43

PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 14, Fasc. 2 (1993), pp. 157–172

ON STEINHAUS' RESOLUTION OF THE ST. PETERSBURG PARADOX

BY

SÁNDOR CSÖRGŐ (ANN ARBOR, MICHIGAN)
AND
GORDON SIMONS (CHAPEL HILL, NORTH CAROLINA)

Abstract. We elaborate on an interesting idea of Hugo Steinhaus from 1949 concerning the St. Petersburg paradox.

1. Introduction. Peter tosses a fair coin until it first lands heads and pays Paul 2^k ducats if this happens on the k^{th} toss, $k = 1, 2, \ldots$ What is the fair price for Paul to pay to Peter for the game? It is an infinite number of ducats but, as Nicolaus Bernoulli wrote, "... there ought not to exist any even halfway sensible man who would not sell his chance for forty ducats." This is the *St. Petersburg paradox*(1). If X denotes Paul's gain in the game, then

$$\mathrm{P}\{X = 2^k\} = 1/2^k, \quad k \in N := \{1, 2, \ldots\}, \tag{1.1}$$

and the last 280 years have produced an amazingly large number of ideas for replacing $\mathrm{E}(X) = \infty$ by something more reasonable as the price of a single game. Among a number of recent reviews of the rich history of the problem we mention Jorland [8] and Dutka [4]. Further historical details will be discussed in the introduction and a partially annotated bibliography of our forthcoming (mathematical) monograph [3].

(1) A variant of the problem, the same reward system for obtaining the first *six* on a fair die, was first proposed by the same Nicolaus Bernoulli (1687–1759), nephew of both the famous brothers Jacob (1654–1705) and Johannes (1667–1748) Bernoulli, in a letter of September 9, 1713, to de Montmort, and was published on pp. 401–402 in [9]. The above formulation, with the names of the people involved but with Paul receiving 2^{k-1} ducats with probability 2^{-k} instead of 2^k ducats, $k = 1, 2, \ldots$, was born in a correspondence between Nicolaus Bernoulli and Gabriel Cramer (1704–1752) in the 1720's. This latter form was used by Daniel Bernoulli (1700–1782), son of Johannes, whose discussion in [1] made the problem widely known. It was christened the *St. Petersburg problem* by d'Alembert (1717–1783) in 1768. Following W. Feller (1906–1970), we have doubled Paul's gain and have taken the associated liberty of doubling all the corresponding figures in *all* the citations here and below; for instance, Nicolaus originally said "twenty" and not "forty".

Feller [5] considers a sequence of independent St. Petersburg games, in which the random variables $X_1, X_2, \ldots$ denote Paul's gains in the first, second, ... plays, each distributed as X as shown in (1.1), and proves that

$$S_n/(n \operatorname{Log} n) \xrightarrow{\mathrm{P}} 1 \quad \text{as } n \to \infty, \tag{1.2}$$

where $S_n := X_1 + \ldots + X_n$ denotes Paul's total gain in n games, $n \in N$, Log denotes logarithm to the base 2, here and in what follows, and $\xrightarrow{\mathrm{P}}$ denotes convergence in probability. The price of $n \operatorname{Log} n$ ducats for n games, or $\operatorname{Log} n$ ducats per game if n games are played, is "fair" in the sense of the law of large numbers in (1.2). The result found its way into Feller's classical book [6] and its later editions, where he declares that "... the modern student will hardly understand the mysterious discussions of this 'paradox'."

Of course, Feller [5], [6] himself was well aware of "unfair 'fair' games", and there is more to the story. A complete asymptotic *distributional* theory, non-standard in nature, will be presented in [3]. Having compiled most of a partially annotated historical bibliography for this monograph, presently containing more than 340 entries, one of us came across some old *Mathematical Reviews* material on a compact disk and, as a final check, ran a search. This turned up Feller's review [7], from 1951, of a small note of Steinhaus [10]. Help by friends Jan Mielniczuk and Jacek Koronacki then produced a copy of the note directly from Poland, and it was easy to realize that chance (or providence) has presented us with a small, but true pearl. Steinhaus' note is difficult to find, it is not included in his *Selected Papers* published in 1985 in Warsaw. The later editions of Feller [6] do not refer to it even though he reviewed the note himself. The aim of the present paper is to pay tribute to the memory of a prominent mathematician, one of the leading figures of the Lwów school of mathematics and the founder of the Wrocław school of probability, by elaborating on an interesting idea of his, the content of his short note.

The heart of the St. Petersburg paradox is the non-existence of a fixed number, such as a finite expectation, which represents the fair "entrance fee" for a *single* St. Petersburg game. Steinhaus [10], clearly motivated by Feller [5], attempts to "resolve" this paradox, for a sequence of games, by showing that it is possible to "fairly" approximate the sequence of Paul's potential winnings $X_1, X_2, \ldots$ by a suitable sequence of deterministic entrance fees $x_1, x_2, \ldots$ The latter is what we call the *Steinhaus sequence* in the next section. The idea is distinctly different from Feller's. The "Steinhaus sums" $s(n) = x_1 + \ldots + x_n$ are then the deterministic analogues of Paul's winnings $S_n = X_1 + \ldots + X_n$, $n \in N$, in a sequence of independent St. Petersburg games. These sums are investigated in Section 3. The resulting "Steinhaus games" are compared to the original St. Petersburg games in Section 4. Some exact calculations are presented in the last section, where we also delineate the place of the Steinhaus resolution in our asymptotic theory.

2. The Steinhaus sequence. After describing the game, Steinhaus [10] notes that the classical rule, which determines the fair entrance fee as $E(X)$, becomes illusory for this game since $E(X) = \infty$. Then, referring to Feller [5], he continues as follows (2):

> Feller has analyzed this paradox by going back to the principle which justifies the rule in ordinary games, where $E(X)$ is finite. The "fair" fee has the property of balancing gains and losses if the game is repeated indefinitely. To speak more exactly, the net gain of a partner after n games has to be small in comparison with n, and the probability of its being so has to approach 1 as n increases indefinitely. This principle can be satisfied by a constant fee, equal to $E(X)$, in most popular games. If, however, $E(X)$ is infinite, as in the Petersburg game, it can be satisfied only by a variable fee c_n for the n^{th} repetition of the game. The determination of c_n according to the principle quoted is an application of a sort of *weak law of large numbers.*
>
> Thus, to solve the paradox, we must consider not an individual game but a sequence of games. This point of view once being adopted, we can find another solution, based on a sort of *strong law of large numbers.*

While the c_n here refers to Feller's average price $c_n = \operatorname{Log} n$ per game, Steinhaus generates his deterministic sequence $x_1, x_2, \ldots$ of entrance fees as follows: Begin with an alternating sequence of twos (3) and blank spaces,

2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2 _ 2...

Fill in every second blank space with a four,

2 4 2 _ 2 4 2 _ 2 4 2 _ 2 4 2 _ 2 4 2 _ 2 4 2 _ 2 4 2 _ 2 4 2 _ 2 4 2...,

then every remaining second blank with an eight,

2 4 2 8 2 4 2 _ 2 4 2 8 2 4 2 _ 2 4 2 8 2 4 2 _ 2 4 2 8 2 4 2 _ 2 4 2...,

and so on. As Steinhaus notes, the resulting sequence

2 4 2 8 2 4 2 16 2 4 2 8 2 4 2 32 2 4 2 8 2 4 2 16 2 4 2 8 2 4 2 64 2 4 2...

has 2's with frequency 1/2, has 4's with frequency 1/4, and in general has 2^k's with frequency 2^{-k} for every $k \in N$.

More precisely, we note that

$$\#\{1 \leqslant j \leqslant n \colon x_j = 2^k\} = \lfloor n/2^k + \tfrac{1}{2} \rfloor, \quad k \in N, \tag{2.1}$$

(2) Here and in other citations in the sequel, we have also taken the liberty of smoothing the English a bit; the original uses the phrases *weak* and *strong law of great numbers*, for example. Also, we adapted his text to the classical language that we use: Steinhaus (1887–1972) has B = Banker for our Peter and A for Paul the gambler, has a_n for c_n, and he uses pennies instead of ducats.

(3) The numbers in Steinhaus' sequence are also doubled here to conform to Feller's practice of doubling the payoffs in the classical formulation of the St. Petersburg game.

since the value $x_j = 2^k$ occurs when $j = 2^{k-1} + l2^k$, $l = 0, 1, \ldots$, and hence the number of j's in question is the number of l's for which

$$\tfrac{1}{2} + l \leqslant n/2^k$$

or, what is the same,

$$l + 1 \leqslant n/2^k + \tfrac{1}{2}, \qquad l = 0, 1, \ldots$$

Here and in what follows,

$$\lfloor y \rfloor = \max\{m \in \boldsymbol{Z}:\ m \leqslant y\}, \qquad \lceil y \rceil = \min\{m \in \boldsymbol{Z}:\ m \geqslant y\},$$
$$\langle y \rangle = y - \lfloor y \rfloor, \qquad y \in \boldsymbol{R},$$

where $\boldsymbol{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Thus, for the constants

$$\theta_{n,k} = 1 - 2\langle (n + 2^{k-1})/2^k \rangle, \qquad -1 < \theta_{n,k} \leqslant 1,$$

the proportion

$$\text{(2.2)} \qquad \frac{\#\{1 \leqslant j \leqslant n:\ x_j = 2^k\}}{n} = \frac{1}{n}\left\lfloor \frac{n}{2^k} + \frac{1}{2} \right\rfloor = \frac{1}{2^k} + \frac{\theta_{n,k}}{2n}, \qquad n \in \boldsymbol{N},$$

which clearly converges to 2^{-k} as $n \to \infty$, for every $k \in \boldsymbol{N}$.

Feller's weak law in (1.2) does not hold strongly or with probability 1. This was later noticed by Chow and Robbins [2]. Hence Steinhaus' mention of a strong law might appear strange at a first reading. While he might have anticipated Chow and Robbins' observation, for which there is no evidence, it was not the strong form of Feller's law that he had in mind. Rather, it was an analogue of the Glivenko–Cantelli strong law for sample distribution functions. Indeed, setting

$$\text{(2.3)} \qquad \tilde{F}_n(x) = \frac{1}{n}\,\#\{1 \leqslant j \leqslant n:\ x_j \leqslant x\}, \qquad x \in \boldsymbol{R},$$

for the n^{th} "empirical distribution function" of Steinhaus' sequence, the equation in (2.2) implies that, for $x \geqslant 2$,

$$\lim_{n\to\infty} \tilde{F}_n(x) = \lim_{n\to\infty} \sum_{k=1}^{\lfloor \operatorname{Log} x \rfloor} \frac{1}{n}\left\lfloor \frac{n}{2^k} + \frac{1}{2} \right\rfloor = \sum_{k=1}^{\lfloor \operatorname{Log} x \rfloor} \frac{1}{2^k} = \frac{1}{2}\,\frac{1 - (\frac{1}{2})^{\lfloor \operatorname{Log} x \rfloor}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^{\lfloor \operatorname{Log} x \rfloor}}.$$

In other words, setting

$$\text{(2.4)} \qquad F(x) = \mathrm{P}\{X \leqslant x\} = \begin{cases} 0, & x < 2, \\ 1 - 2^{-\lfloor \operatorname{Log} x \rfloor}, & x \geqslant 2, \end{cases}$$

for the distribution function for a single game, the formula itself coming easily from (1.1), it actually follows that

$$\sup_{x \in \boldsymbol{R}} |\tilde{F}_n(x) - F(x)| \to 0 \qquad \text{as } n \to \infty.$$

This is what Steinhaus informally describes. His interpretation (somewhat adjusted) is:

> ... fixing x_n as the entrance fee for the n^{th} repetition of the Petersburg game, we can predict with probability 1 that the amounts X_n paid by Peter to Paul will yield a sequence with (elements having) the same distribution function as the (limiting empirical distribution function of the) sequence $\{x_j\}_{j=1}^{\infty}$ of fees paid in advance by Paul to Peter. Such an equality justifies calling the game *fair* in a new sense of the word.

This is the content of Hugo Steinhaus' small note, consisting of sixty-six short lines, which he concludes by the following paragraph:

> When asked to estimate the constant fee he would like to pay for the Petersburg game, the average man names in most cases an amount less than 20 ducats. The reason is his taking into account only 20 terms of the series $\mathrm{E}(X) = \sum_{k=1}^{\infty} 2^k (1/2^k) = 1+1+\ldots = \infty$ at most, as the probability of the game extending beyond the 20^{th} trial is less than 1/1000000; his disbelief in such extraordinary occurrences is scarcely influenced by the rich reward promised by Peter if such a case would really happen. The same man would probably not hesitate to repeat the game indefinitely, his fees being determined by the sequence $\{x_j\}_{j=1}^{\infty}$, because he would realize that he pays large fees very rarely, as rarely as he wins, in the long run, amounts that are equal to such fees.

The aim of the present note is to elaborate further on the Steinhaus resolution. Among other things, we shall show that

$$1/2n \leqslant \sup_{x \in \boldsymbol{R}} |\tilde{F}_n(x) - F(x)| < 1/n, \qquad n \in \boldsymbol{N}. \tag{2.5}$$

First, however, the focus of attention will be on Steinhaus sums, the properties of which are even more interesting.

3. The Steinhaus sums. In view of Feller's law of large numbers in (1.2), the first question that arises for the Steinhaus sums $s(n) = x_1 + \ldots + x_n$, $n \in \boldsymbol{N}$, is whether they satisfy the "Feller property"

$$\lim_{n \to \infty} \frac{s(n)}{n \operatorname{Log} n} = 1. \tag{3.1}$$

This turns out to be true and, moreover, the $s(n)$ exhibit second order asymptotics comparable to what is shown for S_n in [3], mentioned in Section 5 below, but with greater precision: We shall explicitly describe a function ξ defined on $[1/2, 1]$, with range $(0, 2]$, such that $s(n)/n = \xi(\gamma_n) + \operatorname{Log} n$ for all $n \in \boldsymbol{N}$, where

$$\gamma_n = n/2^{\lceil \operatorname{Log} n \rceil} \tag{3.2}$$

is an indicator of the position of n between two consecutive powers of 2, so that $1/2 < \gamma_n \leqslant 1$ for all $n \in \boldsymbol{N}$.

The equation $s(n)/n - \operatorname{Log} n = \xi(\gamma_n)$, proved in Theorem 3.3 below, obviously implies (3.1), and it shows that $s(n)/n - \operatorname{Log} n$ behaves periodically, with period length one, in the variable $\operatorname{Log} n$, $n \in \boldsymbol{N}$. Before establishing this equation, we begin with a definition of the function ξ, and describe its properties. Define

$$\xi(\gamma) := 2 - \frac{1}{\gamma} \sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} - \operatorname{Log} \gamma, \quad \tfrac{1}{2} \leqslant \gamma \leqslant 1, \tag{3.3}$$

where the ε_k's, zeros and ones, are the binary digits of γ in the dyadic expansion

$$\gamma = \sum_{k=0}^{\infty} \frac{\varepsilon_k}{2^k}. \tag{3.4}$$

A graph of $\xi(\cdot)$ appears in Figure 1.

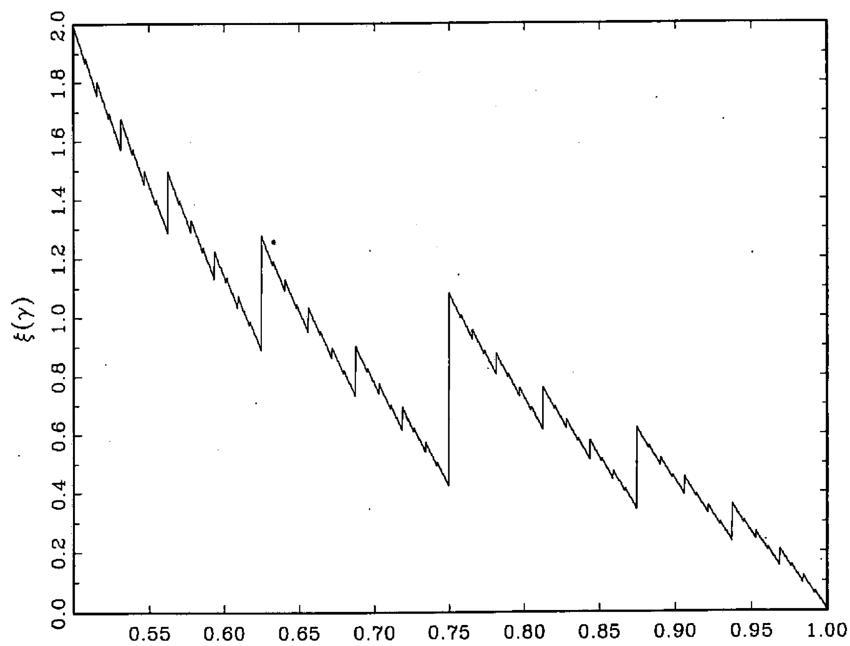

Fig. 1. The graph of $\xi(\gamma)$, $1/2 \leqslant \gamma \leqslant 1$

To avoid ambiguity with dyadic rationals, *affecting the value of* $\xi(\gamma)$, we require an infinite number of the ε_k's to be zero, in which case (3.4) is said to be of *standard form*. Otherwise, it is said to be of *non-standard form*. Thus,

$$(\varepsilon_0, \varepsilon_1, \ldots) = (0, 1, 1, 0, 0, 0, \ldots)$$

and $(0, 1, 0, 1, 1, 1, \ldots)$ give rise to the standard and non-standard forms, respectively, for $\gamma = 3/4$. Each dyadic rational γ in $(1/2, 1]$ has both forms, and, consequently, contributes a point of discontinuity to $\xi(\gamma)$, the only discontinuities. (There is no problem at $\gamma = 1/2$ since all discontinuities occur from the left side.) For any $1/2 \leqslant \gamma < 1$, we have $\varepsilon_0 = 0$ and $\varepsilon_1 = 1$, and $(1, 0, 0, \ldots)$ is the standard-form representation of $\gamma = 1$.

THEOREM 3.1. *The function* $\xi(\cdot)$, *defined in* (3.3),

(i) *assumes the value* 2 *at* $\gamma = 1/2$ *and* $\gamma = 1$;

(ii) *elsewhere, assumes the values* $\{\xi(\gamma): 1/2 < \gamma < 1\} = (0, 2)$, *so that* $\{\xi(\gamma): 1/2 \leqslant \gamma \leqslant 1\} = (0, 2]$;

(iii) *is right-continuous*;

(iv) *is left-continuous except for the dyadic rationals greater than* $1/2$;

(v) *has upward jumps only, for the dyadic rationals, of size* $2 = 2/1$ *at* $\gamma = 1$, *of size* $2/3$ *at* $\gamma = 3/4$, *of size* $2/5$ *and* $2/7$ *at* $\gamma = 5/8$ *and* $7/8$, *respectively, of sizes* $2/9, 2/11$, *and* $2/13$, *at* $\gamma = 9/16, 11/16$, *and* $13/16$, *respectively, etc. Thus* $\xi(\cdot)$ *has unbounded variations even locally.*

Both in the present and the next section, the proofs are deferred to the end of the section.

We shall now describe two different formulae for the Steinhaus sums $s(n)$. The first is a direct consequence of (2.1) and takes the form

$$s(n) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{2^k} + \frac{1}{2} \right\rfloor 2^k = \sum_{k=1}^{\lfloor \operatorname{Log} n \rfloor + 1} \left\lfloor \frac{n}{2^k} + \frac{1}{2} \right\rfloor 2^k, \qquad n \in \boldsymbol{N}. \tag{3.5}$$

The second depends on the binary digits, zeros and ones, in the dyadic expansion of n itself, written as

$$n = \sum_{r=0}^{\infty} a_r 2^r = \sum_{r=0}^{\lfloor \operatorname{Log} n \rfloor} a_r 2^r, \tag{3.6}$$

where, of course, $a_r = 0$ for $r > \lfloor \operatorname{Log} n \rfloor$.

THEOREM 3.2. *In terms of the coefficients* a_r *appearing in* (3.6),

$$s(n) = \sum_{r=0}^{\infty} (r+2) a_r 2^r = \sum_{r=0}^{\lfloor \operatorname{Log} n \rfloor} (r+2) a_r 2^r, \qquad n \in \boldsymbol{N}.$$

A simple corollary to Theorem 3.2 is the "doubling relationship":

$$s(2n)/2n - \operatorname{Log} 2n = s(n)/n - \operatorname{Log} n, \qquad n \in \boldsymbol{N}. \tag{3.7}$$

This is because, by setting $b_0 = 0$ and $b_j = a_{j-1}$, $j \in \boldsymbol{N}$, Theorem 3.2 gives

$$s(2n) = \sum_{j=0}^{\infty} (j+2) b_j 2^j = \sum_{j=1}^{\infty} (j+2) a_{j-1} 2^j = \sum_{r=0}^{\infty} (r+2) a_r 2^{r+1} + \sum_{r=0}^{\infty} a_r 2^{r+1},$$

so that

$$s(2n) = 2s(n) + 2n \qquad \text{for every } n \in \boldsymbol{N}.$$

Another doubling relationship appears in (4.4) below. The present doubling relationship is at the heart of the main result in this section.

THEOREM 3.3. *For every* $n \in N$,

$$s(n)/n - \operatorname{Log} n = \xi(\gamma_n),$$

where γ_n *and* $\xi(\cdot)$ *are defined in* (3.2) *and* (3.3), *respectively.*

Proof of Theorem 3.1. (i) This can easily be shown by direct evaluations. Standard forms of the dyadic expansions must be used for $\gamma = 1/2$ and 1.

(ii) Consider any $\gamma \in (1/2, 1)$. Thus, $\varepsilon_0 = 0$ and $\varepsilon_1 = 1$ in (3.4) and $\operatorname{Log}\gamma > -1$, so that

$$\sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} > \sum_{k=1}^{\infty} \frac{\varepsilon_k}{2^k} = \gamma,$$

and hence

$$\xi(\gamma) = 2 - \frac{1}{\gamma} \sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} - \operatorname{Log}\gamma < 2 - 1 - (-1) = 2.$$

This means that $\xi(\gamma) < 2$ on $(1/2, 1)$.

To show $\xi(\gamma) > 0$ on $(1/2, 1)$, it is enough to verify that

$$\frac{1}{\gamma} \sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} \leqslant 2, \tag{3.8}$$

since $\operatorname{Log}\gamma < 0$ for such γ. To this end, it is convenient to introduce a new function $\eta(\cdot)$ defined for *all* points $(\varepsilon_2, \varepsilon_3, \ldots) \in \{0, 1\}^{\infty}$, with $\varepsilon_0 \equiv 0$ and $\varepsilon_1 \equiv 1$:

$$\eta(\varepsilon_2, \varepsilon_3, \ldots) := \frac{\sum_{k=0}^{\infty} k\varepsilon_k/2^k}{\sum_{r=0}^{\infty} \varepsilon_r/2^r} = \frac{\frac{1}{2} + \sum_{k=2}^{\infty} k\varepsilon_k/2^k}{\frac{1}{2} + \sum_{r=2}^{\infty} \varepsilon_r/2^r}, \qquad (\varepsilon_2, \varepsilon_3, \ldots) \in \{0, 1\}^{\infty}. \tag{3.9}$$

Observe that

$$\eta(\varepsilon_2, \varepsilon_3, \ldots) = \frac{1}{\gamma} \sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k}$$

whenever (3.4) is of standard form. Also, note that $\eta(1, 1, \ldots) = 2$. Thus inequality (3.8) will be established if we show that the function $\eta(\cdot)$ is non-decreasing in each of its components on its whole domain $\{0, 1\}^{\infty}$. To this end, let $(\varepsilon_2, \varepsilon_3, \ldots) \in \{0, 1\}^{\infty}$ and $(\tau_2, \tau_3, \ldots) \in \{0, 1\}^{\infty}$ have common components everywhere except for an index $l \geqslant 2$, where $\varepsilon_l = 0$ and $\tau_l = 1$. It is enough to show that $\eta(\varepsilon_2, \varepsilon_3, \ldots) \leqslant \eta(\tau_2, \tau_3, \ldots)$. By using the formula in (3.9), this holds if and only if

$$\sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} \left(\sum_{r=0}^{\infty} \frac{\varepsilon_r}{2^r} + \frac{1}{2^l} \right) = \sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} \sum_{r=0}^{\infty} \frac{\tau_r}{2^r} \leqslant \sum_{k=0}^{\infty} \frac{k\tau_k}{2^k} \sum_{r=0}^{\infty} \frac{\varepsilon_r}{2^r} = \left(\sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^k} + \frac{l}{2^l} \right) \sum_{r=0}^{\infty} \frac{\varepsilon_r}{2^r},$$

which holds if and only if

$$\sum_{k=0}^{\infty} \frac{k\varepsilon_k}{2^{k+l}} \leqslant \sum_{r=0}^{\infty} \frac{l\varepsilon_r}{2^{r+l}},$$

which holds if and only if

$$\sum_{k=0}^{\infty} \frac{(l-k)\,\varepsilon_k}{2^k} \geqslant 0.$$

But

$$\sum_{k=0}^{\infty} \frac{(l-k)\,\varepsilon_k}{2^k} \geqslant \frac{l-1}{2} - \sum_{k=l+1}^{\infty} \frac{k-l}{2^k} = \frac{l-1}{2} - \frac{1}{2^{l+1}} \sum_{m=1}^{\infty} \frac{m}{2^{m-1}} = \frac{l-1}{2} - \frac{1}{2^{l-1}} \geqslant 0$$

for every $l \geqslant 2$. (Since this inequality is strict for $l \geqslant 3$, the function $\eta(\cdot)$ is strictly increasing in each of the l^{th} components whenever $l \geqslant 3$; and likewise for $l = 2$ except at $(0, 1, 1, 1, \ldots)$, corresponding to $\gamma = 3/4$.) This completes the proof that $\xi(\gamma) > 0$ on $(1/2, 1)$.

It remains to show that the range of $\xi(\gamma)$ on $(1/2, 1)$ is the entire interval $(0, 2)$. This requires some knowledge of parts (iii), (iv) and (v). Briefly, the function $\xi(\cdot)$ makes its way from the value 2 at $\gamma = 1/2$ to $0+$ at $\gamma = 1-$ interrupted by *upward* (never downward) jumping discontinuities at the dyadic rationals, which cause some values of the interval $(0, 2)$ to be visited more than once; no values in $(0, 2)$ can be skipped. The details are omitted.

(iii)–(v) The function $\eta(\cdot)$, defined in (3.9), is also relevant for the proofs of these parts.

The issue of continuity is this: Let γ be a fixed point in $[1/2, 1]$ with standard dyadic expansion defined by $(\varepsilon_0, \varepsilon_1, \ldots)$, and suppose δ, with standard dyadic expansion defined by $(\tau_0, \tau_1, \ldots)$, approaches γ within $[1/2, 1]$. If the convergence to γ is from the right, then τ_k eventually assumes the value of ε_k, $k \geqslant 2$, so that $\eta(\tau_2, \tau_3, \ldots)$ converges to $\eta(\varepsilon_2, \varepsilon_3, \ldots)$. From this it easily follows that $\xi(\delta)$ converges to $\xi(\gamma)$, thereby establishing the right-continuity asserted in part (iii). The same kind of reasoning applies to points δ converging to γ from the left when γ is *not* a dyadic rational. This explains part (iv).

But when δ approaches a dyadic rational γ from the left, the standard-form binary digits τ_k approach the binary digits of γ that are in the *non-standard form*, not in the standard one. For instance, when $\gamma = 3/4$, and δ approaches γ from the left, then

$$(\tau_0, \tau_1, \tau_2, \ldots) \to (0, 1, 0, 1, 1, 1, \ldots), \qquad \text{not } (0, 1, 1, 0, 0, 0, \ldots).$$

This produces an upward jump at $\gamma = 3/4$ of size

$$\xi\left(\frac{3}{4}\right) - \xi\left(\frac{3}{4}-\right) = \eta(0, 1, 1, 1, \ldots) - \eta(1, 0, 0, 0, \ldots) =$$

$$=\frac{4}{3}\left(\frac{1}{2}+\sum_{k=3}^{\infty}\frac{k}{2^k}\right)-\frac{4}{3}\left(\frac{1}{2}+\frac{2}{4}\right)=\frac{4}{3}\left(\frac{1}{2}+1\right)-\frac{4}{3}=\frac{2}{3},$$

as asserted in part (v) of the theorem.

To conclude: The proofs of (iii) and (iv), *as described above*, are straightforward. The proof of (v) has been indicated. While the remaining details for (v) are also straightforward and are similar in kind, they are somewhat tedious and will not be given. ■

Proof of Theorem 3.2. From (3.6),

$$\frac{n}{2^k}+\frac{1}{2}=\sum_{r=k}^{\infty}a_r 2^{r-k}+\left\{\sum_{r=0}^{k-1}a_r 2^{r-k}+\frac{1}{2}\right\},\qquad k\in\boldsymbol{N},$$

where the first sum is an integer, and the bracketed quantity is a number in the interval $[2^{-1}, 1-2^{-k}]$ or in the interval $[2^{-1}, 3\cdot 2^{-1}-2^{-k}]$, according as $a_{k-1}=0$ or 1. Hence

$$\left\lfloor\frac{n}{2^k}+\frac{1}{2}\right\rfloor=a_{k-1}+\sum_{r=k}^{\infty}a_r 2^{r-k}. \tag{3.10}$$

Thus (3.5) becomes

$$s(n)=\sum_{k=1}^{\infty}\left\lfloor\frac{n}{2^k}+\frac{1}{2}\right\rfloor 2^k=\sum_{k=1}^{\infty}\left\{\sum_{r=k}^{\infty}a_r 2^r+a_{k-1}2^k\right\}=\sum_{r=1}^{\infty}\sum_{k=1}^{r}a_r 2^r+\sum_{k=1}^{\infty}a_{k-1}2^k$$

$$=\sum_{r=1}^{\infty}ra_r 2^r+2\sum_{r=0}^{\infty}a_r 2^r=\sum_{r=0}^{\infty}(r+2)\,a_r 2^r,$$

as asserted. ■

Proof of Theorem 3.3. The definition of the function $\xi(\cdot)$, given in (3.3), is described in terms of the binary digits ε_k of the dyadic expansion of γ shown in (3.4). Naturally, the values these digits assume when $\gamma=\gamma_n$ depend on n. The dependence is linked to the values of the binary digits a_r in the dyadic expansion of n appearing in (3.6). The precise relationship must be worked out: By (3.2),

$$n=2^{\lceil \operatorname{Log} n\rceil}\gamma_n\quad\text{and}\quad \operatorname{Log} n=\lceil\operatorname{Log} n\rceil+\operatorname{Log}\gamma_n. \tag{3.11}$$

With (3.6), we have

$$\sum_{k=0}^{\infty}\frac{\varepsilon_k(n)}{2^k}:=\gamma_n=\frac{n}{2^{\lceil\operatorname{Log} n\rceil}}=\frac{1}{2^{\lceil\operatorname{Log} n\rceil}}\sum_{r=0}^{\lceil\operatorname{Log} n\rceil}a_r 2^r=\sum_{k=0}^{\lceil\operatorname{Log} n\rceil}\frac{a_{\lceil\operatorname{Log} n\rceil-k}}{2^k}.$$

Thus,

$$\varepsilon_k(n)=\begin{cases}a_{\lceil\operatorname{Log} n\rceil-k} & \text{for } k=0,1,\ldots,\lceil\operatorname{Log} n\rceil,\\ 0 & \text{for } k>\lceil\operatorname{Log} n\rceil.\end{cases} \tag{3.12}$$

Then, by (3.6) and Theorem 3.2,

$$s(n)-(\lceil \operatorname{Log} n\rceil+2)n = \sum_{r=0}^{\lceil \operatorname{Log} n\rceil}(r-\lceil \operatorname{Log} n\rceil)a_r 2^r = -\sum_{k=0}^{\lceil \operatorname{Log} n\rceil} k a_{\lceil \operatorname{Log} n\rceil - k} 2^{\lceil \operatorname{Log} n\rceil - k}$$
$$= -2^{\lceil \operatorname{Log} n\rceil}\sum_{k=0}^{\infty}\frac{k\varepsilon_k(n)}{2^k}.$$

Combined with (3.11), this yields

$$\frac{s(n)}{n}-\operatorname{Log} n = 2-\frac{2^{\lceil \operatorname{Log} n\rceil}}{n}\sum_{k=0}^{\infty}\frac{k\varepsilon_k(n)}{2^k}+\lceil \operatorname{Log} n\rceil-\operatorname{Log} n = 2-\frac{1}{\gamma_n}\sum_{k=0}^{\infty}\frac{k\varepsilon_k(n)}{2^k}-\operatorname{Log}\gamma_n,$$

which is $\xi(\gamma_n)$, in accordance with the definition given in (3.3). ∎

4. Steinhaus games. Suppose that, instead of a St. Petersburg game, Paul were offered a chance to play a "Steinhaus game", a game with a payoff dictated by the empirical distribution function $\tilde{F}_n$, appearing in (2.3), based upon the first n elements of the Steinhaus sequence. How attractive would this be to Paul?

The previous section provides evidence of what Paul could expect to get from a single play of such a game, namely the amount $s(n)/n$, a little bit more than $\operatorname{Log} n$. This competes with an infinite expectation and, therefore, appears to be less attractive than the St. Petersburg game. This is so. But this is not the only basis upon which Paul should prefer the St. Petersburg game: The amount X that Paul earns is stochastically larger under the St. Petersburg game, i.e.

$$1-F(x) = \mathrm{P}\{X>x\} \geqslant \tilde{\mathrm{P}}_n\{X>x\} = 1-\tilde{F}_n(x), \qquad x\in \boldsymbol{R},$$

necessarily a strict inequality for some x, where $\tilde{F}_n(\cdot)$ and $F(\cdot)$ are the distribution functions for X, described in (2.3) and (2.4), respectively, for the Steinhaus and St. Petersburg games, and $\tilde{\mathrm{P}}_n$ denotes "probability" for the Steinhaus game. This follows directly from

THEOREM 4.1. *For every* $k\in \boldsymbol{N}$,

$$\mathrm{P}\{X>2^k\} = \frac{1}{2^k} \quad \textit{and} \quad \tilde{\mathrm{P}}_n\{X>2^k\} = \frac{1}{2^k}\left\{\frac{1}{n}\sum_{r=k}^{\infty} a_r 2^r\right\}.$$

Thus

$$n\{\tilde{F}_n(2^k)-F(2^k)\} = \frac{1}{2^k}\sum_{r=0}^{k-1} a_r 2^r. \tag{4.1}$$

From this, we obtain a formula for the Kolmogorov distance, justify the bounds described in (2.5), and obtain a "doubling relationship" parallel to (3.7).

THEOREM 4.2. *For every* $n\in \boldsymbol{N}$,

$$n\sup_{x\in\boldsymbol{R}}|\tilde{F}_n(x)-F(x) = \max_{k\in\boldsymbol{N}}\left(\frac{1}{2^k}\sum_{r=0}^{k-1} a_r 2^r\right). \tag{4.2}$$

Hence

$$\tfrac{1}{2} \leqslant n \sup_{x \in \boldsymbol{R}} |\tilde{F}_n(x) - F(x)| < 1 \tag{4.3}$$

and

$$2n \sup_{x \in \boldsymbol{R}} |\tilde{F}_{2n}(x) - F(x)| = n \sup_{x \in \boldsymbol{R}} |\tilde{F}_n(x) - F(x)|. \tag{4.4}$$

This last doubling relationship in (4.4) implies that the expression $n \sup \{|\tilde{F}_n(x) - F(x)|: x \in \boldsymbol{R}\}$ is a function of γ_n. More specifically, we have the following analogue of Theorem 3.3:

THEOREM 4.3. *For every* $n \in \boldsymbol{N}$,

$$n \sup_{x \in \boldsymbol{R}} |\tilde{F}_n(x) - F(x)| = \zeta(\gamma_n), \tag{4.5}$$

where

$$\zeta(\gamma) := \sup_{r=0,1,2,\ldots} \left(2^{r-1} \sum_{k=r}^{\infty} \frac{\varepsilon_k}{2^k}\right), \quad \tfrac{1}{2} \leqslant \gamma \leqslant 1,$$

and where $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$ *are the binary digits of* γ *in the dyadic expansion shown in* (3.4), *and* γ_n *is defined in* (3.2).

A graph of $\zeta(\cdot)$ appears in Figure 2.

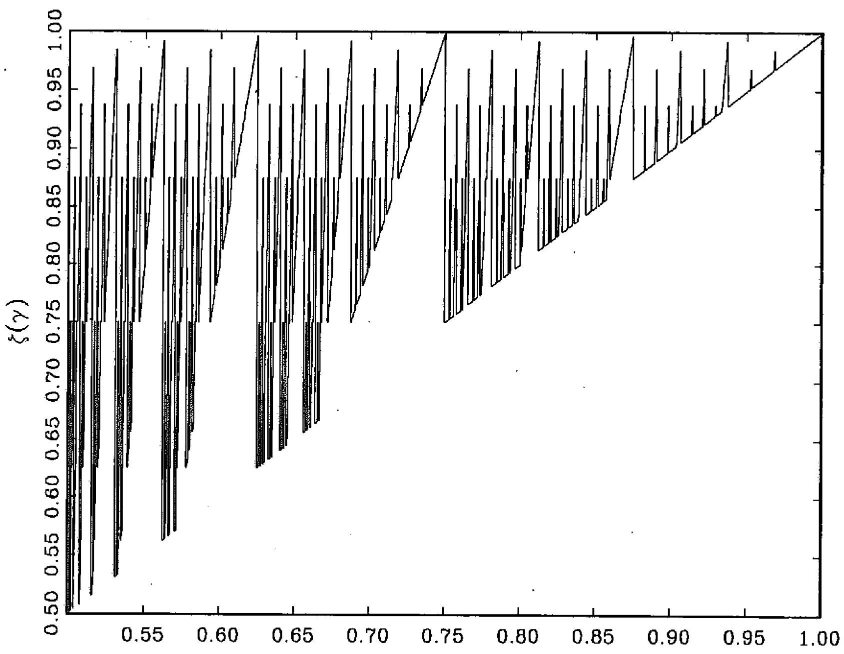

Fig. 2. The graph of $\zeta(\gamma)$, $1/2 \leqslant \gamma \leqslant 1$

Proof of Theorem 4.1. Clearly, by (1.1),

$$\mathrm{P}\{X > 2^k\} = \sum_{r=k+1}^{\infty} \mathrm{P}\{X = 2^r\} = \sum_{r=k+1}^{\infty} \frac{1}{2^r} = \frac{1}{2^k},$$

while

$$\tilde{\mathrm{P}}_n\{X > 2^k\} = \sum_{r=k+1}^{\infty} \tilde{\mathrm{P}}_n\{X = 2^r\} = \frac{1}{n} \sum_{r=k+1}^{\infty} \left\lfloor \frac{n}{2^r} + \frac{1}{2} \right\rfloor,$$

according to (2.1). In turn, with (3.10), the latter sum becomes

$$\begin{aligned}\sum_{r=k+1}^{\infty} \{a_{r-1} + \sum_{j=r}^{\infty} a_j 2^{j-r}\} &= \sum_{j=k}^{\infty} a_j + \sum_{j=k+1}^{\infty} a_j 2^j \sum_{r=k+1}^{j} \left(\frac{1}{2}\right)^r \\ &= \sum_{j=k}^{\infty} a_j + \sum_{j=k+1}^{\infty} a_j 2^j \left(\frac{1}{2^k} - \frac{1}{2^j}\right) \\ &= a_k + \frac{1}{2^k} \sum_{j=k+1}^{\infty} a_j 2^j = \frac{1}{2^k} \sum_{j=k}^{\infty} a_j 2^j.\end{aligned}$$

So the formula stated for $\tilde{\mathrm{P}}_n\{X > 2^k\}$ also follows. Thus, by (3.6),

$$n[\tilde{F}_n(2^k) - F(2^k)] = n[\mathrm{P}\{X > 2^k\} - \tilde{\mathrm{P}}_n\{X > 2^k\}] = \frac{1}{2^k} \sum_{r=0}^{\infty} a_r 2^r - \frac{1}{2^k} \sum_{r=k}^{\infty} a_r 2^r,$$

and this gives (4.1). ∎

Proof of Theorem 4.2. Equation (4.2) follows immediately from (4.1). A maximum is attainable on account of the fact that $a_r = 0$ for $r > \lfloor \mathrm{Log}\, n \rfloor$; cf. (3.6). This implies that the quantity within the brackets on the right side of (4.2) decreases as k increases, provided that k is large enough.

In turn, (4.3) easily follows from (4.2): The middle term in (4.3) is bounded below by $a_{k-1}/2$ for every $k \in N$, and hence bounded below by 1/2. (Some $a_{k-1} = 1$, on account of (3.6).) This lower bound is attainable; equality holds when n is a power of 2. The inequality on the right side of (4.3) arises from the fact that

$$\frac{1}{2^k} \sum_{r=0}^{k-1} a_r 2^r \leqslant \frac{1}{2^k} \sum_{r=0}^{k-1} 2^r = 1 - \frac{1}{2^k} < 1, \quad k \in N.$$

The doubling relationship in (4.4) also follows from (4.2): As in the argument for (3.7), the left side of (4.4) is equal to

$$\max_{k \in N} \frac{1}{2^k} \sum_{r=1}^{k-1} a_{r-1} 2^r = \max_{k \in N} \frac{1}{2^{k-1}} \sum_{j=0}^{(k-1)-1} a_j 2^j = \max_{l \in N} \frac{1}{2^l} \sum_{r=1}^{l-1} a_r 2^r,$$

which, by using (4.2) once more, is equal to the right side of (4.4). ∎

Proof of Theorem 4.3. To establish the claimed equation, it must be shown that the right sides of (4.2) and (4.5) are equal. Let $\varepsilon_0(n)$, $\varepsilon_1(n)$, $\varepsilon_2(n)$, ...

denote the binary digits of γ_n. The key fact is (3.12), yielding

$$\zeta(\gamma_n) = \max_{r=0,1,\ldots} 2^{r-1} \sum_{k=r}^{\infty} \frac{\varepsilon_k(n)}{2^k} = \max_{r=0,1,\ldots} 2^{r-1} \sum_{k=r}^{\lceil \mathrm{Log}\, n \rceil} \frac{a_{\lceil \mathrm{Log}\, n \rceil - k}}{2^k}$$

$$= \max_{r=0,1,\ldots} \frac{1}{2^{\lceil \mathrm{Log}\, n \rceil - r + 1}} \sum_{j=0}^{\lceil \mathrm{Log}\, n \rceil - r} a_j 2^j = \max_{k \in N} \frac{1}{2^k} \sum_{j=0}^{k-1} a_j 2^j,$$

which is the right side of (4.2). ∎

5. The Feller and the Steinhaus resolutions. The Feller and the Steinhaus resolutions require Paul to pay $n \,\mathrm{Log}\, n$ and $s(n) = x_1 + \ldots + x_n$ ducats for his winnings $S_n = X_1 + \ldots + X_n$ in n games, respectively. How attractive are these to Peter? Since, by Theorem 3.3, $s(n) = n[\xi(\gamma_n) + \mathrm{Log}\, n]$ for each $n \in N$, and $\xi(\gamma) > 0$ for all $\gamma \in [1/2, 1]$ by part (ii) of Theorem 3.1, he certainly prefers Steinhaus' price to the Feller premiums. However, the improvement, $n\xi(\gamma_n)$, is slight when γ_n is near and on the left side of a dyadic rational that is itself close to unity. This occurs when n is slightly less than a power of 2. See Figure 1. The best improvement, from Peter's point of view, is in the case when n is a power of 2 or just passed a power of 2, in which case $\xi(\gamma_n) = 2$ or is very close to 2. Table 1 gives the prices that Paul pays, rounded down to integers for the Feller prices, and Peter's winning probabilities for $n = 1, \ldots, 32$ under the two premium systems. The differences shown in the sixth column are, of course, the differences $\mathrm{P}\{S_n \leqslant s(n)\} - \mathrm{P}\{S_n \leqslant n \,\mathrm{Log}\, n\}$ between the probabilities in the fifth and the fourth columns. Our program calculates the probabilities and the differences with precise fifth decimals; then we round off the numbers obtained. Since the number of vectors of integers to be checked grows very rapidly with n, it is difficult to get beyond $n = 40$. (Note added in proof: Meanwhile we were able to get beyond $n = 4000$. The pattern remains and the asymptotics, noted below, take over more closely.)

So, while his winning probabilities do indeed improve, Peter is unhappy even with the largest possible improvement under the Steinhaus plan: It always falls short of the "median resolution" when his winning probabilities would be about 1/2.

The bad news that Table 1 conveys to Peter is not limited to small n: Based on an infinitely divisible "merging approximation" developed in [3], we can describe asymptotic Feller and Steinhaus "winning probabilities" $p_n^{(\mathrm{F})}$ and $p_n^{(\mathrm{S})}$, respectively, and a "median price"

$$m(n) := n[m_n(1/2) + \mathrm{Log}\, n],$$

which satisfy

$$|\mathrm{P}\{S_n \leqslant n \,\mathrm{Log}\, n\} - p_n^{(\mathrm{F})}| = O((\mathrm{Log}^2 n)/n),$$

$$|\mathrm{P}\{S_n \leqslant s(n)\} - p_n^{(\mathrm{S})}| = O((\mathrm{Log}^2 n)/n), \qquad |\mathrm{P}\{S_n \leqslant m(n)\} - \tfrac{1}{2}| = O((\mathrm{Log}^2 n)/n)$$

Table 1. Feller and Steinhaus prices and winning probabilities

n	$\lfloor n \operatorname{Log} n \rfloor$	$s(n)$	$\mathrm{P}\{S_n \leqslant n \operatorname{Log} n\}$	$\mathrm{P}\{S_n \leqslant s(n)\}$	Difference
1	0	2	0.0000	0.5000	0.5000
2	2	6	0.0000	0.5000	0.5000
3	4	8	0.0000	0.3125	0.3125
4	8	16	0.0625	0.4727	0.4102
5	11	18	0.0313	0.3535	0.3223
6	15	22	0.0625	0.3184	0.2559
7	19	24	0.0762	0.2385	0.1624
8	24	40	0.1265	0.4566	0.3301
9	28	42	0.1246	0.3881	0.2634
10	33	46	0.1190	0.3551	0.2361
11	38	48	0.1420	0.2982	0.1561
12	43	56	0.1333	0.3238	0.1905
13	48	58	0.1498	0.2760	0.1263
14	53	62	0.1397	0.2549	0.1151
15	58	64	0.1514	0.2150	0.0635
16	64	96	0.1600	0.4494	0.2895
17	69	98	0.1483	0.4108	0.2625
18	75	102	0.1543	0.3874	0.2331
19	80	104	0.1592	0.3514	0.1922
20	86	112	0.1632	0.3577	0.1945
21	92	114	0.1664	0.3248	0.1584
22	98	118	0.1689	0.3063	0.1373
23	104	120	0.1709	0.2763	0.1054
24	110	136	0.1723	0.3308	0.1585
25	116	138	0.1732	0.3025	0.1293
26	122	142	0.1736	0.2865	0.1129
27	128	144	0.1736	0.2606	0.0870
28	134	152	0.1733	0.2672	0.0940
29	140	154	0.1727	0.2432	0.0705
30	147	158	0.1719	0.2301	0.0582
31	153	160	0.1710	0.2084	0.0374
32	160	224	0.1789	0.4465	0.2676

as $n \to \infty$. While $m(n)$ comes close to giving Peter a valid median price, we find that for every $n \in N$

$$0.2070 < p_n^{(\mathrm{F})} < 0.2073, \qquad 0.2071 < p_n^{(\mathrm{S})} < 0.4421,$$

and

$$2.5844 < m_n(1/2) < 2.6050.$$

For comparison with the latter, we remind the reader that $s(n) = n[\xi(\gamma_n) + \operatorname{Log} n]$, so that

$$\xi(\gamma_n) \leqslant \xi(1/2) = \xi(1) = 2 < m_n(1/2), \qquad n \in N.$$

But Peter is not satisfied even with this median resolution (in spite of its superficially attractive feature: "Why, half the time I win, half the time you win", Paul would argue), because he cannot win more than $m(n)$ ducats if he does, while Paul's winnings may be huge. The details of all this and many further considerations will be in [3].

Acknowledgements. We thank Jacek Koronacki and Jan Mielniczuk for arranging a copy of [10] for us. The authors were partially supported by NSF Grants DMS-9208067 and DMS-8902973, respectively.

REFERENCES

[1] D. Bernoulli, *Specimen theoriae novae de mensura sortis*, Commentarii Academiae Scientiarum Imperialis Petropolitanae, Tomus V, Ad annos 1730 et 1731, pp. 175–193; published in 1738. [Reprinted in: *Die Werke von Daniel Bernoulli*, Vol. II, Birkhäuser, Basel 1982, pp. 223–234. English translation by L. Sommer: *Exposition of a new theory of the measurement of risk*, Econometrica 22 (1954), pp. 23–36; reprinted in: W. J. Baumol and S. M. Goldfeld (Eds.), *Precursors in Mathematical Economics: An Anthology*, Series of Reprints of Scarce Works on Political Economy, No. 19, London School of Economics and Political Science, London 1968, pp. 15–26; and in: A. N. Page (Ed.), *Utility Theory: A Book of Readings*, Wiley, New York 1968, pp. 199–214.]

[2] Y. S. Chow and H. Robbins, *On sums of independent random variables with infinite means and "fair" games*, Proc. Nat. Acad. Sci. U.S.A. 47 (1961), pp. 330–335.

[3] S. Csörgő and G. Simons, *Resolution of the St. Petersburg Paradox*, book manuscript in progress.

[4] J. Dutka, *On the St. Petersburg paradox*, Arch. Hist. Exact Sci. 39 (1988), pp. 13–39.

[5] W. Feller, *Note on the law of large numbers and "fair" games*, Ann. Math. Statist. 16 (1945), pp. 301–304.

[6] — *An Introduction to Probability Theory and Its Applications*, Vol. I, Wiley, New York 1950. [Second and third editions: 1957, 1968; Chapter X, Section 4, in all editions.]

[7] — *Review of Steinhaus* [10], Mathematical Reviews 12 (1951), p. 619.

[8] G. Jorland, *The Saint Petersburg paradox 1713–1937*, in: *The Probabilistic Revolution*, Vol. I: L. Krüger, L. D. Daston and M. Heidelberger (Eds.), *Ideas in History*, MIT Press, Cambridge, Massachusetts, pp. 157–190.

[9] P. R. de Montmort, *Essay d'analyse sur les jeux de hazard*, seconde edition, Revûe & augmentée de plusieurs Lettres, Jacque Quillau, Paris 1713. [Third edition: reprint of the second by Chelsea, New York 1980.]

[10] H. Steinhaus, *The so-called Petersburg paradox*, Colloq. Math. 2 (1949), pp. 56–58.

Department of Statistics
The University of Michigan
1444 Mason Hall
Ann Arbor, MI 48109-1027, U.S.A.

Department of Statistics
The University of North Carolina
322 Phillips Hall
Chapel Hill, NC 27599-3260, U.S.A.

Received on 10.7.1993