# ARE MANY 1-1-FUNCTIONS ON THE POSITIVE INTEGERS ONTO?

MARCEL F. NEUTS, Purdue University

**Introduction.** A simple "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 $1, 2, \cdots$ and at every throw he adds the next ten numbers to the urn so that at the $n$th throw the balls numbered $10n-9$, $10n-8$, $10n (n \geqq 1)$ 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 $1, 2, 3, \cdots$ 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 $[k/10]+1$st throw and is removed after the $k$th throw. Though the number of balls in the urn tends to infinity, any given ball is eventually thrown out. No ball stays in the urn forever. This is one of the paradoxes of infinity, stated by Georg Cantor and discussed as the "Tristam Shandy paradox" by Russell [7].

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

$$T_k = k - 1 - \left[\frac{k}{10}\right], \qquad k = 1, 2, \cdots.$$

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. Thus if he removed successively the balls $10, 20, 30, \cdots$ 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 *at random from among those still in the urn.* 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 $a_1 < a_2 < \cdots$ be a strictly increasing sequence of positive integers and let $\mathfrak{F}$ be the family of all functions from the positive integers into the positive integers which satisfy

$$(1) \qquad \begin{aligned} f(n) &\leqq a_n, & n &\geqq 1, \\ f(n) &\neq f(\nu), & \nu &\neq n. \end{aligned}$$

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

$$(2) \qquad \begin{aligned} P\{f(1) = k\} &= \frac{1}{a_1}, & 1 &\leqq k \leqq a_1 \\ &= 0, & k &> a_1 \end{aligned}$$

and, for all $n > 1$,

$$(3) \qquad P\{f(n) = k \mid f(1), \cdots, f(n-1)\} = \frac{1}{a_n - n + 1}$$
$$1 \leqq k \leqq a_n,\ k \neq f(\nu) \quad \nu = 1, \cdots, n-1,$$

and zero elsewhere.

This assignment of probabilities corresponds to the following scheme: for every $n \geqq 1$, the value of $f(n)$ is chosen at random from among the numbers $1, 2, \cdots, a_n$ which have not been chosen previously. That the requirements (2) and (3) determine a unique probability measure on the class of all subsets of $\mathfrak{F}$ may be proved from first principles or by appealing to the general theorem 8.3.A, p. 137 in Loève [**3**]. The uniqueness of the probability measure $P$ also follows from property 1 below and the classical extension theorem for measures.

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

PROPERTY 1.

$$(4) \qquad P\{f(1) = \alpha_1, \cdots, f(m) = \alpha_m\} = [a_1(a_2 - 1) \cdots (a_m - m + 1)]^{-1},$$

is $\alpha_i \leqq a_i$, for $i = 1, \cdots, m$ and no two $\alpha_i$'s are equal. For all other $m$-tuples $(\alpha_1, \cdots, \alpha_m)$, this probability is zero.

*Proof.* Use the chain rule of conditional probability; then

$$\begin{aligned} &P\{f(1) = \alpha_1, \cdots, f(m) = \alpha_m\} = \\ &P\{f(1) = \alpha_1\} P\{f(2) = \alpha_2 \mid f(1) = \alpha_1\} \cdots \\ &\cdots P\{f(m) = \alpha_m \mid f(1) = \alpha_1, \cdots, f(m-1) = \alpha_{m-1}\}, \end{aligned}$$

which yields (4) upon substitution.

*Remarks.* The space of functions $\mathfrak{F}$ with the probability assignment $P(\cdot)$ may be identified with the following urn scheme. Suppose that the urn contains initially $a_1$ balls, numbered $1, \cdots, a_1$. One ball is drawn out and new balls, numbered $a_1+1, \cdots, a_2$ are added. Again a ball is drawn out at random and removed, and balls, numbered $a_2+1, \cdots, a_3$ are added and so on. If we denote by $X_n$ the number of the $n$th ball drawn, then the sequence $\{X_1, X_2, \cdots\}$ defines a function in $\mathfrak{F}$. We see that the sequence $a_1, a_2, \cdots$ characterizes the set $\mathfrak{F}$ and the probability assignment $P(\cdot)$. The scheme, discussed in the Introduction, corresponds to $a_n = 10n$.

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

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

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

$$\forall k\colon P(B_k) = 1. \tag{6}$$

THEOREM 1. (a) *If* $P(B_{k_0}) = 1$ *for some* $k_0 \geqq 1$, *then* (6) *holds.* (b) *Property* (6) *holds if and only if*

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

*Proof.* Let $k_0$ be a positive integer and $n^* = \min\{n\colon a_n \geqq k_0\}$; then

$$P(B_{k_0}^c) = P\left[\bigcap_{n=n^*}^{\infty} (X_n \neq k_0)\right] = \prod_{n=n^*}^{\infty}\left(1 - \frac{1}{a_n - n + 1}\right), \tag{8}$$

so that $P(B_{k_0}^c) = 0$ 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 $k_0$, which proves part (a).

COROLLARY. *If* (7) *holds, then for any nonvoid set of indices* $\{k_1, k_2, \cdots\}$ *we have*

$$P\left\{\bigcap_{i=1}^{\infty} B_{k_i}\right\} = 1. \tag{9}$$

*Proof.*

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

but

$$0 \leqq P\left\{\bigcup_{i=1}^{\infty} B_{k_i}^c\right\} \leqq \sum_{i=1}^{\infty} P(B_{k_i}^c) = 0$$

by Theorem 1.

*Remark.* The corollary says that, with probability 1, 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 $\mathfrak{F}$ are onto."

**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 a 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:

$$a_n = \sum_{\nu=2}^{n} \phi(\nu) \quad n \geqq 2. \tag{10}$$

Set $a_1 = 1$. $\phi(\nu)$ is Euler's $\phi$-function, i.e., $\phi(\nu)$ is the number of integers $a$, with $1 \leqq a \leqq \nu$ which are relatively prime to $\nu$.

Therefore, for $n \geqq 2$, we have

$$\frac{1}{a_n - n + 1} = \left[ \sum_{\nu=2}^{\infty} \phi(\nu) - (n-1) \right]^{-1}. \tag{11}$$

However, it is known that

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{\nu=1}^{n} \phi(\nu) = \frac{3}{\pi^2} \cdot \tag{12}$$

[See **1**, Vol. 3, p. 172, formula (32).]

Therefore

$$\frac{1}{a_n - n + 1} \sim \frac{1}{n^2} \cdot \frac{\pi^2}{3}, \tag{13}$$

so that the series in (7) *converges.*

*Remark.* An interesting problem is to find an expression for the probability that a function in $\mathfrak{F}$ is onto if condition (7) is not satisfied.

**Functions of at most linear growth.** The class $\mathfrak{F}$ of functions corresponding to

$$a_n = a + b(n-1) \qquad a \geqq 1, b \geqq 1, n \geqq 1, \tag{14}$$

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:

$$P\{T > \nu\} = \prod_{\alpha=n}^{n+\nu} \left[ 1 - \frac{1}{a + (b-1)(\alpha-1)} \right]$$

$$= \prod_{\alpha=n}^{n+\nu}\left(\frac{a-b}{b-1}+\alpha\right)\left(\frac{a-b+1}{b-1}+\alpha\right)^{-1}$$

(15)
$$= \frac{\Gamma\left(\frac{a-b}{b-1}+n+\nu+1\right)\Gamma\left(\frac{a-b+1}{b-1}+n\right)}{\Gamma\left(\frac{a-b}{b-1}+n\right)\Gamma\left(\frac{a-b+1}{b-1}+n+\nu+1\right)}$$

$$= \frac{B\left[\frac{a-b}{b-1}+n+\nu+1,\ \frac{1}{b-1}\right]}{B\left[\frac{a-b}{b-1}+n,\ \frac{1}{b-1}\right]}, \qquad b>1.\ \nu\geqq 0$$

in terms of Euler's gamma and beta functions [1]. The case $b=1$ is trivial and leads to a geometric distribution for $T$. The expected value of the random variable $T$ is given by ($b>1$)

$$E(T) = \sum_{\nu=0}^{\infty} P[T>\nu] = \frac{1}{B\left[\frac{a-b}{b-1}+n,\ \frac{1}{b-1}\right]}$$

(16)
$$\cdot \sum_{\nu=0}^{\infty} \int_0^1 u^{(a-b/b-1)+n+\nu}(1-u)^{(1/b-1)-1}du$$

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

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 1, the *expected number of drawings required is infinite.*

To illustrate the enormous growth of waiting times in terms of $n$, we consider an extremely simple case of (14) and appeal to some results which were proved in the theory of record observations.

Let $a=2$ and $b=2$ so that the number of balls in the urn at the $n$th drawing is $n+1$ ($n\geqq 1$).

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 $1/(n+1)$. 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 $p_n = 1/(n+1)$.

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

$$(L_m)^{1/m} \to e \tag{17}$$

with probability 1 and that

$$P\{\log L_m \leqq m + t\sqrt{m}\} \to \int_{-\infty}^{t} e^{-u^2/2} \frac{du}{\sqrt{2\pi}}, \tag{18}$$

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

However if we set $\Delta_m = L_m - L_{m-1}$, $m \geqq 1$, $L_0 = 0$, then Neuts [5] has shown that

$$(\Delta_m)^{1/m} \to e \text{ in probability} \tag{19}$$

and

$$P\{\log \Delta_m \leqq m + t\sqrt{m}\} \to \int_{-\infty}^{t} e^{-u^2/2} \frac{du}{\sqrt{2\pi}}, \tag{20}$$

so that the limiting behavior of $L_m = \Delta_1 + \Delta_2 + \cdots + \Delta_m$ is practically the same as that of the last term $\Delta_m$. 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 [7] has investigated the sequence $L_m$, $m = 1, 2, \cdots$ further and has shown, in particular, that the limiting distribution of $(L_{m+1})/(L_m)$ exists for $m \to \infty$, 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 by Professor Louis J. Cote are also gratefully acknowledged.

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.

**References**

1. A. Erdelyi, Higher Transcendental Functions, vol. 3, McGraw-Hill, New York, 1953.
2. W. Feller, An Introduction to Probability and its Applications, vol. I, 2nd ed., Wiley, New York, 1947.
3. M. Loève, Probability Theory, 3rd ed., Van Nostrand, Princeton, 1955.
4. F. G. Foster and A. Stuart, Distribution-free tests in time-series based on the breaking of records, J. Roy. Statist. Soc., Ser. B., 16 (1954) 1–22.
5. M. F. Neuts, Waitingtimes between record observations, J. Appl. Prob., 4 (1967) 206–208.
6. Alfred Rényi, Théorie des éléments saillants d'une suite d'observations, Coll. Combin. Meth. in Prob. Th., Math. Inst., Aarhus University, Denmark, (1962).
7. M. N. Tata, Contributions to the Theory of Record Observations, Ph.D. Thesis, Purdue University, unpublished, 1967.

---

# WEAK SUFFICIENT CONDITIONS FOR FATOU'S LEMMA AND LEBESGUE'S DOMINATED CONVERGENCE THEOREM

H. R. VAN DER VAART, North Carolina State University and ELIZABETH H. YEN, Columbia University

**0. Introduction.** In many expositions the Lebesgue-Stieltjes integral, $\int f(x)\mu(dx)=\int f d\mu$, or briefly $\int f$, of a measurable function $f$ is defined as the limit of a sequence of integrals $\int s_n d\mu$, where the $s_n$ are simple functions which in some sense tend to $f$ as $n\to\infty$. So, when we are interested in the limit of a sequence $\int f_n d\mu$ where all $f_n$ are measurable (rather than simple) functions, we have to deal with a double limit process. The *monotone convergence theorem* (MCT), *Fatou's Lemma*, and Lebesgue's *Dominated Convergence Theorem* (DCT) belong in this category. In the literature these results are discussed under a variety of mostly too restrictive conditions (cf. Section 2 below), which we have found tend to obscure their true nature in the mind of many students. The aim of this note is to present Fatou's Lemma as a special case of the MCT, and the DCT as a special case of Fatou's Lemma, being as general as possible as to conditions of boundedness and finiteness and also to indicate a method by which to construct the dominating function in the DCT. Of these objectives the last one seems to have some novelty. However, our main concern is pedagogical.

**1. Notations and terminology.** All functions discussed are assumed to be defined on a totally $\sigma$-finite measure space $(X, \mathfrak{A}, \mu)$ into the extended real number system $R^*$. (For the properties of $R^*$ see for instance [4], p. 2). All functions discussed will be *measurable* (i.e., if $B$ is $\{+\infty\}$, $\{-\infty\}$, or a Borel subset of the real line $R$, then $f^{-1}(B)\in\mathfrak{A}$). Given $\phi$: $X\to R^*$, the symbols $\phi^+$ and $\phi^-$ have the usual meaning: $\phi^+=\frac{1}{2}(\phi+|\phi|)$, $\phi^-=\frac{1}{2}(-\phi+|\phi|)$, so that $\phi=\phi^+-\phi^-$. Integration is always over some set $A$ belonging to the $\sigma$-algebra $\mathfrak{A}$. For our purposes the choice of $A$ is irrelevant (all properties stated concerning integrands are to hold on $A$), and we shall omit all reference to it. Whenever we write $\int\phi$, or $\int\phi d\mu$, or $\int\phi(x)\mu(dx)$, we imply that the *integral* is *defined*, either as a finite number, or as $+\infty$, or as $-\infty$, and we shall call such a $\phi$ integrable. In fact, we shall say