$

Lemma 3. For an arbitrary h $(0<h<1)$, there is a constant $c>0$ which satisfies

$$1<\lim_{x\to\infty}\frac{e^{-cx}}{P(L(x)\geqq h)}.$$

Proof. The statement follows from an application of a theorem for terminating a renewal process in Feller (1971), p. 376 to Lemma 1 with Lemma 2.

Theorem. For an arbitrary h $(0<h<1)$, there exists $a(h)>0$ for which

$$\lim_{x\to\infty}\frac{1}{x}\int_0^x e^{a(h)(x+1)}P(L(x)\geqq h)dx=1.$$

Proof. We consider

$$xf(x+1)=\int_0^x f(x-y)f(y)dy$$

with

$$f(x)=\begin{cases} 0 & \text{for} \quad 0 \leqq x < h \\ 1 & \text{for} \quad h \leqq x < 1 \\ 0 & \text{for} \quad x = 1. \end{cases}$$

Putting

$$g(s)=\int_0^\infty e^{-sx} f(x+1)dx,$$

we have

$$\left(\frac{e^{-sh}-e^{-s}}{s}+e^{-s}g(s)\right)^2=-\frac{d}{ds}g(s).$$

Put $\alpha(s)=g(-s)$. Then

$$\left(\frac{e^{sh}-e^{s}}{s}+e^{s}\alpha(s)\right)^2=\frac{d}{ds}\alpha(s).$$

From Lemma 3, $\alpha(s)$ blows up at a certain point $a(h)>0$. Hence

$$\frac{\alpha'(s)}{(\alpha(s))^2} \xrightarrow[s\to -0+a(h)]{} e^{2a(h)}.$$

So

$$\frac{-1}{\alpha(s)} \approx (s-a(h))e^{2a(h)} \quad \text{for} \quad s \to -0+a(h).$$

(See p. 139 of de Bruijn (1958)). Hence

$$(s+a(h))g(s) \xrightarrow[s\to +0-a(h)]{} e^{-2a(h)}.$$

Thus we have

$$\int_0^x f(y+1)e^{a(h)(y+2)}dy \approx x.$$

(See the Tauberian theorem on p. 443 of Feller (1971).) Thus the result follows.

3. Numerical calculations for $a(h)$

We give an algorithm to estimate the lower bound and the upper bound of $a(h)$. Since $0 \leqq g(s) \leqq \int_0^\infty e^{-sx}dx = 1/s$, $g(\infty)=0$. Putting $t=e^{-2s}$, and defining $\phi(t)$ by $\phi(t)\equiv g(s)$, we have

$$2\phi'(t)=[\phi(t)+C(t)]^2$$

with $\phi(0)=0$, where

$$C(t)=\frac{2(1-t^{\frac{1}{2}(h-1)})}{\log t}.$$

The existence and uniqueness of the solution for this intial value problem is proved from a result on p. 4 of Cesari (1963).

$\phi(t)$ blows up at a certain point T. Assuming C is a constant,

$$\frac{d}{dt}\left(-\frac{1}{\phi(t)+C}\right)=\frac{1}{2}.$$

Hence, for $\tau_1<\tau_2$,

$$\frac{1}{\phi(\tau_1)+C}-\frac{1}{\phi(\tau_2)+C}=\frac{1}{2}(\tau_2-\tau_1).$$

We estimate the upper bound T. The minimum of $C(t)$ in (τ_{k-1},τ_k) is $C(\tau_k)$. We determine ϕ_k recursively by the following:

$$\frac{1}{\phi_0+C(\tau_1)}-\frac{1}{\phi_1+C(\tau_1)}=\frac{1}{2}\tau,$$

$$\frac{1}{\phi_1+C(\tau_2)}-\frac{1}{\phi_2+C(\tau_2)}=\frac{1}{2}\tau,$$

$$\vdots$$

$$\frac{1}{\phi_{k-1}+C(\tau_k)}-\frac{1}{\phi_k+C(\tau_k)}=\frac{1}{2}\tau,$$

where $\tau_k=k\tau$. We continue until there is no $\phi_k>0$ satisfying the equation. The τ_k determined by such a procedure is an upper bound $\bar{T}$ of T.

We discuss the lower bound of T. For $t\leqq\tau_0<1$,

$$C(t)\leqq\left(-\frac{t^{\frac{1}{2}(h-1)}}{\log\tau_0}+\frac{1}{\log\tau_0}\right)\times 2.$$

For $0<t\leqq\tau_0$, if there is an $\varepsilon(\tau_0)$ such that

$$\phi(t)\leqq\frac{2\varepsilon(\tau_0)}{-\log\tau_0}(t^{\frac{1}{2}(h-1)}-1),$$

then

$$2\phi'(t)\leqq\frac{4(1+\varepsilon(\tau_0))^2}{(\log\tau_0)^2}(t^{h-1}-2t^{\frac{1}{2}(h-1)}+1)$$

for t with $0<t\leqq\tau_0$. Under the above assumption

$$\phi(t) \leqq \frac{2(1+\varepsilon(\tau_0))^2}{(\log \tau_0)^2}\left(\frac{1}{h}t^h - 2\frac{1}{\frac{1}{2}h+\frac{1}{2}}t^{\frac{1}{2}(h+1)} + t\right)$$

$$= \frac{2(1+\varepsilon(\tau_0))^2}{(\log \tau_0)^2}\left(\frac{1}{h}t^h - \frac{4}{h+1}t^{\frac{1}{2}(h+1)} + t\right) \equiv \bar{\phi}(t).$$

If for $0 < t < \tau_0$,

$$\phi(t) \leqq \frac{2\varepsilon(\tau_0)}{-\log \tau_0}(t^{\frac{1}{2}(h-1)} - 1),$$

and

$$\bar{\phi}(\tau_0) \leqq \frac{2\varepsilon(\tau_0)}{-\log \tau_0}(\tau_0^{\frac{1}{2}(h-1)} - 1),$$

then

$$\phi(t) \leqq \bar{\phi}(t) \quad \text{for} \quad 0 < t < \tau_0.$$

Putting $\bar{\phi}(\tau_0) = \phi_0$, we determine ϕ_k recursively by following:

$$\frac{1}{\phi_0 + C(\tau_0)} - \frac{1}{\phi_1 + C(\tau_0)} = \frac{1}{2}\tau$$

$$\frac{1}{\phi_1 + C(\tau_1)} - \frac{1}{\phi_2 + C(\tau_1)} = \frac{1}{2}\tau$$

$$\vdots$$

We continue until there is no $\phi_k > 0$ for which

$$\frac{1}{\phi_{k-1} + C(\tau_{k-1})} - \frac{1}{\phi_k + C(\tau_{k-1})} = \frac{1}{2}\tau$$

where $\tau_k = \tau_0 + k\tau$. The τ_{k-1} determined by such a procedure is a lower bound of T.

Using the relation $t = e^{-2s}$, we can obtain the bounds for $a(h)$, that is to say, $-\frac{1}{2}\log \bar{T}$ is a lower bound and $-\frac{1}{2}\log \underline{T}$ is an upper bound for $a(h)$.

In Table 1, we give the lower bounds and upper bounds for $h = 0.01\,k$, $k = 1, 2, \cdots, 99$, in which $\varepsilon(\tau_0) = 10^{-2}$, $\tau_0 = 10^{-6}$ and $\tau = 10^{-5}$. The calculation is performed by the significands of 15 decimal places. In Figure 1 we plot the values of Table 1. Obviously $a(0) = 0$ and $a(1) = \infty$.

4. Applications and discussion

Random packing models for elections in Japan are introduced in the previous works by Itoh and Ueda (1978), (1979), and Itoh (1978). Consider a stick of length $x \geqq 2d$. The stick is divided into two sticks with length x_1 and x_2 such that $x_1 \geqq d$ and $x_2 \geqq d$. Each possible division is assumed to be equally probable.

TABLE 1

h	$a(h)$ lower bound		$a(h)$ upper bound	h	$a(h)$ lower bound		$a(h)$ upper bound
0.01	−0.20147	---	0.07052	0.51	0.68108	---	0.68124
0.02	−0.06793	---	0.08533	0.52	0.69376	---	0.69390
0.03	−0.00489	---	0.09989	0.53	0.70659	---	0.70671
0.04	0.03619	---	0.11421	0.54	0.71957	---	0.71968
0.05	0.06737	---	0.12830	0.55	0.73271	---	0.73281
0.06	0.09316	---	0.14216	0.56	0.74603	---	0.74612
0.07	0.11559	---	0.15582	0.57	0.75954	---	0.75963
0.08	0.13578	---	0.16928	0.58	0.77326	---	0.77333
0.09	0.15437	---	0.18256	0.59	0.78719	---	0.78726
0.10	0.17174	---	0.19567	0.60	0.80135	---	0.80141
0.11	0.18818	---	0.20860	0.61	0.81576	---	0.81581
0.12	0.20385	---	0.22139	0.62	0.83043	---	0.83048
0.13	0.21892	---	0.23403	0.63	0.84538	---	0.84543
0.14	0.23347	---	0.24653	0.64	0.86064	---	0.86068
0.15	0.24758	---	0.25891	0.65	0.87621	---	0.87624
0.16	0.26133	---	0.27117	0.66	0.89212	---	0.89216
0.17	0.27474	---	0.28332	0.67	0.90841	---	0.90844
0.18	0.28788	---	0.29537	0.68	0.92508	---	0.92511
0.19	0.30077	---	0.30732	0.69	0.94218	---	0.94221
0.20	0.31344	---	0.31919	0.70	0.95974	---	0.95976
0.21	0.32593	---	0.33098	0.71	0.97778	---	0.97780
0.22	0.33826	---	0.34270	0.72	0.99635	---	0.99637
0.23	0.35044	---	0.35435	0.73	1.01550	---	1.01551
0.24	0.36250	---	0.36595	0.74	1.03526	---	1.03527
0.25	0.37446	---	0.37750	0.75	1.05569	---	1.05570
0.26	0.38632	---	0.38901	0.76	1.07685	---	1.07686
0.27	0.39810	---	0.40047	0.77	1.09881	---	1.09882
0.28	0.40981	---	0.41191	0.78	1.12164	---	1.12165
0.29	0.42146	---	0.42333	0.79	1.14543	---	1.14544
0.30	0.43307	---	0.43472	0.80	1.17028	---	1.17029
0.31	0.44464	---	0.44611	0.81	1.19630	---	1.19630
0.32	0.45619	---	0.45749	0.82	1.22362	---	1.22363
0.33	0.46771	---	0.46887	0.83	1.25241	---	1.25242
0.34	0.47923	---	0.48026	0.84	1.28284	---	1.28285
0.35	0.49075	---	0.49166	0.85	1.31514	---	1.31515
0.36	0.50227	---	0.50308	0.86	1.34957	---	1.34957
0.37	0.51381	---	0.51453	0.87	1.38645	---	1.38646
0.38	0.52537	---	0.52601	0.88	1.42620	---	1.42620
0.39	0.53696	---	0.53753	0.89	1.46931	---	1.46932
0.40	0.54858	---	0.54909	0.90	1.51645	---	1.51646
0.41	0.56025	---	0.56071	0.91	1.56849	---	1.56849
0.42	0.57197	---	0.57238	0.92	1.62658	---	1.62658
0.43	0.58375	---	0.58412	0.93	1.69238	---	1.69239
0.44	0.59560	---	0.59593	0.94	1.76831	---	1.76832
0.45	0.60752	---	0.60782	0.95	1.85811	---	1.85812
0.46	0.61953	---	0.61979	0.96	1.96807	---	1.96807
0.47	0.63162	---	0.63186	0.97	2.10998	---	2.10998
0.48	0.64382	---	0.64403	0.98	2.31035	---	2.31035
0.49	0.65612	---	0.65631	0.99	2.65391	---	2.65391
0.50	0.66854	---	0.66871				

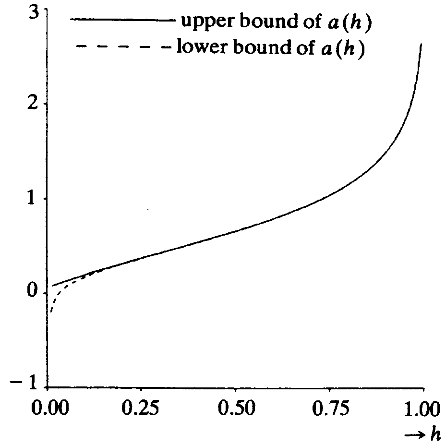

Figure 1

TABLE 2

$x = 25$

h	$\exp\{-a(h)(x+1)\}$ upper bound		lower bound		frequency of $L(x) \geqq h$ per 10000 computer experiments
0.01	188.342682	---	0.1598697	---	4624
0.02	5.8480263	---	0.1087598	---	2600
0.03	1.1356907	---	0.0744817	---	1517
0.04	0.3902792	---	0.0513327	---	930
0.05	0.1734716	---	0.0355891	---	595
0.06	0.0887375	---	0.0248158	---	357
0.07	0.0495199	---	0.0173982	---	237
0.08	0.0292939	---	0.0122601	---	160
0.09	0.0180695	---	0.0086805	---	111
0.10	0.0115015	---	0.0061746	---	68
0.11	0.0075021	---	0.0044107	---	53
0.12	0.0049905	---	0.0031633	---	35
0.13	0.0033730	---	0.0022773	---	25
0.14	0.0023108	---	0.0016452	---	20
0.15	0.0016010	---	0.0011925	---	13
0.16	0.0011199	---	0.0008671	---	10
0.17	0.0007901	---	0.0006322	---	7
0.18	0.0005615	---	0.0004622	---	3
0.19	0.0004016	---	0.0003387	---	2
0.20	0.0002889	---	0.0002488	---	2
0.21	0.0002088	---	0.0001831	---	2
0.22	0.0001515	---	0.0001350	---	1
0.23	0.0001104	---	0.0000997	---	0
0.24	0.0000807	---	0.0000738	---	0
0.25	0.0000591	---	0.0000546	---	0

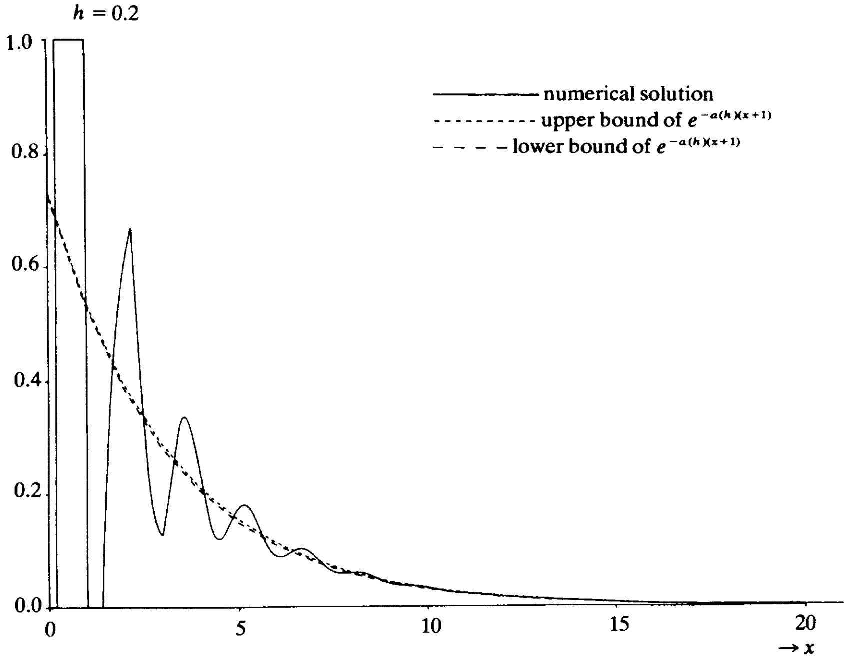

Figure 2

Such division is continued until all sticks are shorter than $2d$. The sticks obtained by such procedure correspond to gaps generated by one-dimensional random packing. (See Ney (1962).) Consider an election in a certain constituency. The length x corresponds to the total votes obtained by candidates of a certain political party, for example the Liberal Democratic Party. The party nominates a candidate if he can obtain at least d votes. The length of each stick, which results from the above procedure of division, corresponds to the votes obtained by a candidate nominated by the party. We must remark that in Japan each voter can vote for one candidate at an election even if the fixed number is more than one. Evidence for this model is given in the previous papers by Itoh and Ueda (1978), (1979) and Itoh (1978).

The present results give one of the checks for the computer experiments which are used in the model fitting. Numerical calculations show that $f(x)$ does not oscillate very much for larger x. (See Figure 2.) For smaller x, we can calculate directly from Equation (3). For larger x for example $x = 20, 21, \cdots, 30$, the approximation by $e^{-a(h)(x+1)}$ fits well to the result of computer experiments for $P(L(x) \geqq h)$, as shown in Table 2 for $x = 25$. A stricter treatment of this discussion is our future problem.

We remark that the maximum of lengths of gaps can be discussed in the same way.

Acknowledgement

The author is most grateful to Dr J. M. Hammersley, who suggested to him an essential idea for numerical calculations of $a(h)$. He thanks Miss S. Ueda for computer programs. Thanks are also due to the referee for helpful comments.

References

BANKÖVI, G. (1962) On gaps generated by a random space filling procedure, *Publ. Math. Inst. Hungar. Acad. Sci.* **7**, 395–407.

CESARI, L. (1963) *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations.* Springer-Verlag, Berlin.

DE BRUIJN, N. G. (1958) *Asymptotic Methods in Analysis.* North-Holland, Amsterdam.

DVORETZKY, A. AND ROBBINS, H. (1964) On the 'parking problem'. *Publ. Math. Inst. Hungar. Acad. Sci.* **9**, 209–225.

FELLER, W. (1971) *An Introduction to Probability Theory and its Applications,* II. Wiley, New York.

ITOH, Y. (1978) Random packing model for nomination of candidates and its application to elections of the House of Representatives in Japan. *Proc. Internat. Conf. Cybernetics and Society* **1**, 432–435.

ITOH, Y. AND UEDA, S. (1978) Note on random packing models for an analysis of elections (in Japanese). *Proc. Inst. Statist. Math.* **25**, 23–27.

ITOH, Y. AND UEDA, S. (1979) A random packing model for elections. *Ann. Inst. Statist. Math.* **31**, 157–167.

NEY, P. E. (1962) A random space filling problem. *Ann. Math. Statist.* **33**, 702–718.

RÉNYI, A. (1958) On a one-dimensional problem concerning random space-filling. *Publ. Math. Inst. Hungar. Acad. Sci.* **3**, 109–127.