ANNALES DE L'I. H. P., SECTION B

JEAN LEMAIRE

Une nouvelle définition de l'équitabilité pour les jeux de hasard. Application au paradoxe de Saint-Pétersbourg

Annales de l'I. H. P., section B, tome 9, nº 2 (1973), p. 193-204

<http://www.numdam.org/item?id=AIHPB_1973__9_2_193_0>

Ann. Inst. Henri Poincaré, Section B :
Vol. IX, n° 2, 1973, p. 193-204. *Calcul des Probabilités et Statistique.*

Une nouvelle définition de l'équitabilité pour les jeux de hasard. Application au paradoxe de Saint-Pétersbourg

par

Jean LEMAIRE

Assistant à l'Université de Bruxelles

PREMIÈRE PARTIE

UNE NOUVELLE DÉFINITION DE L'ÉQUITABILITÉ

RÉSUMÉ. — La considération de l'espérance mathématique d'un jeu comme mise équitable conduit à de nombreux paradoxes dont le plus célèbre est le jeu de Saint-Pétersbourg. Nous introduisons une espérance mathématique tronquée dépendant du nombre de parties à jouer et nous démontrons une condition suffisante d'équitabilité pour la mise ainsi définie. Plusieurs contre-exemples montrent que le théorème ne peut être étendu.

SUMMARY. — The introduction of the mathematical expectation as a fair stake for chance games leads to numerous paradoxes, such as the celebrated Petersburg game. We propose here to introduce a truncated mathematical expectation, depending on the number of plays. A sufficient condition for the equitability of this new stake is proved and several counter-examples show that the theorem cannot be improved.

1. INTRODUCTION ET ÉNONCÉ DU THÉORÈME PRINCIPAL

Considérons un jeu de hasard Γ opposant un joueur A au banquier B. Les deux adversaires se mettent d'accord pour jouer Γ n fois. La partie est composée de la succession des n jeux, joués de manière uniforme et indépendante. Une et une seule partie sera jouée.

Avant chaque jeu, A dépose une mise $m(n) \geq 0$. Les résultats du jeu constituent un ensemble exhaustif de k (fini ou dénombrable) événements, caractérisés par leur probabilité p_i et le gain $g_i \geq 0$ qui y est associé (en général, une classe comportera le gain $g_i = 0$). Posons $\mathrm{K} = \{ 1, \ldots, k \}$ et ordonnons les p_i par valeurs décroissantes.

L'espérance mathématique de gain du jeu vaut

$$\mathrm{E} = \sum_{i=1}^{k} p_i g_i.$$

La théorie classique des jeux définit un jeu équitable comme étant un jeu dont la mise $m = m(n)$ vaut $\sum_{\mathrm{K}} p_i g_i$.

Ce critère est fort critiquable, et de nombreux paradoxes (dont le célèbre jeu de Saint-Pétersbourg) démontrant sa faiblesse ont vu le jour (voir notamment [*1*, *2* et *3*]).

Les lois des grands nombres permettent d'affirmer que le gain moyen réel converge vers l'espérance mathématique lorsque le nombre de parties n augmente indéfiniment. En pratique n est fini, et il y aura toujours une différence entre les deux nombres Lévy. [*6*] a démontré le théorème suivant :

« Pour mesurer une probabilité avec une précision telle que l'erreur relative soit de l'ordre de $1/q$, il faut continuer les expériences jusqu'à ce que l'événement ait été réalisé environ q^2 fois ».

Donc, pour que la probabilité soit déterminée avec deux chiffres significatifs, il faut que l'événement ait été observé 10.000 fois environ. Pour n petit, les probabilités les plus faibles sont mal approchées par les fréquences relatives correspondantes. Comme ce sont, en général, ces événements rares qui conduisent aux plus grosses valeurs du gain la différence entre le gain observé et le gain théorique sera très sensible. Si l'on ne joue qu'un petit nombre de parties, le gain sera relativement faible. Au fur et à mesure que n augmente, des événements plus rares vont apparaître et le gain moyen va croître et tendre vers l'espérance mathématique du jeu.

Ceci conduit naturellement à la conclusion suivante : il ne faut considérer qu'une certaine partie de la distribution, dépendant de n. La fraction tronquée doit diminuer lorsque n augmente, pour disparaître à la limite.

Nous proposons de remplacer l'espérance mathématique par la fonction

$$E^*(n) = \sum_{\substack{i \in K \\ p_i \geq P_0(n)}} p_i g_i = f(n),$$

où $P_0(n)$ est une fonction uniforme dont les valeurs sont comprises entre 0 et 1.

Le critère de l'espérance mathématique correspond au cas $P_0(n) \equiv 0$.

Nous pouvons imposer les deux conditions suivantes à $P_0(n)$:

— elle doit être monotone non croissante ; en effet, lorsque le nombre de jeux n augmente, les fréquences tendent vers les probabilités théoriques ; il faut tenir compte d'une fraction plus importante de la distribution ;

— $P_0(\infty) = \lim_{n \to \infty} P_0(n) = 0$. En effet, en nous appuyant sur la loi des grands nombres, nous pouvons exiger de retrouver l'espérance mathématique classique à la limite.

Bien que le problème ne se pose que pour n naturel, nous supposons $f(n)$ et $P_0(n)$ définies pour tous les réels de l'intervalle $[1, \infty)$.

ÉTUDE DE $f(n)$

— $f(n)$ est une fonction non-négative ;

— c'est une fonction non-décroissante de n ;

— $f(n)$ est une fonction en escalier dont les points de discontinuité sont appelés les points critiques : ce sont les points n_x tels que

$$P_0(n_x) = p_x.$$

Nous rangeons ces points par ordre croissant et nous les notons $n_0, n_1, \ldots, n_x, \ldots, n_y$ (y infini ssi k l'est). En général $n_0 = 1$.

Lorsque ce n'est pas le cas (c'est-à-dire si $f(n)$ est encore nulle au point $n = 1$), nous posons $n_0 = 0$. Il y a un saut positif en chacun de ces points (sauf au point éventuel de gain nul, qui n'est pas considéré comme un point critique).

— $f(n)$ tend vers l'espérance mathématique du jeu. Lorsque k est fini, cette espérance est atteinte au point n_y. Sinon, elle n'est atteinte qu'asymptotiquement.

DÉFINITION 1. — La mise $m(n)$ adoptée par le joueur est toute fonction continue non-décroissante passant par les points $\{ (n_x, f(n_x)) ; x = 1, \ldots, y \}$.

Soit X_i la variable aléatoire représentant le gain au cours de la i^e répétition du jeu. Soit $S_n = X_1 + \ldots + X_n$. Les lois des grands nombres permettent d'affirmer que

$$P\left(\left|\frac{S_n}{nm} - 1\right| > \varepsilon\right) \to 0 \qquad \text{pour } n \to \infty.$$

Nous généralisons naturellement cette propriété par la définition suivante :

DÉFINITION 2. — Le jeu est équitable si et seulement si

$$P\left(\left|\frac{S_n}{nm(n)} - 1\right| > \varepsilon\right) \to 0 \qquad \text{pour } n \to \infty.$$

Lorsque k est fini, $m(n)$ est constante à partir du point n_y et nous retrouvons la définition classique.

THÉORÈME 1. — Lorsque k est infini, le jeu est équitable si $P_0(n)$ est de la forme $\frac{A}{n}$, où A est une constante de l'intervalle [0, 1].

2. DÉMONSTRATION DU THÉORÈME

Ce théorème est une conséquence immédiate des lemmes 1 à 5 et du fait que A doit faire partie de l'intervalle unitaire pour que $P_0(n)$ soit compris entre 0 et 1.

Il suffit évidemment de faire la démonstration en considérant $f(n)$ au lieu de $m(n)$.

DÉFINITION 3. — Le jeu satisfait à la condition A si et seulement si

$$A(n) = \circ(nf(n)),$$

où

$$A(n) = \max_{\substack{i \\ p_i \geq P_0(n)}} g_i.$$

LEMME 1. — Si le jeu satisfait à la condition A et si $P_0(n) = \circ\left(\frac{1}{n}\right)$, il est équitable.

Pour n fixé, considérons la variable aléatoire X'_p obtenue en ne tenant compte que des gains de probabilité supérieure ou égale à $P_0(n)$ (les autres gains étant annulés pour avoir une variable non-défective).

Soit X''_p la variable aléatoire où n'interviennent que les gains de probabilité inférieure à $P_0(n)$.

Évidemment

$$X_p = X'_p + X''_p.$$

Par définition

$$E(X'_p) = f(n).$$

D'autre part

$$\begin{aligned}\operatorname{Var}(X'_p) &\leq E(X'^2_p)\\ &= \sum_{p_i \geq P_0(n)} p_i g_i^2\\ &= \sum_{p_i \geq P_0(n)} p_i g_i \max_{p_i \geq P_0(n)} g_i\\ &= A(n).f(n).\end{aligned}$$

$X'_1 + \ldots + X'_n$ est donc une variable aléatoire de moyenne $nf(n)$ et de variance σ_n^2 inférieure ou égale à $nf(n)A(n)$.

En appliquant l'inégalité de Bienaymé-Tchebycheff, il vient

$$\begin{aligned}P\left\{\left|\frac{X'_1 + \ldots + X'_n}{nf(n)} - 1\right| > \varepsilon\right\} &= P\{\,|X'_1 + \ldots + X'_n - nf(n)| > \varepsilon nf(n)\,\}\\ &\leq \frac{n\sigma_n^2}{\varepsilon^2 n^2 f(n)^2}\\ &\leq \frac{A(n)}{\varepsilon^2 nf(n)} \to 0\end{aligned}$$

en vertu de la condition A.

Pour achever la démonstration, il faut prouver que les variables X''_p deviennent négligeables, c'est-à-dire que

$$P\{X''_1 + \ldots + X''_n > 0\} \to 0.$$

Puisque les p_i sont de somme égale à 1, nous avons pour n suffisamment grand

$$P\left\{X''_p > 0 = \sum_{p_i < P_0(n)} p_i < B.P_0(n)\right\},$$

où B est une constante positive.

Par conséquent

$$P(X''_1 + \ldots + X''_n > 0) < BnP_0(n) \to 0$$

par hypothèse.

LEMME 2. — Si le jeu satisfait à la condition A, il est équitable si $P_0(n)$ est de la forme $\frac{A}{n}$.

La propriété a été prouvée pour les $P_0(n)$ du type $\frac{A}{n^{1+\varepsilon}}$. Nous devons la démontrer pour $\varepsilon = 0$. Par commodité, posons $A = 1$.

Utilisons la propriété

$$P\{\,|c-d|>f\,\} \leq P\left\{|e-d|>\frac{f}{2}\right\} + P\left\{|c-e|>\frac{f}{2}\right\},$$

en posant

$$c = X_1 + \ldots + X_n,$$

$$d = n\sum_{p_i \geq \frac{1}{n}} p_i g_i,$$

$$e = n\sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i = n\sum_{p_i \geq \frac{1}{n}} p_i g_i + n\sum_{\frac{1}{n^{1+\varepsilon}} \leq p_i < \frac{1}{n}} p_i g_i,$$

$$f = \varepsilon e = \varepsilon n \sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i.$$

Il vient

$$P\left\{\left|X_1 + \ldots + X_n - n\sum_{p_i \geq \frac{1}{n}} p_i g_i\right| > \varepsilon n \sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i\right\}$$

$$\leq P\left\{\left|n \sum_{\frac{1}{n^{1+\varepsilon}} \leq p_i < \frac{1}{n}} p_i g_i\right| > \frac{\varepsilon n}{2} \sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i\right\}$$

$$+ P\left\{\left|X_1 + \ldots + X_n - n \sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i\right| > \frac{\varepsilon n}{2} \sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i\right\}.$$

Ce dernier terme tend vers 0 par application du lemme 1. D'autre part, en considérant la suite $\left\{ n_x + \frac{1}{2}\,;\ x = 1, \ldots, y \right\}$ et une suite décroissante de ε_x telle que

$$\varepsilon_x \leq \frac{1}{2}\{\,P_0(n_x) - P_0(n_{x+1})\,\}\,;$$

le premier terme du second membre est nul pour chaque terme de la suite, par passage à la limite, nous voyons donc que

$$P\left\{\left|X_1 + \ldots + X_n - n\sum_{p_i \geq \frac{1}{n}} p_i g_i\right| > \varepsilon n \sum_{p_i \geq \frac{1}{n^{1+\varepsilon}}} p_i g_i\right\} \rightarrow 0.$$

Pour la suite $\left\{ n_x + \frac{1}{2}, \varepsilon_x \right\}$, ce terme ne diffère pas de

$$P\left\{\left|X_1 + \ldots + X_n - n\sum_{p_i \geq \frac{1}{n}} p_i g_i\right| > \varepsilon n \sum_{p_i \geq \frac{1}{n}} p_i g_i\right\}.$$

Ce dernier terme tend vers zéro, ce qui démontre le lemme.

Il nous reste à étudier la condition A, qui exprime que les gains ne peuvent augmenter trop vite par rapport à $f(n)$.

En posant $a(n) = \dfrac{\log f(n)}{\log n}$, toute fonction $f(n)$ peut s'écrire sous la forme $f(n) = n^{a(n)}$.

LEMME 3. — $f(n)$ est de la forme $Cn^{a(n)}$ ($a(n)$ donné) si et seulement si

$$g_x = \frac{C\left(n_x^{a(n_x)} - n_{x-1}^{a(n_{x-1})}\right)}{P_0(n_x)}$$

pour tous les $x \in K$ tels que $P_0(1) > p_x$.

Premier cas : $n_0 = 1$.

a) Au point n_0 :

$$\begin{aligned}\sum_{p_i \geq P_0(1)} p_i g_i &= C \\ &= p_1 g_1 + p_2 g_2 + \ldots + p_j g_j,\end{aligned}$$

(= définition de j) ce qui détermine la valeur de C.

b) Au point n_1 :

$$\begin{aligned}\sum_{p_i \geq P_0(n_1)} p_i g_i &= Cn_1^{a(n_1)} \\ &= p_1 g_1 + \ldots + p_j g_j + p_{j+1} g_{j+1} \\ &= C + p_{j+1} g_{j+1}\end{aligned}$$

Donc $$g_{j+1} = \frac{C\left(n_1^{a(n_1)} - 1\right)}{p_{j+1}} = \frac{C\left(n_1^{a(n_1)} - n_0^{a(n_0)}\right)}{P_0(n_1)}$$

car $p_{j+1} = P_0(n_1)$.

c) Au point n_x :

$$\sum_{p_i \geq P_0(n_x)} p_i g_i = Cn_x^{a(n_x)}$$
$$= p_1 g_1 + \ldots + p_j g_j + p_{j+1} g_{j+1} + \ldots + p_{x-1} g_{x-1} + p_x g_x.$$

Or $p_k g_k = C\left(n_k^{a(n_k)} - n_{k-1}^{a(n_{k-1})}\right)$ par hypothèse de récurrence ($j < k < x$); donc

$$\begin{aligned} Cn_x^{a(n_x)} &= C + C\left(n_1^{a(n_1)} - 1\right) + C\left(n_2^{a(n_2)} - n_1^{a(n_1)}\right) + \ldots \\ &\qquad + C\left(n_{x-1}^{a(n_{x-1})} - n_{x-2}^{a(n_{x-2})}\right) + p_x g_x \\ &= Cn_{x-1}^{a(n_{x-1})} + p_x g_x, \end{aligned}$$

les autres termes se simplifiant 2 à 2. Donc

$$p_x g_x = C\left(n_x^{a(n_x)} - n_{x-1}^{a(n_{x-1})}\right)$$

et

$$\begin{aligned} g_x &= \frac{C\left(n_x^{a(n_x)} - n_{x-1}^{a(n_{x-1})}\right)}{p_x} \\ &= \frac{C\left(n_x^{a(n_x)} - n_{x-1}^{a(n_{x-1})}\right)}{P_0(n_x)}. \end{aligned}$$

Deuxième cas : $n_0 = 0$.

Démonstration similaire, on a dans ce cas

$$C = \frac{P_0(n_1) . g_1}{n_1^{a(n_1)}}.$$

Posons $P_1(n) = [P_0(n)]^{-1}$.

LEMME 4. — Quelle que soit la forme de $f(n)$, si $P_1(n)$ est du type Dn^b (où b est une constante positive), la condition A est vérifiée si $b < 1$.

En vertu du lemme précédent,

$$\begin{aligned} A(n) &= \max_{p_i \geq P_0(n)} g_i \\ &= \max_{p_i \geq P_0(n)} \frac{C\left(n_i^{a(n_i)} - n_{i-1}^{a(n_{i-1})}\right)}{P_0(n_i)} \\ &\approx \max_{\frac{1}{n^b} \leq \frac{1}{n_i^b}} \frac{n_i^{a(n_i)} - n_{i-1}^{a(n_{i-1})}}{\dfrac{1}{Dn_i^b}} \\ &\approx \max_{n_i \leq n} Dn_i^b\left(n_i^{a(n_i)} - n_{i-1}^{a(n_{i-1})}\right) \\ &\leq \max_{n_i \leq n} Dn_i^b n_i^{a(n_i)}, \end{aligned}$$

et

$$\begin{aligned} A(n)/nf(n) &\approx \max_{n_i \leq n} Dn_i^{b+a(n_i)}/n^{a(n_i)+1} \\ &\leq Dn^{b-1} \to 0 \qquad \text{si } b < 1. \end{aligned}$$

LEMME 5. — La condition A est satisfaite avec une probabilité 1 si $P_1(n)$ est linéaire.

La démonstration de ce lemme est assez similaire à celle du lemme 2. On emploie la même relation en posant cette fois

$$c = \max_{p_i \geq \frac{1}{n^{1-\varepsilon}}} g_i\,;$$

$$e = n \sum_{p_i \geq \frac{1}{n^{1-\varepsilon}}} p_i g_i\,;$$

$$d = n \sum_{p_i \geq \frac{1}{n}} p_i g_i\,;$$

$$f = \varepsilon d.$$

La fin de la démonstration s'appuie sur le fait qu'il existe toujours une suite $\{ n_x, \varepsilon_x \}$ telle que

$$\max_{p_i \geq \frac{1}{n^{1-\varepsilon}}} g_i = \max_{p_i \geq \frac{1}{n}} g_i.$$

3. QUELQUES CONTRE-EXEMPLES

Le théorème 1 présente une condition suffisante d'équitabilité. Nous allons montrer par des contre-exemples que ce résultat ne peut être étendu. Nous utilisons un théorème de Feller [*4*].

THÉORÈME 2. — Soit $a_n \to \infty$ une suite de nombres positifs. La condition nécessaire et suffisante pour qu'il existe une suite $\{ b_n \}$ telle que

$$P\{ | S_n - b_n | > \varepsilon a_n \} \to 0 \qquad n \to \infty$$

est que, pour tout $\delta > 0$, on ait simultanément

$$n \int_{|x| > \delta a_n} dF(x) \to 0 \quad \text{et} \quad a_n^{-2} . n \int_{|x| < a_n} x^2 dF(x) \to 0.$$

Les contre-exemples suivants montrent qu'il peut ne pas y avoir équitabilité lorsque l'on choisit un $P_0(n)$ tronquant plus vite ou moins vite que la fonction $\frac{A}{n}$.

1) *Jeu non-équitable lorsque* $P_1(n) = \sqrt{n}$

$$\begin{cases} p_i = 2^{-i} \\ g_i = 2^{2i} \end{cases} \quad i = 1, \ldots, \infty.$$

On vérifie que l'on a dans ce cas

$$f(n) = \sum_{p_i \geq \frac{1}{\sqrt{n}}} p_i g_i = \sum_{n \geq 2^{2i}} 2^i.$$

Donc $m(n) = 2\left(2^{\frac{\operatorname{Log}_2 n}{2}} - 1\right)$. Il vient

$$n \int_{|x| > \delta a_n} dF(x) = n \sum_{g_i > \delta a_n} 2^{-i} = n \sum_{g_i > 2^{\frac{\operatorname{Log}_2 n}{2}}} 2^{-i},$$

en posant $\delta = 1/2$ et $a_n = m(n)$.

Si a est un nombre entier, nous avons $\sum_{i \geq a} 2^{-i} = \frac{2}{2^a}$. En choisissant une suite croissante de n telle que $\frac{\operatorname{Log} n}{4}$ soit entier, nous avons

$$n \int_{|x| > \delta a_n} dF(x) = \frac{2n}{2^{\frac{\operatorname{Log} n}{4}}} > \frac{2n}{2^{\operatorname{Log} n}} = \frac{2n}{n} = 2.$$

Par conséquent la première condition nécessaire n'est pas vérifiée et il n'y a pas convergence. Remarquons que pour cet exemple la condition A est satisfaite. En effet

$$A(n) = \max_{\frac{1}{\sqrt{n}} < \frac{1}{2^i}} 2^{2i} = \max_{2^{2i} < n} 2^{2i} < n$$

et $\qquad A(n)/nf(n) < 1/f(n) \to 0 \qquad$ car $\qquad m(n) \to \infty.$

2) *Jeu non-équitable lorsque* $P_1(n) = n^2$

$$\begin{cases} p_i = 2^{-i} \\ g_i = 2^{\frac{3i-1}{2}} \end{cases} \quad i = 1, \ldots, \infty.$$

Dans ce cas $m(n) = \frac{\sqrt{2}}{2} n.$

Il vient

$$n\int_{|x|>\frac{\delta\sqrt{2}}{2}} dF(x) = n \sum_{2^{3i-1/2}>\frac{\delta\sqrt{2}}{2}n} 2^{-i}$$

$$= n \sum_{2^{3i}>n^2} 2^{-i} \qquad \text{(en posant } \delta = 1\text{)}$$

$$= n \sum_{i>\frac{2}{3}\mathrm{Log}_2 n} 2^{-i} \approx n \frac{2}{2^{\frac{2}{3}\mathrm{Log}_2 n}} = \frac{n}{n^{2/3}} \not\to 0.$$

La non-convergence provient du fait que la condition A n'est pas vérifiée. En effet

$$A(n) = \max_{p_i \geq \frac{1}{n^2}} g_i = \max_{n \geq p_i^{-1/2}} 2^{\frac{3i-1}{2}} \approx \max_{n \geq 2^{i/2}} 2^{3i/2} \approx n^3$$

et

$$A(n)/nf(n) \approx n^3/n^2 = n \not\to 0.$$

La deuxième partie est consacrée entièrement à l'application du théorème 1 au paradoxe de Saint-Pétersbourg ainsi qu'aux résultats d'une simulation de ce jeu. Cet exemple montre que la condition est suffisante mais non nécessaire.

4. EXTENSIONS

Le théorème a été présenté sous la forme d'un jeu de hasard à deux joueurs, mais le champ de ses applications est évidemment beaucoup plus large et s'étend à tous les processus stochastiques. La généralisation à des variables aléatoires non-équidistribuées (mais indépendantes) et pouvant prendre des valeurs non-négatives s'effectue sans difficultés majeures. La propriété est également vraie dans le cas de variables aléatoires continues, pour autant que le gain ne soit pas exponentiel. La démonstration, ainsi que l'utilisation du théorème en réassurance est donnée dans [5].

BIBLIOGRAPHIE

[1] D. Bernoulli, Specimen theoriæ novæ de mensura sortis. *Commentarii academiæ Scientiarum Imperialis Petropolitanæ*, vol. **5**, 1738, p. 175-193. Traduction par L. Sommer dans *Econometrica*, vol. **22**, 1954, p. 23-25.

[2] E. Borel, Le paradoxe de Saint-Pétersbourg. *C. R. Acad. Sci.* (Paris), vol. **229**, 1949, p. 406-407.

[3] E. Borel, Sur une propriété singulière de la limite d'une espérance mathématique. *C. R. Acad. Sci.* (Paris), vol. **229**, 1949, p. 429-431.

[4] W. Feller, *An introduction to probability theory and its applications*. New York, 1950, vol. **1**.

[5] J. Lemaire, Sur la détermination d'un contrat optimal de réassurance, *à paraître* au *Bulletin ASTIN*.

[6] P. Levy, *Calcul des probabilités*. Paris, 1925, p. 113-133.

(Manuscrit reçu le 30 janvier 1973).