The Annals of Probability
1980, Vol. 8, No. 1, 157–163

THE ASYMPTOTIC BEHAVIOR OF SPACINGS UNDER KAKUTANI'S MODEL FOR INTERVAL SUBDIVISION

BY RONALD PYKE

University of Washington

If $X_1, X_2, \cdots$ are random variables with values in (0, 1), let $D_{n1}, \cdots, D_{n,n+1}$ denote the $n+1$ spacings given by the first n observations, $X_1, \cdots, X_n$. If G_n^* denotes the empirical distribution function of the normalized spacings $\{(n+1)D_{ni}\}$, it is proved in this paper that under the Kakutani model in which X_m is a uniform random variable over the largest spacing determined by $X_1, \cdots, X_{m-1}$, with probability one $G_n^* \to G$ uniformly, where G is the uniform distribution function on (0, 2). This is in sharp contrast to the known exponential limiting distribution when the X_i are independent uniform random variables on (0, 1).

1. Introduction. Let $\{X_n: n \geqslant 1\}$ be a sequence of random variables (rv's) taking values in (0, 1) and let $X_{n1} \leqslant \cdots \leqslant X_{nn}$ represent the ordered values of $\{X_1, \cdots, X_n\}$. Define the spacings

$$D_{ni} = X_{ni} - X_{n,i-1}, \qquad 1 \leqslant i \leqslant n+1, \quad \text{with } X_{n0} = 0, \quad X_{n,n+1} = 1,$$

and let $D_{n1}^* \leqslant \cdots \leqslant D_{n,n+1}^*$ denote the ordered spacings.

We are interested in two probability models for $\{X_n: n \geqslant 1\}$. The first is the usual model for the random subdivision of the unit interval, namely that in which the X_n's are independent $U(0, 1)$ rv's. We will refer to this as the U-model (for *usual* or *uniform*). The second model will be referred to as the K-model (for Kakutani) in which X_1 is a $U(0, 1)$ rv, and conditionally given $\{X_1, \cdots, X_{n-1}\}$, X_n is uniformly distributed over the largest subinterval formed by $X_1, \cdots, X_{n-1}$, namely the one whose length is $D_{n-1,n}^*$. Although in [3], Kakutani does not specifically consider this probabilistic context, the K-model is clearly motivated by his 'α-maximal refinements'.

If F_n denotes the empirical df of $\{X_1, \cdots, X_n\}$, Van Zwet [9] established that the Glivenko-Cantelli result ($F_n \to F$ uniformly, with probability one) holds in the K-model, with F equal to the U(0, 1) distribution function (df), as was conjectured by Kakutani. For the U-model, this result is the classical one proved by Glivenko [2] and Cantelli [1] in 1933. Unfortunately, these results fail to exhibit any distinguishability between the two models.

The K-method of random division of the interval [0, 1] should, by its nature, result in 'more uniform' spacings than those of the U-method. This is intuitively clear since the largest spacings are always broken down in the former, whereas in the latter, the largest interval may remain untouched for several iterations while at

Received March 30, 1978.

AMS 1970 *subject classifications.* Primary 60F15; secondary 60K99.

Key words and phrases. Kakutani model, spacings, normalized spacings, Glivenko-Cantelli theorem, empirical distribution function.

the same time, the smaller intervals are consequently being divided into even smaller intervals. In this paper, we consider the empirical df of the normalized spacings $\{(n+1)D_{ni}\}$ rather than of the subdivision points $\{X_{ni}\}$ and show that the corresponding Glivenko-Cantelli results manifest the greater uniformity of the spacings under the K-model.

In 1955, Blum (cf. the appended note to [10] in which convergence in probability was established) showed that under the U-model, the empirical df's of the normalized spacings converge uniformly, with probability 1, to the Exp(1) df, $H(x) = 1 - e^{-x}$, $x > 0$. Under the K-model, we show below that these empirical df's again converge uniformly, but to a $U(0, 2)$ df. These limits are quite different. In particular, the one limit has an unbounded support while the other has the bounded support, (0, 2).

The proof given by Van Zwet introduces the brilliant idea of focusing attention on the subdivision process at the random times N_s, defined in (1.1) below. The reader of this paper will quickly observe the author's dependence upon this idea and the general method of Van Zwet's proof.

Before stating the main result, we introduce further notation as follows. For $s > 0$ and $I^+ = \{0, 1, 2, \cdots\}$, define

$$N_s = \min\{n \in I^+ : D^*_{n,n+1} \leqslant s\} \tag{1.1}$$

where for $n = 0$, we set $D^*_{01} = D_{01} = 1$. Interpret $\min \varnothing = +\infty$. The empirical df of the spacings $\{D_{n1}, \cdots, D_{n,n+1}\}$ is denoted by

$$G_n(x) = (n+1)^{-1}\Sigma_{i=1}^{n+1}\varepsilon(x - D_{ni}) \tag{1.2}$$

where $\varepsilon(u) = 1$ or 0 according as $u \geqslant 0$ or $u < 0$. In what follows, it is simpler to work at first with the counting function

$$K_n(x) = (n+1)G_n(x) = {}^{\#}\{D_{ni} : D_{ni} \leqslant x\}. \tag{1.3}$$

We shall study these functions at the random sample sizes N_s, and so introduce

$$G(x, s) = G_{N_s}(x) \quad \text{and} \quad K(x, s) = K_{N_s}(x). \tag{1.4}$$

In the construction of the sequence $\{X_i\}$ write U for X_1 to emphasize its $U(0, 1)$ df and observe that $K_1(\cdot)$ assigns unit mass to the values U and $1 - U$. It is then straightforwardly checked that the following recursion relation holds:

$$\begin{aligned} K(x, s) &= \int_0^1 K(x/y, s/y)\, dK_1(y) \\ &= K^{(1)}(x/U, s/U) + K^{(2)}(x/(1-U), s/(1-U)) \end{aligned} \tag{1.5}$$

where the superscripts on K indicate that the two functions are different independent versions of the rv's indicated by the notation (cf. Lootgieter [4]). To check (1.5), note that the two intervals $(0, u]$ and $(u, 1]$ will be divided independently by the same scheme until all spacings in each are less than s. But the division of $(0, u]$ until all spacings are less than s is the same as performing a division of $(0, 1]$ until all spacings are less than s/u. A similar observation holds for $(u, 1]$. The boundary conditions on K are $K(x, s) = N_s + 1$ if $x \geqslant s$.

The empirical df we wish to study is that of the normalized spacings, $(n+1)D_{ni}$; it is defined by

$$G_n^*(y) = G_n(y/(n+1)) = (n+1)^{-1}K_n(y/(n+1)). \tag{1.6}$$

If one takes $n = N_s$, and writes $G^*(y, s) = G_{N_s}^*(y)$ then

$$G^*(y, s) = (N_s + 1)^{-1}K(y/(N_s + 1), s). \tag{1.7}$$

A key result in [9] relative to our study is the fact that

$$sN_s \to 2 \quad \text{a.s.} \qquad \text{as } s \text{ converges to zero.} \tag{1.8}$$

(cf. [7], (2.8) in which sN_s is shown to converge a.s. over the sequence $\{n^{-2}: n \geqslant 1\}$. Since $N_s \nearrow$ as $s \searrow$, one has for $(m+1)^{-2} < s \leqslant m^{-2}$ that

$$(m+1)^{-2}N_{m^{-2}} \leqslant sN_s \leqslant m^{-2}N_{(m+1)^{-2}}$$

which yields the convergence of sN_s.) An immediate consequence of this result is the possibly surprising result that the limit of G_n^*, if it exists, will have its support contained in the bounded interval, (0, 2]. To see this, let $M_n := D_{n,n+1}^*$ denote the maximum spacing at the nth stage.

LEMMA 1.

$$nM_n \to 2 \quad \text{a.s.}$$

PROOF. Notice that $M_n > s$ if and only if $N_s > n$. Thus, M is the inverse of N. In particular, $N_{M_n} = n$ so that $nM_n = M_nN_{M_n} \to 2$ a.s. by (1.8). □

2. The main result. The first preliminary step in the proof of Theorem 1 below is the evaluation of the mean and variance of $K(x, s)$. First of all, set $\mu(x, s) = E[K(x, s)]$. Notice that all moments of K are finite since $K(x, s) \leqslant N_s + 1$ and N_s was shown to have finite moments by Van Zwet [9]. From (1.5) and the fact that X_1 is a $U(0, 1)$ rv, we obtain

$$\mu(x, s) = 2\int_0^1 \mu(x/u, s/u)\, du. \tag{2.1}$$

According to Van Zwet ([9], (2.5)),

$$E(N_s + 1) = 2/s, \qquad 0 < s < 1. \tag{2.2}$$

But $K(x, s) = N_s + 1$ for $x \geqslant s$. Moreover, for $s \geqslant 1$, $N_s = 0$ so that $\mu(x, s) = \varepsilon(x-1)$, where $\varepsilon(u) = 1$ or according as $u \geqslant 0$ or $u < 0$. Substitution of these values into (2.1) yields for $0 < x \leqslant s < 1$,

$$\mu(x, s) = 2\int_s^1 \mu(x/u, s/u)\, du + 2\int_0^s \varepsilon(x/u - 1)\, du.$$

In the second integral, the integrand is 0 on $(x, s]$ and 1 on $(0, x]$. Hence

$$\mu(x, s) = 2\int_s^1 \mu(x/u, s/u)\, du + 2x. \tag{2.3}$$

Let $r = x/s$ and introduce the new function

$$g(x, r) = x^{-1}\mu(x, x/r), \qquad 0 < x \leqslant 1, \quad x < r \leqslant 1.$$

Then (2.3) transforms into

$$g(x, r) = 2\int_{x/r}^{1} u^{-1} g(x/u, r)\, dy + 2 = 2\int_{x}^{r} v^{-1} g(v, r)\, dv + 2.$$

Differentiation of this with respect to x yields

$$g_1(x, r) = -(2/x) g(x, r) \tag{2.4}$$

whose solution is $g(x, r) = x^{-2} c(r)$ for some function c depending only on r. This translates back into the solution for μ;

$$\mu(x, s) = x g(x, x/s) = x^{-1} c(x/s).$$

For $x \leqslant 1$, $K(x, s) \searrow K_1(x)$ as $s \nearrow 1$. Hence $\lim_{s \to 1-} \mu(x, s) = E[K_1(x)]$ by the monotone convergence theorem. Now $\{D_{11}, D_{12}\}$, the spacings after the first stage, have the same distribution as $\{U, 1 - U\}$ where U is $U(0, 1)$. Hence $P[D_{11} \leqslant x] = P[D_{12} \leqslant x] = x$ for $0 < x < 1$ so that $E[K_1(x)] = 2x$ if $0 < x < 1$. Therefore,

$$\lim_{s \to 1-} \mu(x, s) = \lim_{s \to 1-} x^{-1} c(x/s) = x^{-1} c(x+) = 2x,$$

and hence $c(x +) = 2x^2$, for $0 < x < 1$. This implies that $c(x) = 2x^2$ and so

$$\mu(x, s) = x^{-1} 2(x/s)^2 = 2x/s^2 \tag{2.5}$$

for $0 < x \leqslant s \leqslant 1$. This checks with (2.2) for $x = s$.

Consider now the variance of $K(x, s)$, denoted by

$$v(x, s) = E\big[K(x, s) - 2x/s^2\big]^2, \qquad 0 < x \leqslant s \leqslant 1.$$

For convenience, we restrict out attention to $s < \frac{1}{2}$. By (1.5)

$$\begin{aligned} v(x, s) &= \int_0^1 E\big[K(x/u, s/u) + K(x/(1-u), s/(1-u)) - 2x/s^2\big]^2\, du \\ &= 2\int_0^{\frac{1}{2}} \big[v(x/u, s/u) + v(x/(1-u), s/(1-u))\big]\, du \\ &\quad + 2\int_0^{\frac{1}{2}} \big[\mu(x/u, s/u) + \mu(x/(1-u), s/(1-u)) - 2x/s^2\big]^2\, du. \end{aligned}$$

But, for $0 < x \leqslant s \leqslant 1$, $\mu(x, s) = 2x/s^2$. Thus for $s < \frac{1}{2}$ and $u \geqslant s$ the intergrand in the second integral is zero. For $u < s$, $\mu(x/u, s/u) = \varepsilon(x/u - 1) = \varepsilon(x - u)$ and $\mu(x/(1-u), s/(1-u)) = 2x(1-u)/s^2$, since $s < 1 - u$ when $u, s < \frac{1}{2}$. Thus,

$$\begin{aligned} v(x, s) &= 2\int_0^1 v(x/u, s/u)\, du + 2\int_x^s (2xu/s^2)^2\, du + 2\int_0^x (1 - 2xu/s^2)^2\, du \\ &= 2x\int_x^\infty w^{-2} v(w, sw/x)\, dw + 8x^2(s^3 - x^3)/3s^4 + 2x - 4x^3/s^2 + 8x^5/3s^4 \\ &= 2x\int_x^\infty w^{-2} v(w, sw/x)\, dw + 2x + 8x^2/3s - 4x^3/s^2. \end{aligned}$$

Set $r = x/s$ and $\beta(x, r) = x^{-1} v(x, x/r)$ so that

$$\beta(x, r) = 2\int_x^\infty w^{-1} \beta(w, r) + 2 + 8r/3 - 4r^2. \tag{2.6}$$

Observe now that $v(x, s) = 0$ if $s > 1$, or equivalently, $\beta(x, r) = 0$ if $x > r$. Thus (2.6) reduces to

$$\beta(x, r) = 2\int_x^r w^{-1} \beta(w, r)\, dw + 2 + 8r/3 - 4r^2.$$

Differentiation with respect to x yields

$$\beta_1(x, r) = (-2/x)\beta(x, r)$$

whose solution, as for (2.4), is

$$\beta(x, r) = x^{-2}b(r), \qquad 0 < x < r/2 < \tfrac{1}{2},$$

for some value $b(r)$ independent of x. Therefore,

$$v(x, s) = x\beta(x, x/s) = x^{-1}b(x/s). \tag{2.7}$$

Although the evaluation of $b(\cdot)$ is of importance for a study of the limiting df of G_n, it is only necessary at this point to know the order of magnitude of the variance as $s \to 0$.

Consider the approximate empirical df

$$\bar{G}(y, s) = (s/2)K(ys/2, s) \tag{2.8}$$

obtained from $G(y, s)$ by substituting $s/2$ for $N_s + 1$ therein. (Recall (1.8).) For $0 < y < 2$ and $0 < s < \frac{1}{2}$, it follows from (2.5) and (2.7) that

$$E\big[\bar{G}(y, s)\big] = y/2, \qquad \operatorname{Var}\big(\bar{G}(y, s)\big) = c_y s$$

where $c_y = (y/2)b(y/2)$ is constant in s. By Chebychev's inequality and the Borel-Cantelli lemma, $\bar{G}(y, s) \to y/2$ a.s. when s converges to 0 over a convergent sequence, such as $\{m^{-2}: m \geqslant 2\}$ as used by Van Zwet [9]. For $s' < s \leqslant s''$ one obtains

$$(s'/s'')\bar{G}(ys'/s'', s'') \leqslant \bar{G}(y, s) \leqslant (s''/s')\bar{G}(ys''/s', s')$$

directly from the definition of $\bar{G}$ and the monotoneity properties, $K(\cdot, s) \nearrow$ and $K(x, \cdot) \searrow$. For $s' = (m + 1)^{-2}$ and $s'' = m^{-2}$, $s'/s'' = m^2/(m + 1)^2 \to 1$. This completes the proof of

LEMMA 2. *For* $0 < y < 2$, $\bar{G}(y, s) \to y/2$ a.s. *as* $s \to 0$.

This brings us to the main result concerning the uniform convergence of the empirical df's G_n^* of the normalized spacings.

THEOREM 1. *With probability* 1, G_n^* *converges uniformly, as* $n \to \infty$, *to the* $U(0, 2)$ df G; $G(y) = y/2$ *for* $0 < y < 2$.

PROOF. By Polya's result (cf. [6], page 120), Lemma 2 implies the uniform convergence of $\bar{G}(\cdot, s)$ to G as $s \to 0$. It remains to observe that because of (1.6) and (2.8)

$$G_{N_s}^*(y) = G^*(y, s) = \frac{2}{s(N_s + 1)}(s/2)K((2y/s(N_s + 1)(s/2), s)$$

$$= \frac{2}{s(N_s + 1)}\bar{G}(y(2/s(N_s + 1)), s)$$

and this has the same limit as $\bar{G}$ since as $s \to 0$, $sN_s \to 2$ a.s.. Since $N_s \to \infty$ a.s. over all the integers the proof is complete. □

3. Remarks. By considering the empirical df of the spacings rather than the points of subdivision, we have indicated the differences that exist between the two models of subdivision. Another indication of the increased uniformity of the K-model is seen as follows. Define density functions

$$f_n(x) = 1/(n+1)D_{ni} \quad \text{if} \quad X_{n,i-1} < x \leqslant X_{ni}.$$

These are the density functions of the df obtained by making F_n piecewise linear. Consider the L_1-distance between f_n and the uniform density,

$$\begin{aligned} d_n &:= \int_0^1 |f_n(x) - 1|\,dx = 2\int_0^1 [f_n(x) - 1]^+\,dx \\ &= 2\big(^{\#}\{i: (n+1)D_{ni} \leqslant 1\}/(n+1) - \Sigma\{D_{ni}: (n+1)D_{ni} \leqslant 1\}\big) \\ &= 2\int_0^1 (1-y)\,dG_n^*(y). \end{aligned}$$

Thus for the Kakutani model

$$d_n \to 2\int_0^1 (1-y)\,dG(y) = \int_0^1 (1-y)\,dy = \tfrac{1}{2} \quad \text{a.s.},$$

while for the usual model

$$d_n \to 2\int_0^1 (1-y)\,dH(y) = 2\int_0^1 (1-y)e^{-y}\,dy = 2/e \approx .736 \quad \text{a.s.}$$

Although much more could be attempted for the normalized spacings under the K-model, in parallel for example to the extensive literature for the U-model (cf. the survey [5]), it is not clear that this would be the best direction. Of greater importance might be a search for general methods which would facilitate the study of the vast spectrum of models 'between' the two studied here. A general model might be defined by probabilities $\{p_{ni}: 1 \leqslant i \leqslant n+1\}$ which after the nth stage would give the probability of choosing the next observation, X_{n+1}, uniformly from the ith interval $(X_{n,i-1}, X_{ni}]$. For example, $p_{nn} = p_{n,n+1} = \frac{1}{2}$ would determine the model in which the largest and the second largest are equally likely to be subdivided. It would be of interest to characterize those models which are 'more uniform' than the usual U-model.

Since this paper was submitted, the paper [7] by Slud appeared. In this paper another proof of Kakutani's conjecture is given by a combinatorial method that extends to the generalization in which the dividing measure of the largest spacing is arbitrary and not necessarily uniform. In addition, rates of convergence are considered. Relative to the spacings of the subdivision, Slud shows also (Proposition 3.2 of [7]) that their entropy $-\sum_{i=1}^{n+1} D_{ni} \log D_{ni}$ is stochastically larger under the K-model than under the U-model. A related paper by Slud is [8].

Acknowledgment. The author greatly acknowledges discussions with Professor A. V. Peterson, and correspondence, through Professor Peterson, with Professor G. F. Hermann about the related deterministic problem of optimal sequential interval splitting.

REFERENCES

[1] CANTELLI, F. P. (1933). Sulla determinazione empirica di une legge di probabilita. *Giorna Ist. Ital. Attuari* **4** 421–424.

[2] GLIVENKO, V. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorna Ist. Ital. Attuari* **4** 92–99.

[3] KAKUTANI, S. (1975). A problem of equidistribution on the unit interval [0, 1]. *Lecture Notes in Math.* **541** 369–376. Springer, Berlin.

[4] LOOTGIETER, J. C. (1977). Sur la repartition des suites de Kakutani. *C. R. Acad. Sci. Paris.* A **285** 403–406. Also, *Ann. Inst. Henri Poinaré* **13** 385–410.

[5] PYKE, R. (1965). Spacings. *J. Roy. Statist. Soc. Ser. B* **27** 395–449.

[6] RAO, C. R. (1973). *Linear Statistical Inference and its Applications, 2nd Ed.* Wiley New York.

[7] SLUD, ERIC (1978a). Entropy and maximal spacings for random partitions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **41** 341–352.

[8] SLUD, ERIC (1978b). On entropy and random spacings of the interval. Tech. Rpt. No. 78-8, MD78-10-ES, Univ. Maryland.

[9] VAN ZWET, W. R. (1978). A proof of Kakutani's conjecture on random subdivision of longest intervals. *Ann. Probability* **6** 133–137.

[10] WEISS, L. (1955). The stochastic convergence of a function of sample successive differences. *Ann. Math. Statist.* **26** 532–536.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195