Random division of an interval*)

*by F. W. Steutel***)*

Summary

The well-known relation between random division of an interval and the Poisson process is interpreted as a Laplace transformation. With the use of this interpretation a number of (in part known) results is derived very easily.

1. Introduction

The study of distributions arising from the random division of an interval has attracted many statisticians and mathematicians over a long period of time. Derivations of the distribution function of the largest subinterval, for instance, are given by WHITWORTH [17] in 1897 and by GIRAULT [7] in 1962 (independent of each other).

These and many other derivations are based on elementary combinatorial methods, which are useful only to obtain simple results. Others (see e.g. [12] and [16]) use geometrical methods, which often involve rather intricate integration in n-dimensional space.

As an often quite effective alternative approach one can use the relation between random division of an interval and the Poisson process. This relation is well-known but has only been used incidentally (MORAN [12], DOMB [4], DWASS [5]) and not always in the most effective manner.

In this paper the relation with the Poisson process is interpreted as a Laplace transformation (formula (1)), the, often quite simple, inversion of which yields many results with surprisingly little effort. This has been done by DWASS for the special case of linear combinations of subintervals. Dwass uses this method more or less "ad hoc" and it seems to be generally unknown (there is no mention of it for instance in [8]).

Here we give a detailed account of it in the general case. Some general formulae are derived and in order to demonstrate how the method works it is applied to derive a number of explicit results, most of which are known. It appears that this method often requires considerably less computation than the ones mentioned above. In the final section it is shown how the method can be used to obtain asymptotic results.

---

*) A first draft of this article appeared as Report S 314 (1963) of the Mathematical Centre, Amsterdam.

**) Technische Hogeschool Twente, Enschede, Netherlands.

2. Notations and definitions

For reasons that will become clear later we consider the interval $(0, t)$ for arbitrary positive t instead of the usual interval $(0,1)$. The interval $(0, t)$ is divided into n subintervals by $n - 1$ random points, i.e. points with coordinates:

$$\underline{y}_1(t),\ \underline{y}_2(t), \cdots, \underline{y}_{n-1}(t)$$

drawn independently from a rectangular distribution on $(0, t)$. Denoting by

$$\underline{y}_{(1)}(t),\ \underline{y}_{(2)}(t), \cdots, \underline{y}_{(n-1)}(t,$$

these coordinates in increasing order and putting

$$\underline{y}_{(0)}(t) = 0,\ \underline{y}_{(n)}(t) = t,$$

for the lengths

$$\underline{x}_1(t),\ \underline{x}_2(t), \cdots, \underline{x}_n(t)$$

of the subintervals we have

$$\underline{x}_j(t) = \underline{y}_{(j)}(t) - \underline{y}_{(j-1)}(t) \qquad (j = 1, 2, \cdots, n).$$

We will be concerned with the distributions of functions of the $\underline{x}_j(t)$. The increasingly ordered interval lengths will be denoted by

$$\underline{x}_{(1)}(t),\ \underline{x}_{(2)}(t), \cdots, \underline{x}_{(n)}(t).$$

For the sake of brevity we will sometimes omit the argument t if $t = 1$ and we write

$$\underline{x}_j,\ \underline{u}_j \quad \text{etc.}$$

instead of

$$\underline{x}_j(1),\ \underline{u}_j(1) \quad \text{etc.}$$

The notation

$$(\underline{a}_1,\ \underline{a}_2, \cdots, \underline{a}_m) \triangleq (\underline{b}_1,\ \underline{b}_2 \cdots, \underline{b}_m)$$

will express the fact that $(\underline{a}_1, \underline{a}_2, \cdots, \underline{a}_m)$ and $(\underline{b}_1, \underline{b}_2, \cdots, \underline{b}_m)$ have the same distribution.

3. Relation to Poisson process

In this and the following sections

$$\underline{u}_1(\tau),\ \underline{u}_2(\tau) \cdots, \underline{u}_n(\tau)$$

will denote independent random variables with distribution function

$$P\{\underline{u}_j(\tau) \leqslant u\} = 1 - e^{-\tau u} \qquad (u \geqslant 0,\ \tau \text{ some constant} > 0).$$

We will sometimes omit the argument τ if $\tau = 1$.

The content of the following lemma is well known:

L e m m a 1: *the conditional distribution of*

$$[\underline{u}_1(\tau),\ \underline{u}_1(\tau) + \underline{u}_2(\tau)\ \cdots,\ \underline{u}_1(\tau) + \underline{u}_2(\tau) + \cdots + \underline{u}_{n-1}(\tau)]$$

given that

$$\underline{u}_1(\tau) + \underline{u}_2(\tau) + \cdots + \underline{u}_n(\tau) = t$$

is the same as the distribution of

$$[\underline{x}_1(t),\ \underline{x}_1(t) + \underline{x}_2(t), \cdots, \underline{x}_1(t) + \underline{x}_2(t) + \cdots + \underline{x}_{n-1}(t)].$$

P r o o f: considering the density functions.

$$f(z_1, \cdots, z_{n-1} \mid t),\ g(z_1, \cdots, z_n) \text{ and } h(z_n)$$

of the random vectors

$$[\underline{u}_1(\tau), \cdots, \underline{u}_1(\tau) + \cdots + \underline{u}_{n-1}(\tau) \mid \underline{u}_1(\tau) + \cdots + \underline{u}_n(\tau) = t],\ ^1)$$

$$[\underline{u}_1(\tau), \cdots, \underline{u}_1(\tau) + \cdots + \underline{u}_n(\tau)]$$

and

$$\underline{u}_1(\tau) + \cdots + \underline{u}_n(\tau)$$

respectively, we have

$$f(z_1, \cdots, z_{n-1} \mid t) = \frac{g(z_1, \cdots, z_{n-1}, t)}{h(t)} =$$

$$= \tau^n \exp[-\tau\{z_1 + (z_2 - z_1) + \cdots + (t - z_{n-1})\}](n-1)!\,\tau^{-n} t^{-n+1} e^{\tau t} = \frac{(n-1)!}{t^{n-1}},$$

where the last member is the density function of

$$[\underline{y}_{(1)}(t), \cdots, \underline{y}_{(n-1)}(t)] = [\underline{x}_1(t), \underline{x}_1(t) + \underline{x}_2(t), \cdots, \underline{x}_1(t) + \cdots + \underline{x}_{n-1}(t)].$$

From lemma 1 we deduce:

T h e o r e m 1: *For any Borel-measurable function $f(x_1, \cdots, x_n)$ satisfying*

$$\int_0^\infty \mathcal{E}\,|f(\underline{x}_1(t), \cdots, \underline{x}_n(t)\,|\,t^{n-1} e^{-\tau t} dt < \infty$$

($\mathcal{E}$ denoting mathematical expectation), we have

$$\int_0^\infty \mathcal{E} f(\underline{x}_1(t), \cdots, \underline{x}_n(t))\,t^{n-1} e^{-\tau t} dt = \frac{(n-1)!}{\tau^n}\,\mathcal{E} f(\underline{u}_1(\tau), \cdots, \underline{u}_n(\tau)). \quad (1)$$

[1]) i.e. $[\underline{u}_1(\tau), \ldots, \underline{u}_1(\tau) + \ldots + \underline{u}_{n-1}(\tau)]$ given that $\underline{u}_1(\tau) + \ldots + \underline{u}_n(\tau) = t$.

P r o o f: lemma 1 is equivalent to

$$[\underline{x}_1(t), \cdots, \underline{x}_n(t)] \cong [\underline{u}_1(\tau), \cdots, \underline{u}_n(\tau) \mid \underline{u}_1(\tau) + \cdots + \underline{u}_n(\tau) = t],$$

i.e.

$$P\{\underline{x}_1(t) \leqslant x_1, \cdots, \underline{x}_n(t) \leqslant x_n\} = P\{\underline{u}_1(\tau) \leqslant x_1, \cdots, \underline{u}_n(\tau) \leqslant x_n \mid \underline{u}_1(\tau) + \cdots + \underline{u}_n(\tau) = t\}. \quad (2)$$

Multiplying both sides of (2) by $\frac{\tau^n}{(n-1)!} t^{n-1} e^{-\tau t}$ and integrating one finds

$$\frac{\tau^n}{(n-1)!} \int_0^\infty P\{\underline{x}_1(t) \leqslant x_1, \cdots, \underline{x}_n(t) \leqslant x_n\} t^{n-1} e^{-\tau t} dt = P\{\underline{u}_1(\tau) \leqslant x_1, \cdots, \underline{u}_n(\tau) \leqslant x_n\}, \quad (3)$$

from which (1) follows immediately.

Relation (1) expresses the fact that

$$\frac{(n-1)!}{\tau^n} \mathscr{E} f(\underline{u}_1(\tau), \cdots, \underline{u}_n(\tau))$$

is the Laplace transform with respect to t of

$$t^{n-1} \mathscr{E} f(\underline{x}_1(t), \cdots, \underline{x}_n(t)).$$

Hence the required expectation in the left-hand side of (1) is obtained by inversion of the right-hand side.

We note that (1) can be used to derive distribution functions, as we always have

$$P\{g(\underline{x}_1, \cdots, \underline{x}_n) \leqslant z\} = \mathscr{E} f(\underline{x}_1, \cdots, \underline{x}_n),$$

where

$$f(x_1, \cdots, x_n) = \begin{cases} 1 & \text{if } g(x_1, \cdots, x_n) \leqslant z, \\ 0 & \text{otherwise.} \end{cases}$$

For the joint distribution of $\underline{x}_1(t), \cdots, \underline{x}_n(t)$ we find from (1) (with $\iota(x) = 1$ for $x \geqslant 0$ and $\iota(x) = 0$ for $x < 0$)

$$\int_0^\infty P\{\underline{x}_1(t) > x_1, \cdots, \underline{x}_n(t) > x_n\} t^{n-1} e^{-\tau t} dt =$$

$$= \int_0^\infty \mathscr{E} \prod_{j=1}^n (1 - \iota(x_j - \underline{x}_j(t)) t^{n-1} e^{-\tau t} dt =$$

$$= \frac{(n-1)!}{\tau^n} \mathscr{E} \prod_{j=1}^n (1 - \iota(x_j - \underline{u}_j(\tau)) \frac{(n-1)!}{\tau^n} e^{-\tau(x_1 + \ldots + x_n)},$$

and by inversion (see formula (7))

(where we take $\sum_{j=1}^{n} x_j \leqslant t$ and $x_j \geqslant 0$ for each j)

$$t^{n-1} P\{\underline{x}_1(t) > x_1, \cdots, \underline{x}_n(t) > x_n\} = \left(t - \sum_1^n x_j\right)^{n-1}.$$

In the same way one finds from (3)

$$t^{n-1} P\{\underline{x}_1(t) \leqslant x_1, \cdots, \underline{x}_n(t) \leqslant x_n\} = t^{n-1} - \sum_{j_1=1}^{n} (t - x_{j_1})^{n-1} +$$

$$+ \sum_{1 \leq j_1 < j_2 \leq n} (t - x_{j_1} - x_{j_2})^{n-1} + \cdots + (-)^n \sum_{1 \leq j_1 < j_2 \ldots < j_n \leq n} (t - x_{j_1} - \cdots - x_{j_n})^{n-1}.$$

A formula equivalent to the inversion of (1) occurs in POLLACZEK [14] [2]), where it is derived as a formal identity without interpretation. A special case of (1) is used in DWASS [5].

4. Linear combinations

MAULDON [9], [10] and DWASS [5] consider the distributions of linear combinations of the $\underline{x}_j(t)$ and $\underline{x}_{(j)}(t)$. As for the ordered variables $\underline{u}_{(j)}(\tau)$ we have, writing

$$\underline{v}_j(\tau) = \frac{\underline{u}_{(1)}(\tau)}{n} + \frac{\underline{u}_{(2)}(\tau)}{n-1} + \cdots + \frac{\underline{u}_{(j)}(\tau)}{n-j+1}$$

that

$$[\underline{u}_{(1)}(\tau), \underline{u}_{(2)}(\tau), \cdots, \underline{u}_{(n)}(\tau)] \triangleq [\underline{v}_1(\tau), \underline{v}_2(\tau), \cdots, \underline{v}_n(\tau)]$$

(see e.g. RENYI [15]), by (3) the $\underline{x}_{(j)}(t)$ satisfy

$$[\underline{x}_{(1)}(t), \underline{x}_{(2)}(t), \cdots, \underline{x}_{(n)}(t)] \triangleq [\underline{w}_1(t), \cdots, \underline{w}_n(t)],$$

where

$$\underline{w}_j(t) = \frac{\underline{x}_1(t)}{n} + \frac{\underline{x}_2(t)}{n-1} + \cdots + \frac{\underline{x}_j(t)}{n-j+1}. \tag{4}$$

Hence we may restrict ourselves to linear combinations of the $\underline{x}_j(t)$. If

$$\underline{a}_n(t) = \sum_{j=1}^{n} \alpha_j \underline{x}_j(t),$$

where the α_j are real constants, then by (1)

---

[2]) Prof. dr. J. TH. RUNNENBURG, to whom I am indebted for some useful suggestions, drew my attention to this fact.

$$\int_0^\infty \mathscr{E} e^{-s\underline{a}_n(t)} t^{n-1} e^{-\tau t} dt = \frac{(n-1)!}{\tau^n} \mathscr{E} \exp\{-s\sum \alpha_j \underline{u}_j(\tau)\} = (n-1)! \prod_{j=1}^{n} \frac{1}{s\alpha_j + \tau}. \quad (5)$$

The distribution function of $\underline{a}_n(t)$ can be found by partial fraction expansion of the last member of (5) followed by LAPLACE-inversion with respect to τ and LAPLACE-STIELTJES-inversion with respect to s. The general formula is quite complicated and will not be given. A special case of relation (5) has been used by DWASS [5].

As

$$\underline{a}_n(t) \doteq \underline{a}_n t$$

for the first member of (5) one may write

$$\int_0^\infty \int_0^\infty e^{-sat} dF_n(a)\, t^{n-1} e^{-\tau t} dt = (n-1)! \int_0^\infty (sa+\tau)^{-n} dF_n(a),$$

$F_n(a)$ denoting the distribution function of $\underline{a}_n$, i.e. by (5) we have

$$\mathscr{E}(s\underline{a}_n + \tau)^{-n} = \prod_{j=1}^{n} (s\alpha_j + \tau)^{-1}. \quad (6)$$

Formula (6) was obtained for some special linear combinations of the $\underline{x}_j$ by MAULDON [9] using straightforward but rather complicated analytical methods. In BARTON and DAVID [1] formula (6) occurs interpreted as a generating function of the $\underline{x}_{(j)}$ (c.f. (4)). In MAULDON [10] the inversion of transforms of the type $\mathscr{E}(s\underline{a}_n + \tau)^{-n}$ is considered more generally (see also MAULDON [11]).

To obtain explicit results it is more practical not to use the transforms (5) and (6) but to start directly from (1). If, for instance, we wish to compute the distribution function of $\underline{x}_{(n-k)}(t)$ we have by (1)

$$\int_0^\infty P\{\underline{x}_{(n-k)}(t) \leqslant z\} t^{n-1} e^{-\tau t} dt = \frac{(n-1)!}{\tau^n} \sum_{j=0}^{k} \binom{n}{j} e^{-\tau j z}(1-e^{-\tau z})^{n-j}.$$

Using

$$\int_0^\infty \frac{(t-\alpha)^{p-1}}{\Gamma(p)} \iota(t-\alpha) e^{-\tau t} dt = \frac{e^{-\alpha\tau}}{\tau^p} \qquad (\alpha \geqslant 0, p > 0) \quad (7)$$

where

$$\iota(x) = \begin{cases} 0 & (x < 0) \\ 1 & (x \geqslant 0) \end{cases}$$

we obtain

$$P\{\underline{x}_{(n-k)}(t) \leqslant z\} = \sum_{j=0}^{k} \sum_{l=0}^{n-j} \binom{n}{j} \binom{n-j}{l} (-1)^l \left\{1-(j+l)\frac{z}{t}\right\}^{n-1} \iota(t-(j+l)z),$$

a result which may be found in WHITWORTH [17].

With a little more difficulty the distribution function

$$G_{n,k}(z, t) = P\{\underline{x}_{(n)}(t) + \underline{x}_{(n-1)}(t) + \cdots + \underline{x}_{(n-k+1)}(t) \leqslant z\}$$

as studied in MAULDON [9] may be found by inverting directly

$$\frac{(n-1)!}{\tau^n} P\{\underline{u}_{(n)}(\tau) + \underline{u}_{(n-1)}(\tau) + \cdots + \underline{u}_{(n-k+1)}(\tau) \leqslant z\},$$

which equals (not trivially, by using properties of the Poisson process)

$$\binom{n}{k}\frac{(n-1)!}{(k-2)!}\int_0^z e^{-\tau u}u^{k-2}du \int_0^{z-u} \frac{e^{-\tau x}\left(1-e^{-\frac{\tau x}{k}}\right)^{n-k}}{\tau^{n-k}}\, dx. \tag{8}$$

Application of

$$\sum_{j=0}^{n-k}\binom{n-k}{j}(-1)^{n-k-j}\int_0^\infty \frac{(aj+b-t)^{n-k-1}}{(n-k-1)!}\,\iota(aj+b-t)e^{-\tau t}dt =$$
$$= \frac{e^{-\tau b}(1-e^{-\tau a})^{n-k}}{\tau^{n-k}}$$

and putting $j = p - k$ yields MAULDON's result:

$$G_{n,k}(z,t) =$$
$$= \sum_{p=k+1}^{n} (-1)^{n-p} \frac{n!}{p(p-k)!(n-p)!k!(p-k)^{k-1}(n-k)^{n-k-1}}\left(p\frac{z}{t}-k\right)^{n-1}\iota\left(p\frac{z}{t}-k\right).$$

Inversion of (8) with the use of (7) leads to a more complicated result, which is also given in MAULDON [9].

5. Moments

Formula (1) can be used to obtain the moments of some functions of $\underline{x}_1(t), \cdots, \underline{x}_n(t)$ very easily. Substituting $\tau = 1$ in (1) and using the obvious relation $\underline{x}_j(t) \doteq \underline{x}_j t$ we find

T h e o r e m 2:

if $f(x_1, \cdots, x_n)$ is homogeneous of the degree p, i.e. if

$f(\lambda x_1, \cdots, \lambda x_n) = \lambda^p f(x_1, \cdots, x_n)$ *and if* $\mathscr{E}\,|f(\underline{x}_1, \cdots, \underline{x}_n)| < \infty$,

then

$$\mathscr{E} f(\underline{x}_1, \cdots, \underline{x}_n) = \frac{(n-1)!}{\Gamma(n+p)}\mathscr{E} f(\underline{u}_1, \cdots, \underline{u}_n). \tag{9}$$

For instance we have:

$$\mathscr{E}\underline{x}_1^{a_1-1}\cdots\underline{x}_n^{a_n-1} = (n-1)!\,\frac{\Gamma(a_1)\cdots\Gamma(a_n)}{\Gamma(a_1+\cdots+a_n)}$$

as given in KENDALL and MORAN [8], where it is derived by a geometrical method.

MORAN ([12] and [13]) considers the distribution of

$$\underline{q}_n(t) = \sum_{j=1}^{n} \underline{x}_j^2(t)$$

which obviously is homogeneous of degree 2. Application of (9) yields

$$\mathscr{E}\underline{q}_n^l = \frac{(n-1)!}{(n+2l-1)!}\mathscr{E}(\underline{u}_1^2+\cdots+\underline{u}_n^2)^l = \frac{(n-1)!\,l!}{(n+2l-1)!}\sum_{j_1+\ldots+j_n=l}\frac{(2j_1)!\cdots(2j_n)!}{j_1!\cdots j_n!}, \tag{10}$$

which is derived in MORAN [12] with some difficulty.
SHERMAN [16] introduces the test statistic

$$\underline{\omega}_{n-1} = \tfrac{1}{2}\sum_{j=1}^{n}\left|F(\underline{z}_{(j)}) - F(\underline{z}_{(j-1)}) - \frac{1}{n}\right|,$$

where $(\underline{z}_{(1)}, \cdots, \underline{z}_{(n)})$ is an ordered sample from a population with distribution function F, $F(\underline{z}_{(0)}) = 0$ and $F(\underline{z}_{(n+1)}) = 1$. Obviously one has

$$\underline{\omega}_{n-1} \cong \tfrac{1}{2}\sum_{j=1}^{n}\left|\underline{x}_j - \frac{1}{n}\right|,$$

the moments of which may be found as follows:

$$\frac{(n-1)!}{\tau^n}\mathscr{E}\exp\left\{-s\sum\left|\underline{u}_j(\tau)-\frac{1}{n}\right|\right\} = \frac{(n-1)!}{(\tau-s)^n}\left(e^{-\frac{s}{n}} - \frac{2s}{\tau+s}\,e^{-\frac{\tau}{n}}\right)^n =$$

$$= \frac{(n-1)!}{(\tau-s)^n}\sum_{k=0}^{n}\binom{n}{k}(-)^k 2^k s^k(\tau+s)^{-k} e^{-\frac{n-k}{n}s}\, e^{-\frac{k\tau}{n}} =$$

$$= (n-1)!\sum_{k=0}^{n}\binom{n}{k}$$

$$(-)^k 2^k e^{-\frac{k\tau}{n}}\sum_{j=0}^{\infty}(-)^j\left(\frac{n-k}{n}\right)^j\frac{1}{j!}\sum_{h=0}^{\infty}\binom{-k}{h}\sum_{r=0}^{\infty}(-)^r\binom{-n}{r}s^{k+j+h+r}\tau^{-(n+k+h+r)}.$$

Differentiating l times with respect to s and putting $s = 0$ we have

$$\frac{(n-1)!}{\tau^n}\mathscr{E}\left(\left|\sum \underline{u}_j(\tau)-\frac{1}{n}\right|\right)^l =$$

$$= (-)^l l!(n-1)!\overline{\sum_{k+j+h+r=l}}(-)^{k+j+r}\binom{n}{k}\binom{-k}{h}\binom{-n}{r}2^k\left(\frac{n-k}{n}\right)^j\frac{1}{j!}e^{-\frac{k}{n}}\tau^{-(n+k+h+r)}.$$

Inversion (see (1)) and taking $t = 1$ gives

$$\mathscr{E}\sum\left|\underline{x}_j-\frac{1}{n}\right|^l = l!(n-1)!\overline{\sum_{k+j+h+r=l}}(-)^h\binom{n}{k}\binom{-k}{h}$$

$$\binom{-n}{r}\frac{1}{j!}2^k\left(1-\frac{k}{n}\right)^{n+k+h+r+j-1}\frac{1}{(n+k+h+r-1)!} =$$

$$= \frac{1}{\binom{n+l-1}{n-1}}\sum_{k=0}^{n}\binom{n}{k}2^k\left(1-\frac{k}{n}\right)^{n+l-1}\overline{\sum_{j+h+r=l-k}}(-)^h\binom{-k}{h}\binom{-n}{r}\binom{n+l-1}{j}.$$

It is easily verified that

$$\overline{\sum_{j+h+r=l-k}}(-)^h\binom{-k}{h}\binom{-n}{r}\binom{n+l-1}{j} = 2^{l-k}\binom{l-1}{k-1}$$

and thus we have

$$\mathscr{E}\underline{\omega}_{n-1}^l = \frac{1}{\binom{n+l-1}{l}}\sum_{k=0}^{n}\binom{n}{k}\binom{l-1}{k-1}\left(1-\frac{k}{n}\right)^{n+l-1}\quad (l = 0,\ 1,\ 2,\cdots),$$

with the usual convention that $\binom{a}{b}$ vanishes for $b < 0$ and $b > a$. This result has been obtained in SHERMAN [16] by integration in n-dimensional space and a considerably greater amount of computation.

6. Asymptotic results

Making use of the fact that for $1 \geqslant a_n \geqslant n^{-\frac{1}{3}}$, say,

$$\frac{n^n}{(n-1)!}\int_{1-a_n}^{1+a_n} t^{n-1}e^{-nt}\,dt = 1 + e(n)\quad (n \to \infty), \tag{11}$$

where $e(n)$ denotes an expontially small contribution [3]), one can prove asymptotic properties for functions of the $\underline{x}_k$ using asymptotic properties of functions of the independent $\underline{u}_k$.

Defining

[3]) i.e. $|e(n)| \leqslant \exp(-n^p)$ for some $p > 0$.

$$\delta_n(t-1) = \frac{n^n}{(n-1)!}\, t^{n-1}\, e^{-nt},$$

we have

$$\left\{\begin{aligned} &\int_0^\infty \delta_n(t-1)\, dt = 1 \\ &\lim_{n\to\infty} \delta_n(t-1) = 0 \quad (t \neq 1). \end{aligned}\right. \tag{12}$$

Putting $\tau = n$ in (1) we obtain a formula of the form

$$\int_0^\infty E_n(t)\, \delta_n(t-1)\, dt = E_n{}^*(n). \tag{13}$$

From (12) and (13) one expects that under suitable conditions the relation

$$\lim_{n\to\infty} E_n(1) = \lim_{n\to\infty} E_n{}^*(n)$$

will hold. We prove the following lemma:

L e m m a 2:

If the $E_n(t)$ are uniformly bounded and monotonic in t and if

$\lim\limits_{n\to\infty} \int_0^\infty E_n(\alpha t)\, \delta_n(t-1)\, dt = E(\alpha)$ *exists and is continuous in α for all $\alpha > 0$ then* $\lim\limits_{n\to\infty} E_n(\alpha) = E(\alpha)$.

P r o o f: from the conditions it follows that $E(\alpha)$ is monotonic and bounded as well as continuous. This implies that the integral converges to $E(\alpha)$ uniformly in α.

For all $\delta \geqslant n^{-\frac{1}{4}}$ and all $\alpha > 0$ we therefore have (see (11))

$$\int_{1-\delta}^{1+\delta} E_n(\alpha t)\, \delta_n(t-1)\, dt = E(\alpha) + \varepsilon_n(\delta, \alpha),$$

where $\varepsilon_n(\delta, \alpha) \to 0$ uniformly in δ and α if $n \to \infty$. If $E_n(t)$ is non-decreasing (and similarly if $E_n(t)$ is non-increasing) it follows that

$$E_n(\alpha - \delta\alpha) \leqslant E(\alpha) + \varepsilon_n(\delta, \alpha)) \leqslant E_n(\alpha + \delta\alpha)$$

or equivalently

$$E\left(\frac{\alpha}{1+\delta}\right) + \varepsilon_n\left(\delta, \frac{\alpha}{1+\delta}\right) \leqslant E_n(\alpha) \leqslant E\left(\frac{\alpha}{1-\delta}\right) + \varepsilon_n\left(\delta, \frac{\alpha}{1-\delta}\right).$$

Now taking $\delta = n^{-\frac{1}{4}}$ and letting $n \to \infty$ by the continuity of $E(\alpha)$ we have $\lim_{n\to\infty} E_n(\alpha) = E(\alpha)$.

We use this lemma to prove the asymptotic normality of the statistic $\underline{q}_n = \sum \underline{x}_j^2$. This is also proved in MORAN [13] and DARLING [3].

We put

$$E_n\left(\frac{t}{z}\right) = P\{(\sum \underline{x}_j^2(t) - t^2\mu)\sigma^{-1} \leqslant z\} = P\left\{(\sum \underline{x}_j^2 - \mu)\sigma^{-1} \leqslant \frac{z}{t^2}\right\},$$

where μ en σ^2 denote the mean and variance of $\underline{q}_n$. Using $\underline{u}_j(n) \triangleq n^{-1}\underline{u}_j$ we have (c.f. (1))

$$\int_0^\infty E_n\left(\frac{t}{z}\right)\delta_n(t-1)\,dt = P\{[\sum \underline{u}_j^2 - \mu(\sum \underline{u}_j)^2]\,n^{-2}\sigma^{-1} \leqslant z\}. \qquad (14)$$

By theorem 2 the mean and variance of $\sum \underline{u}_j^2 - \mu(\sum \underline{u}_j)^2$ are 0 and

$$n(n+1)(n+2)(n+3)\,\sigma^2.$$

Therefore the right-hand side of (14) tends to the standard normal distribution function by a theorem in CRAMÉR [2] (page 366) concerning the asymptotic normality of continuous functions of the sample moments. From lemma 2 it now follows that $\underline{q}_n$ is asymptotically normal with (see (10))

$$\mu \sim 2n^{-1},\ \sigma^2 \sim 4n^{-3}.$$

Extension of this proof to $\sum \underline{x}_j^k \quad (k > 2)$ is immediate.

Finally we give short derivations of some limit theorems proved in FLATTO and KONHEIM [6] (see also [3]).

If

$$Q_{n,m}(a, b) = P\{a \leqslant \underline{x}_j \leqslant b \text{ for exactly } m \text{ values of } j\},$$

$$P_{n,m}(a, b) = \sum_{k=0}^{m} Q_{n,k}(a, b),$$

then putting $\tau = n$ in (1) we have

$$\int_0^\infty P_{n,m}\left(\frac{a}{tn^2}, \frac{b}{tn^2}\right)\delta_n(t-1)\,dt =$$

$$= \sum_{k=0}^{m}\binom{n}{k}\left\{e^{-\frac{a}{n}} - e^{-\frac{b}{n}}\right\}^k\left\{1 - e^{-\frac{a}{n}} + e^{-\frac{b}{n}}\right\}^{n-k} \sim \sum_{k=0}^{m}\frac{(b-a)^k}{k!}e^{-b+a}.$$

Taking α in lemma 2 as a vector variable this lemma can be applied with

$$E_n(\alpha t) = P_{n,m}\left(\frac{a}{tn^2}, \frac{b}{tn^2}\right),$$

which is monotonic and uniformly bounded in t.

It follows that

$$\lim_{n\to\infty} Q_{n,m}\left(\frac{a}{n^2}, \frac{b}{n^2}\right) = \frac{(b-a)^m}{m!} e^{-b+a}.$$

A slightly adapted version of lemma 2 can be used to prove in the same way

$$\lim_{n\to\infty} Q_{n,m}\left(\frac{a+\log n}{n}, \frac{b+\log n}{n}\right) = \frac{(e^{-a}-e^{-b})^m}{m!} \exp(e^{-a}-e^{-b}).$$

In the final section of [6] it is proved that the mean of the number $\underline{N}_\alpha$ of random arcs of length α needed to cover a circle of unit circumference has the asymptotic behaviour

$$\mathscr{E}\underline{N}_\alpha \sim \frac{1}{\alpha} \log \frac{1}{\alpha} \qquad (\alpha \to 0).$$

We find the stronger result

$$\mathscr{E}\underline{N}_\alpha = \frac{1}{\alpha}\left\{\log\frac{1}{\alpha} + \operatorname{loglog}\frac{1}{\alpha} + \gamma + o(1)\right\} \qquad (\alpha \to 0), \tag{15}$$

where γ is EULER's constant.

Clearly the probability $p_n(\alpha; t)$, that a circle of circumference t is covered by n or fewer arcs equals

$$p_n(\alpha; t) = P\{\underline{x}_{(n)}(t) \leqslant \alpha\},$$

while $\mathscr{E}\underline{N}_\alpha$ is given by

$$\mathscr{E}\underline{N}_\alpha = \sum_{n=1}^{\infty} \{1-p_n(\alpha)\},$$

where $p_n(\alpha) = p_n(\alpha; 1)$.

Putting $\tau = n$ in (1) by (13) we have

$$\int_{1-a_n}^{1+a_n} \{1-p_n(\alpha; t)\}\, \delta_n(t-1) = 1-(1-e^{-\alpha n})^n + e(n), \tag{16}$$

where $a_n = n^{-\frac{1}{3}}$, say.

As $p_n(\alpha; t) = p_n(\alpha t^{-1})$ is decreasing with respect to t, from (16) we deduce the inequality

$$\begin{aligned} &1-[1-\exp\{-n\alpha/(1-\alpha^{\frac{1}{3}})\}]^n + e(n) \leqslant 1-p_n(\alpha) \leqslant \\ &\leqslant 1-[1-\exp\{-n\alpha/(1+\alpha^{\frac{1}{3}})\}]^n + e(n) \qquad \left(n > \frac{1}{\alpha}, \alpha \to 0\right). \end{aligned}$$

Hence,

as $p_n(\alpha) = 0$ for $n \leqslant \frac{1}{\alpha}$ and $\sum_{n > \frac{1}{\alpha}} e(n) = o(1)$ $(\alpha \to 0)$ we have

$$\frac{1}{\alpha} + \sum_{n > \frac{1}{\alpha}} (1 - [1 - \exp\{-n\alpha/(1-\alpha^{\frac{1}{4}})\}]^n) + o(1) \leqslant \mathscr{E}\underline{N}_\alpha \leqslant$$

$$\leqslant \frac{1}{\alpha} + \sum_{n > \frac{1}{\alpha}} (1 - [1 - \exp\{-n\alpha/(1+\alpha^{\frac{1}{4}})\}]^n) + o(1) \qquad (\alpha \to 0). \qquad (17)$$

We will prove that

$$\sum_{n > \frac{1}{\alpha}} \{1 - (1 - e^{-\alpha n})^n\} = \frac{1}{\alpha}\left\{\log\frac{1}{\alpha} + \operatorname{loglog}\frac{1}{\alpha} + \gamma - 1 + o(1)\right\} \qquad (\alpha \to 0). \ (18)$$

From (17) and (18) it is not hard to see that (15) holds.

Proof of (18):

as $(1 - e^{-\alpha u})^u$ is increasing for $u > \frac{\log 2}{\alpha}$ we have (with $x = \alpha u$)

$$\sum_{n > \frac{1}{\alpha}} \{1 - (1 - e^{-\alpha n})^n\} = \int_{\frac{1}{\alpha}}^{\infty} \{1 - (1 - e^{-\alpha u})^u\}\, du + 0(1) =$$

$$= \beta \int_1^{\infty} \{1 - (1 - e^{-x})^{\beta x}\}\, dx + 0(1), \qquad (19)$$

where $\beta = \frac{1}{\alpha}$ and $0(1) \leqslant 2$. For $\log 2 \leqslant x \leqslant \log\beta$ one has

$(1 - e^{-x})^{\beta x} = 0\left(\frac{1}{\beta}\right)$ $\quad(\beta \to \infty)$, and therefore

$$\int_1^{\log\beta} \{1 - (1 - e^{-x})^{\beta x}\}\, dx = \log\beta - 1 + o(1) \qquad (\beta \to \infty). \qquad (20)$$

Further we have (with $z = x - \log\beta$, $v = e^{-z}$ and $y = v \log\beta$)

$$\int_{\log\beta}^{\infty} \{1 - (1 - e^{-x})^{\beta x}\}\, dx = \int_0^{\infty} \left\{1 - \left(1 - \frac{1}{\beta} e^{-z}\right)^{\beta z + \beta\log\beta}\right\} dz =$$

$$= \int_0^{\infty} \{1 - \exp[-(z + \log\beta) e^{-z}]\}\, dz + o(1) = \int_0^1 \frac{1 - v^v e^{-v\log\beta}}{v}\, dv + o(1) =$$

$$= \int_0^1 \frac{1 - e^{-v\log\beta}}{v}\, dv + o(1) = \int_0^{\log\beta} \frac{1 - e^{-y}}{y}\, dy + o(1). \qquad (21)$$

Now

$$\int_0^{\log\beta} \frac{1-e^{-y}}{y}\,dy = \int_0^1 \frac{1-e^{-y}}{y}\,dy - \int_1^{\log\beta} \frac{e^{-y}}{y}\,dy + \int_1^{\log\beta} \frac{dy}{y} =$$

$$= \log\log\beta + \gamma + o(1) \qquad (\beta \to \infty) \tag{22}$$

(see e.g. WHITTAKER and WATSON [18, p. 236]). From (19), (20), (21) and (22) the required (18) follows.

References

[1] BARTON, D. E. and F. N. DAVID, Some notes on ordered random intervals, J. Royal. Stat. Soc. (B) **18** (1956).

[2] CRAMÉR, H., Mathematical Methods of Statistics, Princeton University Press (1946).

[3] DARLING, D. A., On a class of problems related to the random division of an interval, Ann. Math. Stat. **24** (1953).

[4] DOMB, C., On the use of a random parameter in combinatorial problems, Proc. Physical Society (A) **65** (1952).

[5] DWASS, M., The distribution of linear combinations of random divisions of an interval, Trabajos de Estadistica XII-I and II (1961).

[6] FLATTO, L. and A. G. KONHEIM, The random division of an interval and the random covering of a circle, SIAM Review (4) (1962).

[7] GIRAULT, M., Loi de probabilité du plus grand intervalle dans une partition "au hasard", Revue de statistique appliquée, Vol X-2 (1962).

[8] KENDALL, M. G. and P. A. P. MORAN, Geometrical probability, Charles Griffin, London (1963)

[9] MAULDON, J. G., Random division of an interval, Proc. Cambridge Phil. Soc. **47** (1951),

[10] MAULDON, J. G., An inversion formula for a generalized transform, Mathematika. Vol. 4 (1957).

[11] MAULDON, J. G., A generalization of the Beta-distribution, Ann. Math. Stat. **30** (1959).

[12] MORAN, P. A. P., The random division of an interval I, J. Roy. Stat. Soc. suppl. 9 (1947).

[13] MORAN, P. A. P., The random division of an interval III, J. Roy. Stat. Soc. XV (1953).

[14] POLLACZEK, F., Problèmes stochastiques posés par le phénomène de formation d'une queue d'attente à un guichet et par des phénomènes apparentés, Gauthier - Villars, Paris (1957).

[15] RENYI, A., On the theory of order statistics, Acta Math. Acad. Sci. Hung. IV (1953).

[16] SHERMAN, B., A random variable relating to the spacing of sample values, Ann. Math. Stat. **21** 339–361 (1950).

[17] WHITWORTH, W. A., Exercises on Choice and Chance, Deighton Bell & Co., Cambridge (1897). (Republished by Hafner, New York (1959)).

[18] WHITTAKER E. T. and G. N. WATSON, Modern Analysis, Cambridge University Press (4th edition) (1958).