Are many functions on the positive integers one-to-one?

by

Marcel F. Neuts *

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 127

October 1967

* This paper contains an expanded version of a talk given by the author under the "Program of Visiting Lecturers in Statistics - 1967-68" under the sponsorship of the American Statistical Association, Biometric Society and the Institute of Mathematical Statistics.

Are many functions on the positive integers one-to one?

by

Marcel F. Neuts*

Purdue University

Introduction

A sinple "paradox" relating to the enumeration of the elements in a countable set may be described in the following way.

Every second a genie throws ten balls into an urn. The balls are numbered 1, 2,... and at every throw he adds the next ten numbers to the urn so that at the n-th throw the balls numbered $10n-9$, $10n-8$, $10n$ $(n \geq 1)$ are added. This goes on forever.

Another genie removes one ball from the urn after each addition, but he must guarantee that every ball will eventually be thrown out. If he can see the balls, there is of course no problem. He can remove the balls 1, 2, 3,... successively and for any natural number k, he knows when it enters the urn and when it is removed. It enters the urn at the $[\frac{k}{10}] + 1$st throw and is removed after the k-th throw.[1]

For every k, the length of time T_k spent in the urn by the ball k is given by

---

* This paper contains an expanded version of a talk given by the author under the "Program of Visiting Lecturers in Statistics - 1967-68" under the sponsorship of the American Statistical Association, Biometric Society and the Institute of Mathematical Statistics.

1 This and the paradox discussed here are related to the so-called "Tristam Shandy Paradox" - See Russell [6].

$$T_k = k - 1 - [\frac{k}{10}] , \qquad k = 1,2,\ldots$$

There are, of course, many more rules which will guarantee the eventual removal of every given ball. Clearly, there are also rules which will leave one or more, even infinitely many balls in the urn. Say if he removed successively the balls 10, 20, 30,... all numbers which are not multiples of ten would stay in forever.

To compound the sad fate of the second genie, we assume next, that he cannot see the numbers on the balls and that the balls are, in fact, completely indistinguishable. The problem is now, whether or not there is a way in which the second genie can remove every ball from the urn. Or, to state the "paradox": Does the ability of the second genie to enumerate all the balls depend on the enumeration already given?"

We must still describe a rule, but one that does not depend on the numbering of the balls at all. The first such procedure that comes to mind is to draw at each removal the ball <u>at random from among those still in the urn</u>. This rule is appealing, because every ball in the urn at every drawing is given the same chance of being removed. Before the n-th removal there are $9n + 1$ balls in the urn. We assume that, independent of the past, any one of these balls has a probability $(9n + 1)^{-1}$ of being taken out.

This rule will be satisfactory for genie II, if we can show that, with probability one, every given ball is eventually removed from the urn.

Since the balls are completely indistinguishable, the genie must rely on chance and a chance procedure with the stated property is the best one can wish for.

We will prove below that "random removal" has this property, but first we leave the world of fairy tales and formulate a more general mathematical problem.

Mathematical formulation

Let $a_1 < a_2 < \dots$ be a strictly increasing sequence of positive integers and let $\mathcal{F}$ be the family of all functions from the positive integers into the positive integers, which satisfy:

(1) $$f(n) \le a_n, \qquad n \ge 1,$$
$$f(n) \ne f(\nu), \qquad \nu \ne n \quad .$$

On the class of all subsets of $\mathcal{F}$, we can define probabilities satisfying:

(2) $$P\{f(1) = k\} = \frac{1}{a_1}, \qquad 1 \le k \le a_1$$
$$= 0, \qquad k > a_1$$

and for all $n > 1$:

(3) $$P\{f(n) = k | f(1), \dots, f(n-1)\} =$$
$$\frac{1}{a_n - n+1}, \qquad 1 \le k \le a_n, \quad k \ne f(\nu), \quad \nu = 1, \dots, n-1$$

and zero elsewhere.

This assignment of probabilities corresponds to the following scheme: For every $n \geq 1$, the value of $f(n)$ is chosen at random from among the numbers $1,2,\dots,a_n$ which have not been chosen previously. That the requirements (2) and (3) determine a unique probability measure on the class of all subsets of $\mathcal{F}$ may be proved from first principles or by appealing to the general theorem 8.3.A, p. 137 in Loève [3][2].

This assignment of probabilities corresponds to the requirement, which, loosely stated, says that all functions in $\mathcal{F}$ "equally probable". To see this we prove:

Property 1:

Let $(\alpha_1,\dots,\alpha_m)$ be any m-tuple of natural numbers, no two of which are equal, with $\alpha_i \leq a_i$ for $i = 1,\dots,m$ then:

(4) $$P\{f(1) = \alpha_1,\dots,f(m) = \alpha_m\} =$$

$$[a_1(a_2-1) \dots (a_i-i+1) \dots (a_m-m+1)]^{-1},$$

regardless of the m-tuple chosen and is equal to zero for all other m-tuples.

Proof: Use the chain rule of conditional probability, then:

$$P\{f(1) = \alpha_1,\dots,f(m) = \alpha_m\} =$$

$$P\{f(1) = \alpha\}\ P\{f(2) = \alpha_2|f(1) = \alpha_1\} \dots$$

$$\dots P\{f(m) = \alpha_m|f(1) = \alpha_1,\dots,f(m-1) = \alpha_{m-1}\},$$

which yields (4) upon substitution.

[3] The uniqueness of the probability measure P also follows from property 1 below and the classical extension theorem for measures.

Remarks

The space of functions $\mathcal{F}$ with the probability assignment $P(\cdot)$ may be identified with the following urn scheme. Suppose that the urn contains initially a_1 balls, numbered $1,\ldots,a_1$. One ball is drawn out and new balls, numbered $a_1+1,\ldots, a_2$ are added. Again a ball is drawn out at random and removed, and balls, numbered $a_2+1,\ldots, a_3$ are added and so on. If we denote by X_n the number of the n-th ball drawn, then the sequence $\{X_1, X_2, \ldots\}$ defines a function in $\mathcal{F}$. We see that the sequence $a_1, a_2,\ldots$ characterizes the set $\mathcal{F}$ and the probability assignment $P(\cdot)$. The scheme, discussed in the Introduction, corresponds to $a_n = 9n+1$.

Let the event that $X_n = k$ be denoted by $\{X_n = k\}$, then $\bigcup_{n=1}^{\infty} \{X_n = k\} = B_k$ is the event that for some n the number k is drawn at the n-th drawing. Since the events $\{X_n = k\}$ are disjoint, we have:

(5) $$P(B_k) = \sum_{n=1}^{\infty} P\{X_n = k\} .$$

We are interested in conditions on the sequence $\{a_n\}$ under which:

(6) $$\forall k:\ P(B_k) = 1.$$

Theorem 1

a. If $P(B_{k_o}) = 1$ for some $k_o \geq 1$, then (6) holds.

b. Property (6) holds if and only if

(7) $$\sum_{n=1}^{\infty} \frac{1}{a_n - n + 1} = \infty .$$

Proof:

Let k_o be a positive integer and $n^* = \min\{n: a_n \geq k_o\}$ then

(8) $$P(B_{k_o}^c) = P[\bigcap_{n=n^*}^{\infty} (X_n \neq k_o)] = \prod_{n=n^*}^{\infty} (1 - \frac{1}{a_n - n+1})$$

So that $P(B_{k_o}^c) = o$ if and only if the infinite product diverges, or equivalently if the sum (7) does.

However the divergence of this sum is independent of the value of k_o which proves part a .

Corollary

If (7) holds, then for any nonvoid set of indices $\{k_1, k_2, \ldots\}$ we have:

(9) $$P\{\bigcap_{i=1}^{\infty} B_{k_i}\} = 1$$

Proof:

$$P\{\bigcap_{i=1}^{\infty} B_{k_i}\} = 1 - P\{\bigcup_{i=1}^{\infty} B_{k_i}^c\}$$

but

$$o \leq P\{\bigcup_{i=1}^{\infty} B_{k_i}^c\} \leq \sum_{i=1}^{\infty} P(B_{k_i}^c) = o$$

by Theorem 1.

Remark

The corollary says that with probability one all positive integers appear in an infinite sequence of drawings in an urn corresponding to a sequence $\{a_n\}$ which satisfies (7). We can therefore say that if and only if condition (7) is satisfied "almost all functions in the class $\mathcal{F}$ are one-to-one".

An example of a class of functions, which do not satisfy condition (7)

It is, of course, easy to give examples of such classes of functions, just by choosing a_n a fast growing sequence. The following example is of some particular interest as it relates to the familiar proof of the countability of the set of all rational numbers.

Let E_n be the set of all rational numbers in $(0,1)$ which can be written as irreducible fractions with denominator at most equal to n. The number of elements in E_n is given by:

$$(10) \qquad a_n = \sum_{\nu=2}^{n} [\varphi(\nu) - 1] = \sum_{\nu=2}^{n} \varphi(\nu) - n + 1, \qquad n \geq 2 .$$

Set $a_1 = 1$. $\varphi(\nu)$ if Euler's φ-function.[3]

Therefore, for $n \geq 2$, we have:

$$(12) \qquad \frac{1}{a_n - n+1} = [\sum_{\nu=2}^{\infty} \varphi(\nu) - 2(n-1)]^{-1} .$$

However, it is known that:

$$(13) \qquad \lim_{n \to \infty} \frac{1}{n^2} \sum_{\nu=1}^{n} \varphi(\nu) = \frac{3}{\pi^2} .$$

[3]Euler's $\varphi(\cdot)$-function is defined as follows: $\varphi(\nu)$ is the number of integers a, with $1 \leq a \leq \nu$ which are relatively prime to ν.

See Erdelyi, Magnus, Oberhettinger, Tricomi [1]. Vol. 3, p. 172, formula (32).

Therefore:

(14) $$\frac{1}{a_n - n + 1} \sim \frac{1}{n^2} \cdot \frac{\pi^2}{3} ,$$

so that the series in (7) converges.

Remark

An interesting open problem is to find an expression for the probability that a function is one-to-one if condition (7) is not satisfied.

Functions of at most linear growth.

The class $\mathcal{F}$ of functions corresponding to

(16) $$a_n = a + b(n-1) \qquad a \geq 1,\ b \geq 1,\ n \geq 1 ,$$

is of particular interest.

Since $a_n - n + 1 = a + (b-1)(n-1)$, the series in (7) diverges. Consider any ball in the urn just before the n-th drawing and let T be the additional number of drawings required before this ball is removed, then:

(17) $$P\{T > \nu\} = \prod_{\alpha=n}^{n+\nu} [1 - \frac{1}{a + b(\alpha-1)}], \qquad \nu \geq 0 ,$$

$$= \frac{\Gamma(\frac{a-b-1}{b} + n + \nu + 1)}{\Gamma(\frac{a-b}{b} + n + \nu + 1)} \cdot \frac{\Gamma(\frac{a-b}{b} + n)}{\Gamma(\frac{a-b-1}{b} + n)}$$

$$= \frac{B[\frac{a-b-1}{b} + n + \nu + 1, \frac{1}{b}]}{B[\frac{a-b-1}{b} + n, \frac{1}{b}]}$$

in terms of Euler's gamma and beta functions. The expected value of the random variable T is given by:

$$(18) \qquad E(T) = \sum_{\nu=0}^{\infty} P\{T > \nu\} =$$

$$\frac{1}{B[\frac{a-b-1}{b} + n, \frac{1}{b}]} \sum_{\nu=0}^{\infty} \int_0^1 u^{\frac{a-b-1}{b} + n + \nu} (1-u)^{\frac{1}{b} - 1} du$$

$$= \frac{1}{B[\frac{a-b-1}{b} + n, \frac{1}{b}]} \int_0^1 u^{\frac{a-b-1}{b}} (1-u)^{\frac{1}{b} - 2} du$$

$$= +\infty ,$$

since the integral on the right diverges.

This leads to the observation, that though the ball in the urn at time n will be drawn out eventually with probability one, the expected number of drawings required is infinite.

To illustrate the enormous growth of waitingtimes in terms of n, we consider an extremely simple case of (16) and appeal to some results which were proved in the theory of record observations.

Let $a = 2$ and $b = 1$, so that the number of balls in the urn at the n-th drawing is $n + 1$. $(n \geq 1)$. We note that this is the slowest possible growing sequence a_n.

Consider the following process. Before the first drawing, mark one of the two balls and continue drawing until the marked ball is drawn. When this happens, mark one of the balls in the urn just before the next drawing and continue drawing until this ball is drawn. When this happens, mark again one of the balls in the urn and so on.

It is easy to see that by this procedure, we generate a sequence of independent Bernoulli trials in which the probability of success at the n-th trial is $\frac{1}{n+1}$. Success is defined as the drawing of a previously marked ball.

Suppose now that we define the random variable L_m as the total number of drawings required until the m-th marked ball is drawn out. Equivalently L_m is the number of trials until the m-th success in a sequence of independent Bernoulli trials in which the probability of success at the n-th trial is $p_n = \frac{1}{n+1}$.

The random variable L_m was studied by Foster and Stuart [4] and by Alfred Rényi [6] in connection with the study of recordbreaking observations. They proved among other things that:

$$(19) \qquad (L_m)^{1/m} \to e$$

with probability one and that:

$$(20) \qquad P\{\log L_m \le m + t\sqrt{m}\} \to \int_{-\infty}^{t} e^{-u^2/2} \frac{du}{\sqrt{2\pi}} ,$$

so that the limiting distribution of $\frac{\log L_m - m}{\sqrt{m}}$ is a unit normal distribution.

However if we set $\Delta_m = L_m - L_{m-1}$, $m \ge 1$, $L_o = o$, then Neuts [5] has shown that:

(21) $(\Delta_m)^{1/m} \rightarrow e$ in probability

and

(22) $$P\{\log \Delta_m \leq m + t\sqrt{m}\} \rightarrow \int_{-\infty}^{t} e^{-u^2/2} \frac{du}{\sqrt{2\pi}} ,$$

so that the limiting behavior of $L_m = \Delta_1 + \Delta_2 + \ldots + \Delta_m$ is practically the same as that of the last term Δ_m. This shows that for large m, the waitingtime between the last two successes completely overshadows even the sum of all the previous waitingtimes.

M. N. Tata [8] has investigated the sequence L_m, $m = 1,2,\ldots$ further and has shown, in particular, that the limiting distribution of $\frac{L_{m+1}}{L_m}$ exists for $m \rightarrow \infty$, but even it has an infinite expected value. This shows that the penalty paid for making the balls indistinguishable is in the waitingtimes involved.

To end this discussion in the world of fairy tales, where it started, we may say that the genie II will exhibit the k-th ball, less than k drawings after it was placed in the urn, provided he knows the numbering on the balls. If he has to go by chance, he can still be certain to draw out any given ball eventually, but the number of drawings involved in each case will be large with considerable probability. Since the genies were doomed to this activity for an infinite length of time, anyway, it probably does not matter to them whether they are guided by knowledge or by chance.

Acknowledgement

The 'paradox' of the genies was mentioned several years ago to the author by Professor Samuel Kaplan. He should certainly be thanked for this stimulating conversation piece.

Constructive comments for Professor Louis J. Cote are also gratefully acknowledged.

References

[1] Erdelyi, Magnus, Oberhettinger, Tricomi (1953). Higher Transcendental Functions, Vol. 3, McGraw-Hill, New York.

[2] Feller, W. (1947). "An Introduction to Probability and its Applications, Vol. I, 2nd Ed." J. Wiley, New York.

[3] Loève, M. (1955). Probability Theory, 3rd Ed., Van Nostrand, Princeton, New Jersey.

[4] Foster, F.G. and Stuart, A. (1954). Distribution-free tests in time-series based on the breaking of records. J. Roy. Stat. Soc. Ser. B, **16**, pp. 1-22.

[5] Neuts, M.F. (1967). Waitingtimes between record observations. J. Appl. Prob. **4**, 206-208.

[6] Rényi, Alfred (1962). Théorie des éléments saillants d'une suite d'observations. Coll. Combin. Meth in Prob. Th., Math. Inst. Aarhus University, Denmark.

[7] Russell, Bertrand (1903). The Principles of Mathematics. Cambridge University Press.

[8] Tata, M.N. (1967). On Outstanding Values in a Sequence of Random Variables. Ph.D. Thesis, Purdue University, Department of Statistics (unpublished).