A Counterexample to Richard von Mises's Theory of Collectives

Jean Ville
Translation and introduction by Glenn Shafer
www.glennshafer.com, gshafer@rutgers.edu

November 20, 2005

Abstract

This document presents a translation from French into English of a passage from pp. 55–63 of Jean Ville's *Étude critique de la notion de collectif*, which was published in 1939.

1 Introduction

In the passage translated here, Jean Ville gives an example of an event that is assigned probability zero by classical probability theory but cannot be ruled out by conditions of the type Richard von Mises used to define his concept of a collective. Ville and others considered the existence of such examples a shortcoming of von Mises's theory, and this motivated Ville to introduce martingales into probability theory.

For various historical reasons, Ville's counterexample to von Mises theory has not always been noticed. Von Mises's approach is still sometimes discussed as if it were in contention to serve as a foundation for probability theory, without reference to the counterexample. One reason for this neglect may be linguistic. Ville's book has never been translated into English, and the only detailed explanation of his counterexample that has appeared in English, by Loveland, carries its own considerable overhead. So this translation may help right the balance.

Ville's work first appeared in his doctoral dissertation, which he defended in 1939 at the University of Paris. His book, *Étude critique de la notion de collectif*, consisting of the dissertation and an additional introductory chapter, was published after the defense. Ville had obtained his results, however, during a stay in Vienna in 1934–35, where he participated in Karl Menger's colloquium, which was also frequented by von Mises and Abraham Wald, and he had announced his results in the *Comptes rendus* in 1936. We translate the announcement in an appendix.

The passage translated here is concerned only with Ville's counterexample to von Mises's concept of a collective; it does not include any of Ville's positive theory, in which he removed the defect he detected by generalizing von Mises's concept of a *selection*, which uses outcomes of previous trials in deciding whether to include the next trial in a subsequence in which frequencies are to be checked, to the concept of a *martingale*, which uses previous trials to decide what portion of one's current capital to risk on the next trial. Instead of requiring only that each subsequence extracted by a selection have the right frequencies, Ville required that the capital achieved by any martingale should remain bounded. Ville's positive theory is the basis of the modern game-theoretic foundation for probability (see Shafer & Vovk, 2001).

The passage translated begins near the top of p. 55, in the middle of §4 of Chapter II. It includes all of §5 and §6 and concludes at the end of §6 in the middle of p. 63.

Earlier in Chapter II, Ville discusses the concept of collective developed by von Mises and Wald. We may summarize his discussion as follows:

1. In the case Ville emphasizes, where the outcome of each trial is either a 0 or a 1, the question is how to extract a subsequence from a sequence $x = x^1x^2\dots$ of 0s and 1s.

2. As Ville explains, a *selection* S is defined by a sequence $f_0, f_1, f_2, \dots$ of functions. Here f_n is a function of n binary variables and itself takes the values 0 and 1. The subsequence of x extracted by S includes x^n if and only if $f_{n-1}(x^1 \dots x^{n-1}) = 1$.

3. There is a selection that leaves every sequence unchanged; this is the selection for which $f_n(x^1 \dots x^n) = 1$ for all n and all $x^1 \dots x^n$.

4. A *system of selections* $\mathcal{S}$ is a countably infinite set of selections that includes the selection that leaves every sequence unchanged.

5. The sequence x is a *collective* with respect to a system of selections $\mathcal{S}$ and a number $p \in (0,1)$ if for every selection $S \in \mathcal{S}$ that extracts an infinite subsequence from x,[1] the proportion of 1s in an initial segment of the extracted subsequence converges, as the length of initial segment goes to infinity, to p.

6. Ville writes $\mathbf{K}(\mathcal{S}, p)$ for the set of all collectives with respect to $\mathcal{S}$ and p.

7. As Wald had pointed out, $\mathbf{K}(\mathcal{S}, p)$ is always nonempty. For any system of selections $\mathcal{S}$ and any $p \in (0,1)$, there is at least one sequence x that qualifies as a collective with respect to $\mathcal{S}$ and p. If fact, the set of sequences that do qualify have p-measure one.[2] This implies Ville's Theorem 3, with which we begin the translation.

2 Translation

Theorem 3 *Suppose $p \in (0,1)$, and suppose A is a set of sequences of 0s and 1s represented by a subset of $(0,1)$. Then a necessary condition for being able to exclude a sequence x from A by a requirement of irregularity based on the notion of selection (i.e., a necessary condition for finding a system of selections $\mathcal{S}$ such that no $x \in \mathbf{K}(\mathcal{S}, p)$ is in A) is that A have p-measure zero.*

The converse of this theorem comes naturally to mind, but it is false. In fact:

Theorem 4 *Given any number p between 0 and 1, one can construct a set G of p-measure zero such that for any countably infinite system of selections $\mathcal{S}$, the sets $x \in \mathbf{K}(\mathcal{S}, p)$ and G have at least one point in common.*

This theorem is a consequence of the following theorem:

[1] If $f_m(x^1 \dots x^m) = 0$ for all m larger than some integer N, then the extracted subsequence stops, remaining finite rather continuing as an infinite sequence.

[2] This means that this set has probability one with respect to the probability distribution for $x^1, x^2, \dots$ obtained by assuming that the x^n are independent random variables, each equal to 1 with probability p and 0 with probability $1-p$.

Theorem 4 $'$ *Suppose $\mathcal{S}$ is a countably infinite set of selections, $p \in [0,1]$, and $\phi(\mu)$ is a continuous increasing function that always takes positive values and satisfies*

$$\lim_{\mu\to\infty} \frac{1}{\mu}\phi(\mu) = 0 \quad \textit{and} \quad \lim_{\mu\to\infty} \phi(\mu) = \infty$$

but may diverge to infinity arbitrarily slowly.

Then there exists a collective x of type $\mathbf{K}(\mathcal{S}, p)$ with the following property. For any selection A in the system $\mathcal{S}$ that extracts an infinite sequence from x, there exist two positive numbers α and β such that for every positive integer μ, the number ν of 1s among the first μ terms of the extracted sequence satisfies

$$-\frac{\alpha}{\mu} \leq \frac{\nu}{\mu} - p < \frac{\alpha}{\mu} + \beta\frac{\phi(\mu)}{\mu}. \tag{19}$$

First let us show that Theorem $4'$ implies Theorem 4.

If we consider a sequence of independent trials in which an event can happen with probability p, it is well known (see, for example, Lévy [1], p. 258ff) that *if f_μ represents the frequency of success in the first μ trials, then for any positive number δ, the inequality*

$$|f_\mu - p| > \frac{\delta}{\sqrt{\mu}} \tag{20}$$

happens infinitely many times with probability one. So if we set

$$\phi(\mu) = \mu^\epsilon \qquad\qquad (0 < \epsilon < \frac{1}{2})$$

in the statement of Theorem $4'$, and we designate by G the set of points x corresponding to sequences for which the inequality (20) happens only a finite number of times (we may take $\delta = 1$, for example), we see that G has p-measure zero and at least one point in common with the set $\mathbf{K}(\mathcal{S}, p)$ for any $\mathcal{S}$, thus establishing Theorem 4.

We will demonstrate Theorem $4'$ by constructing a collective of type $\mathbf{K}(\mathcal{S}, p)$ satisfying its conclusions. For this construction, we need to set out some preliminary definitions:

5. The composition of selections

Definition 4 *Given two selections A_1 and A_2, defined by sequences of functions $\{f_i^1\}$ and $\{f_i^2\}$, respectively (Def. 2, p. 42), let A be the selection defined by the sequence $\{g_i\}$, where*

$$g_i = 1 - (1 - f_i^1)(1 - f_i^2).$$

We call A the sum of A_1 and A_2 and write

$$A = A_1 + A_2.$$

The selection A chooses a term if and only if either A_1 *or* A_2 chooses it. From the preceding definition, one can easily figure out definitions of the selections

$$A_1 + A_2 + \cdots + A_n \quad \text{and} \quad \sum_{i=1}^{\infty} A_i.$$

Definition 5 *Given two selections A_1 and A_2, defined by $\{f_i^1\}$ and $\{f_i^2\}$, respectively, let B be the selection defined by $\{g_i\}$, where*

$$g_i = f_i^1 f_i^2.$$

We call B the inner product of A_1 and A_2 and write

$$A = A_1 A_2.$$

B is the selection that chooses a term if and only if A_1 *and* A_2 choose it simultaneously. The operations of sum and inner product are commutative and associative.

Definition 6 *If A_1 is a selection defined by $\{f_i^1\}$, the selection C defined by $\{g_i\}$, where*

$$g_i = 1 - f_i^1,$$

is called A_1's complementary selection; we write

$$C = 1 - A_1.$$

C is the selection that chooses a term if and only if A_1 does not choose it.

Definition 7 *If A_1 is a selection, and a is a number equal to 0 or 1, we write aA_1 for the selection that is identical to A_1 if $a = 1$ and equal to $1 - A_1$ if $a = 0$.*

The operations introduced by definitions 4 to 7 together constitute what we call the *composition* of selections. We will define yet one more operation on selections.

Definition 8 *Given a selection A_1 defined by $\{f_i^1\}$ and a positive integer m, we define a selection E by $\{g_i\}$, where*

$$g_i = f_i^1 \quad \textit{if} \sum_{j=0}^{i} f_j^1 \leq m,$$

$$g_i = 0 \quad \textit{if} \sum_{j=0}^{i} f_j^1 > m,$$

and we call it A_1's initial segment of length m, and we write

$$E = A_1^{(m)}.$$

The selection E chooses a term only if that term is chosen by A_1 and A_1 has chosen fewer than m earlier terms.

6. Construction of a collective of type $\mathbf{K}(\mathcal{S}, p)$. Let us now construct a collective that satisfies the conditions of Theorem 4′. We begin with a number p, a countably infinite system $\mathcal{S}$ of selections $A_1, A_2, \ldots, A_n, \ldots$, and a function $\phi(\mu)$ satisfying the hypothesis. We will define new selections, study them, use their properties to construct a particular sequence, show that this sequence is a collective with respect to the new selections, and then show that it is a collective with respect to $\mathcal{S}$.

1° *Composing selections $A_1, A_2, \ldots, A_n$.* The equation $2^\alpha = \phi(\mu)$ or

$$\alpha \log 2 = \log \phi(\mu) \tag{21}$$

determines an inverse $\mu = \psi(\alpha)$. Let $\{m_i\}$ be a sequence of integers such that

$$m_i > \psi(i+1) - \psi(i). \tag{22}$$

For each finite sequence $a^1 a^2 \dots a^m$ of 0s and 1s, we define two selections $B_{a^1 a^2 \dots a^m}$ and $C_{a^1 a^2 \dots a^m}$ as follows:

$$B_{a^1} = a^1 A_1, \qquad C_{a^1} = B_{a^1}^{(m_1)},$$

$$B_{a^1 a^2} = (1 - C_{a^1}) a^1 A_1 a^2 A_2, \qquad C_{a^1 a^2} = B_{a^1 a^2}^{(m_2)},$$

and, in general,

$$\begin{cases} B_{a^1 a^2 \dots a^{n-1} a^n} = (1 - C_{a^1})(1 - C_{a^1 a^2}) \dots (1 - C_{a^1 a^2 \dots a^{n-1}}) a^1 A_1 a^2 A_2 \dots a^n A_n \\ C_{a^1 a^2 \dots a^n} = B_{a^1 a^2 \dots a^n}^{(m_n)}. \end{cases} \tag{23}$$

The B are the new selections of which we spoke. The C are initial segments of the B.

2° *Properties of the selections B and C.* Consider an arbitrary sequence of 0s and 1s:

$$x = x^1 x^2 \dots x^i \dots \qquad (x^i = 0 \text{ or } 1).$$

Consider a particular term, say x^m. *No two distinct selections $C_{a^1 a^2 \dots a^n}$ and $C_{\alpha^1 \alpha^2 \dots \alpha^\nu}$ can choose x^m at the same time.* Suppose in fact that the two selections do choose x^m and yet there is a position i for which $a^i \neq \alpha^i$. Assume, to fix ideas, that $a^i = 1$ and $\alpha^i = 0$. By the first of the relations in (23), we see that x^m's being chosen by $C_{a^1 a^2 \dots a^n}$ implies its being chosen by A_i, and its being chosen by $C_{\alpha^1 \alpha^2 \dots \alpha^\nu}$ implies its being chosen by $1 - A_i$. By Definition 7, this is a contradiction.

If no such position i exists, then the selections are distinct only if $\nu \neq n$, say $\nu > n$, and the selections are of the form

$$C_{a^1 a^2 \dots a^n}, \qquad C_{a^1 a^2 \dots a^n \alpha^{n+1} \dots \alpha^{n+\nu}}.$$

Again, by the first of the relations in (23), we see that a term chosen by $C_{a^1 a^2 \dots a^n}$ can be chosen neither by $B_{a^1 a^2 \dots a^n \alpha^{n+1} \dots \alpha^{n+\nu}}$ nor, *a fortiori*, by $C_{a^1 a^2 \dots a^n \alpha^{n+1} \dots \alpha^{n+\nu}}$. So there is still a contradiction.

Inversely, each term of x is chosen by one of the selections $C_{a^1 a^2 \dots a^n}$. Indeed, suppose this were not so, and let x^m be a term chosen by none of the selections. For every n, x^m is chosen either by A_n or by $1 - A_n$. Let $a^n A_n$

be the one of the two that does choose x^m, and consider the sequence $\{a^n\}$ and the corresponding sequence

$$C_{a^1}, C_{a^1a^2}, \ldots, C_{a^1a^2\ldots a^n}, \ldots \tag{24}$$

According to the first of the relations in (23), the fact that x^m is chosen by none of the C means that it is chosen by all the $B_{a^1a^2\ldots a^n}$ corresponding to the selections in the sequence (24). According to the second of the relations in (23), each of the $C_{a^1a^2\ldots a^n}$, because they do not choose x^m, must choose m_n terms among the first m terms of x, and this leads to a contradiction, because no two of these selections can choose the same term and $\sum_{n=1}^{\infty} m_n = \infty$.

So we have shown that x^m is chosen by one and only one of the selections in the sequence (24), say $C_{a^1a^2\ldots a^n}$. We notice that it is chosen by the selections $B_{a^1}, B_{a^1a^2}, \ldots, B_{a^1a^2\ldots a^n}$, because it is not chosen by $C_{a^1}, C_{a^1a^2}, \ldots, C_{a^1a^2\ldots a^{n-1}}$. We deduce from this that each of these last selections had already chosen, respectively, $m_1, m_2, \ldots, m_{n-1}$ terms in x.

To summarize: *If we write $C_{a^1a^2\ldots a^n}(x_m)$ for the sequence extracted from the first m terms of x by the selection $C_{a^1a^2\ldots a^n}$, then*

- *each of the first m terms of x belongs to exactly one sequence $C_{a^1a^2\ldots a^n}(x_m)$, and*
- *if the sequence $C_{a^1a^2\ldots a^n}(x_m)$ contains at least one term, then each of the sequences $C_{a^1a^2\ldots a^i}(x_m)$, $(i < n)$, contains exactly m_i terms ($a^1a^2\ldots a^i$ being an initial segment of $a^1a^2\ldots a^n$).*

3° *Construction of a particular sequence x.* Starting with the selections B and C and the number p, we form a sequence x as follows.

First set $x^1 = 1$. Suppose, as an inductive hypothesis, that we have been able to determine $x^1, x^2, \ldots, x^m$ so that *in each nonempty $C_{a^1a^2\ldots a^n}(x_m)$, the number of terms equal to one equals or exceeds by no more than one the product of p and the number of terms in $C_{a^1a^2\ldots a^n}(x_m)$.* No matter what value we choose for x^{m+1}, we can determine which sequence $C_{a^1a^2\ldots a^n}(x_{m+1})$ it will belong to. The subscripts of that sequence will be functions of $x^1, x^2, \ldots, x^m$. Write (m) for these subscripts, so that $C_{(m)}$ is the selection $C_{a^1a^2\ldots a^n}$ such that $C_{(m)}(x_{m+1})$ contains x^{m+1}. If $C_{(m)}(x_m)$ is empty, so that x^{m+1} is the first term of $C_{(m)}(x_{m+1})$, we set $x^{m+1} = 1$. If $C_{(m)}(x_m)$ contains μ terms, of which ν are ones, we have

$$p\mu \leq \nu < p\mu + 1$$

by the inductive hypothesis.

We then have

$$p(\mu+1) \leq \nu < p(\mu+1)+1 \qquad \text{or} \qquad p(\mu+1) \leq \nu+1 < p(\mu+1)+1$$

We set $x^{m+1} = 0$ in the first case, $x^{m+1} = 1$ in the second case. Then all the sequences $C_{a^1a^2\dots a^n}(x_{m+1})$ will have the same property as the sequences $C_{a^1a^2\dots a^n}(x_m)$.

The process of forming the sequence x continues indefinitely.

4° *The sequence x is a collective with respect to B.* For this section §4°, fix n and indices $a^1a^2\dots a^n$, and consider the selection $B_{a^1a^2\dots a^n}$ and the sequence it extracts from the first m terms of x, $B_{a^1a^2\dots a^n}(x_m)$. The terms of a sequence of the form $C_{a^1a^2\dots a^n\alpha^1\alpha^2\dots\alpha^s}(x_m)$ all belong to $B_{a^1a^2\dots a^n}(x_m)$; inversely, every term in the latter sequence belongs to one and only one sequence of that form.[3] Let μ be the number of terms of $B_{a^1a^2\dots a^n}(x_m)$ among which ν is equal to 1; let s_0 be the largest of the values of s for which there are α^i such that the sequence $C_{a^1a^2\dots a^n\alpha^1\alpha^2\dots\alpha^s}(x_m)$ is nonempty, and let $C_{a^1a^2\dots a^n\beta^1\beta^2\dots\beta^{s_0}}(x_m)$ be one such nonempty sequence. The sequences

$$C_{a^1a^2\dots a^n}(x_m), C_{a^1a^2\dots a^n\beta^1}(x_m), \dots, C_{a^1a^2\dots a^n\beta^1\beta^2\dots\beta^{s_0-1}}(x_m)$$

contain $m_n, m_{n+1}, \dots, m_{n+s_0-1}$ terms, respectively. The first relation,

$$\mu \geq m_n + m_{n+1} + \cdots + m_{n+s_0-1} \tag{25}$$

follows. Because every sequence $C_{a^1a^2\dots a^n\alpha^1\alpha^2\dots\alpha^{s_0+1}}(x_m)$ is empty, and there are 2^s sequences of the form $C_{a^1a^2\dots a^n\alpha^1\alpha^2\dots\alpha^s}(x_m)$, we conclude that

$$\mu \leq m_n + 2m_{n+1} + \cdots + 2^{n+s_0}m_{n+s_0}. \tag{26}$$

When $s = 0$, relations (25) and (26) become $\mu \geq 0$ and $\mu \leq m_n$, respectively.

Let r be the number of nonempty sequences $C_{a^1a^2\dots a^n\alpha^1\alpha^2\dots\alpha^s}(x_m)$. (We are still keeping the a^i fixed but allowing the α^j to vary.) From the definition of s_0, we conclude that

$$r \leq 1 + 2 + \cdots + 2^{s_0} < 2^{s_0+1}. \tag{27}$$

[3] *Translator's note:* The exposition now becomes a little more compressed. Towards the end of §2°, Ville had noted that any term in a C will be in all earlier Bs; this is evident from (23). He had not stated so explicitly that every term in a B will be in some later C, but this is also clear from (23), for the new α^i can always be chosen so that the term is in α^iA_i, keeping it in the corresponding B until it appears in a C.

If these nonempty sequences contain $\mu_1, \mu_2, \dots, \mu_r$ terms, respectively, of which $\nu_1, \nu_2, \dots, \nu_r$ are equal to one, we have

$$\begin{cases} \mu_1 + \mu_2 + \cdots + \mu_r = \mu, \\ \nu_1 + \nu_2 + \cdots + \nu_r = \nu. \end{cases} \tag{28}$$

On the other hand, the way the sequence x is constructed guarantees

$$p\mu_j \leq \nu_j < p\mu_j + 1 \qquad (j = 1, 2, \dots, r), \tag{29}$$

and so, by summation,

$$p\mu \leq \nu < p\mu + r < p\mu + 2^{s_0+1}. \tag{30}$$

We will then have, by (22) and (25),

$$\mu + \psi(n) > \psi(n + s_0).$$

This means, by the definition of ψ, that

$$2^{n+s_0} < \phi[\mu + \psi(n)] = \phi(\mu + \rho). \qquad [\rho = \phi(n)]$$

Taking (30) into account, with the selection $B_{a^1a^2\dots a^n}$ unchanged, we have

$$0 \leq \frac{\nu}{\mu} - p < \frac{\phi(\mu + \rho)}{\mu 2^{n-1}}, \tag{31}$$

So if $\lim_{m\to\infty} \mu = \infty$ while n and hence ρ remain fixed, we see that

$$\lim_{m\to\infty} \frac{\nu}{\mu} = p.$$

5° *The sequence x will satisfy the conclusions of Theorem 4′.* Let A_n be one of the selections in the system $\mathcal{S}$, and write $A_n(x_m)$ for the sequence extracted by A_n from the first m terms of x. The terms of $A_n(x_m)$ belong

a. either to the sequences $C_{a^1a^2\dots a^s}(x_m) \quad (s < n)$,

b. or to the sequences $B_{a^1a^2\dots a^{n-1}}$.

If $A_n(x)$, the sequence extracted from x by A_n, is infinite, we can neglect the terms of type (a), of which there are only a finite number, say N.

So we write

$$A_n = D + \sum B_{a^1 a^2 \dots a^{n-1}}, \tag{32}$$

where D is a selection that chooses N terms. The summation is over all values of the indices $a^1, a^2, \dots, a^{n-1}$ (Definition 4). Let μ_0 be the number of terms in $D(x_m)$. Number all the sequences of the form $B_{a^1 a^2 \dots a^{n-1}}(x_m)$ from 1 to s, where $s = 2^{n-1}$, and let $\mu_1, \dots, \mu_s$ be the numbers of terms contained in the corresponding $B_{a^1 a^2 \dots a^{n-1}}(x_m)$. Let $\nu_0, \nu_1, \dots, \nu_s$ be the numbers of terms equal to 1 in these respective sequences. If $A_n(x_m)$ contains μ terms, of which ν are equal to 1, we will have, taking (31) into account,

$$\begin{cases} \mu = \mu_0 + \mu_1 + \cdots + \mu_s & \qquad (\mu_0 = N) \\ \nu = \nu_0 + \nu_1 + \cdots + \nu_s & \qquad (0 \leq \nu_0 \leq \mu_0 = N) \end{cases}$$

$$\begin{cases} p\mu_i \leq \nu_i < p\mu_i + 2^{1-n}\phi(\mu_i + \rho) & \qquad [\rho = \psi(n)], (i = 1, 2, \dots, s), \\ pN \leq \nu_0 + pN \leq pN + N. \end{cases}$$

So, by summation,

$$p\mu \leq \nu + pN < p\mu + N + 2^{1-n} \sum_{i=1}^{s} \phi(\mu_i + \rho),$$

or, because ϕ is an increasing function

$$-\frac{pN}{\mu} \leq \frac{\nu}{\mu} - p < \frac{qN}{\mu} + \frac{\phi(\mu + \rho)}{\mu} \qquad (q = 1 - p). \tag{33}$$

This means that if $\lim_{m\to\infty} \mu = \infty$, then

$$\lim_{m\to\infty} \frac{\nu}{\mu} = p.$$

So x is indeed a collective of type $\mathbf{K}(\mathcal{S}, p)$. Moreover, we obtain (19) by substituting

$$\alpha = n \quad \text{and} \quad \beta = \limsup_{0<\mu<\infty} \frac{\phi(\mu + \rho)}{\phi(\mu)}$$

in (33). Theorem 4′ is thus demonstrated.

Remark. Consider the frequency of 1s in the sequence x that we have just constructed. Each of the first m terms of x belong to a sequence $C_{a^1 a^2 \ldots a^n}(x_m)$. But each of these sequences, by construction, contain a number of 1s equal or exceeding by no more than on the product of p and[4] the number of its terms. It follows that the frequency of 1s among the first m terms of x is always $\geq p$. *For any $\mathcal{S}$ and p, one can construct a collective in which the frequency tends to its limit unilaterally.*

[4] *Translator's note:* The original inadvertently omits the phrase "the product of p and".

A Ville's 1936 announcement

On the notion of a collective.
Note by Mr. Jean-André Ville, presented by Mr. Émile Borel.[5]

We propose to study the notion of irregularity of a sequence (*Regellosigkeit*) used by Mr. von Mises in his definition of a collective.[6]

1. Let us represent by a sequence of 0s and 1s the outcomes of an infinite sequence of independent trials, all with the same probability p.[7] To simplify the exposition, we assume that $p = 1/2$.

Mr. Wald defined a collective relative to a countably infinite system of selections.[8] We can associate a point of the line segment $(0, 1)$ with each collective. It is easily shown that for a given countably infinite system of selections, the set of points in $(0, 1)$ that do not represent collectives with respect to **S** has measure zero. We have shown that the converse if false: *there are sets of measure zero such that for any* **S***, there are points in the set that represent collectives relative to* **S**.

2. To obtain a theory without this asymmetry, we substitute the notion of a martingale for that of a selection. We continue to assume that $p = 1/2$. Consider a player who begins with unit capital and plans to play indefinitely, risking on each round a certain proportion of the amount he then has, a proportion that depends in a specified way on the sequence of outcomes already obtained. The martingale he follows can be defined unambiguously just as Mr. Wald defines a selection function. A sequence x is a collective with respect to a martingale **M** if the capital sequence the player gets by following the martingale has a finite upper bound in the course of play.[9]

With this notion: for any **M**, the set of points representing collectives relative to **M** has measure one, and conversely: *for any set of A of measure zero, there is a martingale* **M** *such that none of the points of A are collectives with respect to* **M**.

[5] *Comptes rendus*, **203**, pp. 26–27, 1936

[6] *Footnote in the original: Wahrscheinlichkeitsrechnung*, Leipzig und Wien, 1931.

[7] *Footnote in the original:* See my note in the *Comptes rendus*, **202**, 1936, p. 1393.

[8] *Footnote in the original: Comptes rendus*, **202**, 1936, p. 180.

[9] Whereas Wald considered a countably infinite set of selections, Ville makes do with a single martingale. This is possible because a countably infinite collection of martingales can be averaged to produce a single martingale. Because it diverges to infinity whenever any of the martingales being averaged do so, this single martingale is all we need.