CHAPITRE X.

DE L'ESPÉRANCE MORALE.

41. On a vu dans le n° 2 la différence qui existe entre l'espérance mathématique et l'espérance morale. L'espérance mathématique résultante de l'attente probable d'un ou de plusieurs biens, étant le produit de ces biens, par la probabilité de les obtenir, elle peut être évaluée par l'analyse exposée dans ce qui précède. L'espérance morale se règle sur mille circonstances qu'il est presque impossible de bien évaluer. Mais nous avons donné dans le numéro cité un principe qui, s'appliquant aux cas les plus communs, conduit à des résultats souvent utiles, et dont nous allons développer les principaux.

D'après ce principe, x étant la fortune physique d'un individu, l'accroissement dx qu'elle reçoit, produit à l'individu un bien moral réciproque à cette fortune; l'accroissement de sa fortune morale peut donc être exprimé par $\frac{k.dx}{x}$, k étant une constante. Ainsi en désignant par y la fortune morale correspondante à la fortune physique x, on aura

$$y = k.\log.x + \log.h,$$

h étant une constante arbitraire que l'on déterminera au moyen d'une valeur de y correspondante à une valeur donnée de x. Sur cela, nous observerons que l'on ne peut jamais supposer x et y nuls ou négatifs, dans l'ordre naturel des choses; car l'homme qui ne possède rien regarde son existence comme un bien moral qui peut être comparé à l'avantage que lui procurerait une fortune

physique dont il est bien difficile d'assigner la valeur, mais que l'on ne peut fixer au-dessous de ce qui lui serait rigoureusement nécessaire pour exister; car on conçoit qu'il ne consentirait point à recevoir une somme modique, telle que cent francs, avec la condition de ne prétendre à rien, lorsqu'il l'aurait dépensée.

Supposons maintenant que la fortune physique d'un individu soit a, et qu'il lui survienne l'expectative d'un des accroissements α, 6, γ, etc. ces quantités pouvant être nulles ou même négatives, ce qui change les accroissements en diminutions. Représentons par p, q, r, etc. les probabilités respectives de ces accroissements, la somme de ces probabilités étant supposée égale à l'unité. Les fortunes morales correspondantes de l'individu pourront être

$$k.\log.(a+\alpha)+\log.h, \quad k.\log.(a+6)+\log.h, \quad k.\log.(a+\gamma)+\log.h, \quad \text{etc.}$$

En multipliant ces fortunes respectivement par leurs probabilités p, q, r, etc. la somme de leurs produits sera la fortune morale de l'individu, en vertu de son expectative; en nommant donc Y cette fortune, on aura

$$Y = kp.\log.(a+\alpha) + kq.\log.(a+6) + kr.\log.(a+\gamma) + \text{etc.} + \log.h.$$

Soit X la fortune physique qui correspond à cette fortune morale, on aura

$$Y = k.\log.X + \log.h.$$

La comparaison de ces deux valeurs de Y donne

$$X = (a+\alpha)^p.(a+6)^q.(a+\gamma)^r.\text{ etc.}$$

Si l'on retranche la fortune primitive a, de cette valeur de X; la différence sera l'accroissement de la fortune physique qui procurerait à l'individu le même avantage moral qui résulte pour lui de son expectative. Cette différence est donc l'expression de cet avantage, au lieu que l'avantage mathématique a pour expression

$$p\alpha + q6 + r\gamma + \text{etc.}$$

De là résultent plusieurs conséquences importantes. L'une d'elles est que le jeu mathématiquement le plus égal est toujours désavantageux. En effet, si l'on désigne par a la fortune physique du joueur avant de commencer le jeu; par p, sa probabilité de gagner, et par μ sa mise; celle de son adversaire doit être, par l'égalité du jeu, $\frac{(1-p).\mu}{p}$; ainsi le joueur gagnant la partie, sa fortune physique devient $a+\frac{1-p}{p}.\mu$, et la probabilité de cela est p. S'il perd la partie, sa fortune physique devient $a-\mu$, et la probabilité de cela est $1-p$; en nommant donc X sa fortune physique, en vertu de son expectative, on aura, par ce qui précède,

$$X=\left(a+\frac{1-p}{p}.\mu\right)^p.(a-\mu)^{1-p};$$

or cette quantité est plus petite que a, c'est-à-dire que l'on a

$$\left(1+\frac{1-p}{p}.\frac{\mu}{a}\right)^p.\left(1-\frac{\mu}{a}\right)^{1-p}<1,$$

ou, en prenant les logarithmes hyperboliques,

$$p.\log.\left(1+\frac{1-p}{p}.\frac{\mu}{a}\right)+(1-p).\log.\left(1-\frac{\mu}{a}\right)<0.$$

Le premier membre de cette équation peut être mis sous la forme

$$\int(1-p).\frac{d\mu}{a}.\left(\frac{1}{1+\frac{1-p}{p}.\frac{\mu}{a}}-\frac{1}{1-\frac{\mu}{a}}\right),$$

quantité qui est évidemment négative.

Il résulte encore de l'analyse précédente, qu'il vaut mieux exposer sa fortune, par parties, à des dangers indépendants les uns des autres, que de l'exposer tout entière au même danger. Pour le faire voir, supposons qu'un négociant ayant à faire venir par mer une somme ε, l'expose sur un seul vaisseau, et que l'observation ait fait connaître la probabilité p de l'arrivée d'un vaisseau du même genre, dans le port; l'avantage mathématique du négociant, résultant de son expectative, sera $p\varepsilon$. Mais si l'on représente par l'unité sa fortune physique, indépendamment de son expectative; sa fortune morale sera, par ce qui précède,

$$kp.\log.(1+\varepsilon)+\log.h,$$

et son avantage moral sera, en vertu de son expectative,

$$(1+\varepsilon)^p-1,$$

quantité plus petite que $p\varepsilon$; car on a

$$(1+\varepsilon)^p<1+p\varepsilon,$$

puisque $\log.(1+\varepsilon)^p$ ou $p.\log.(1+\varepsilon)$ est moindre que $\log.(1+p\varepsilon)$, ce qui est évident, lorsqu'on met ces deux logarithmes sous les formes $\int\frac{p\,d\varepsilon}{1+\varepsilon}$, et $\int\frac{p\,d\varepsilon}{1+p\varepsilon}$.

Supposons maintenant que le négociant expose la somme ε par parties égales sur r vaisseaux. Sa fortune physique deviendra $1+\varepsilon$, si tous les vaisseaux arrivent, et la probabilité de cet événement est p^r. Si $r-1$ vaisseaux arrivent, la fortune physique du négociant devient $1+\frac{(r-1).\varepsilon}{r}$, et la probabilité de cet événement est $rp^{r-1}.(1-p)$. Si $r-2$ vaisseaux arrivent, la fortune physique du négociant devient $1+\frac{r-2}{r}.\varepsilon$, et la probabilité de cet événement est $\frac{r.\overline{r-1}}{1.2}.p^{r-2}.(1-p)^2$, et ainsi de suite; la fortune morale du négociant est donc, par ce qui précède,

$$k.\left\{\begin{array}{l}p^r.\log.(1+\varepsilon)+r.p^{r-1}.(1-p).\log.\left(1+\frac{r-1}{r}.\varepsilon\right)\\+\frac{r.\overline{r-1}}{1.2}.p^{r-2}.(1-p)^2.\log.\left(1+\frac{r-2}{r}.\varepsilon\right)+\text{etc.}\end{array}\right\}+\log.h;$$

expression que l'on peut mettre sous cette forme,

$$kp.\int d\varepsilon.\left[\frac{p^{r-1}}{1+\varepsilon}+\frac{\overline{r-1}.p^{r-2}.(1-p)}{1+\frac{r-1}{r}.\varepsilon}+\frac{\overline{r-1}.\overline{r-2}.p^{r-3}.(1-p)^2}{1.2.\left(1+\frac{r-2}{r}.\varepsilon\right)}+\text{etc.}\right]+\log.h.\quad(a)$$

Si l'on retranche de cette expression celle de la fortune morale du négociant, lorsqu'il expose la somme ε sur un seul vaisseau, et que l'on obtient en faisant $r=1$ dans la précédente, ce qui, abstraction faite de $\log.h$, réduit celle-ci à $kp\int\frac{d\varepsilon}{1+\varepsilon}$, qui est égal à

$$kp.\int d\varepsilon.\left\{\frac{p^{r-1}}{1+\varepsilon}+\frac{\overline{r-1}.p^{r-2}.(1-p)}{1+\varepsilon}+\frac{\overline{r-1}.\overline{r-2}.p^{r-3}.(1-p)^2}{1.2.(1+\varepsilon)}+\text{etc.}\right\},$$

la différence sera

$$kp.(1-p).\frac{r-1}{r}.\int\frac{\varepsilon\,d\varepsilon}{1+\varepsilon}.\left[\frac{p^{r-2}}{1+\frac{(r-1)}{r}.\varepsilon}+\frac{\overline{r-2}.p^{r-3}.(1-p)}{1+\frac{(r-2)}{r}.\varepsilon}+\text{etc.}\right];$$

cette différence étant positive, on voit qu'il y a moralement de l'avantage à partager la somme ε sur plusieurs vaisseaux. Cet avantage s'accroît à mesure que l'on augmente le nombre r des vaisseaux; et si ce nombre est très-grand, l'avantage moral devient à peu près égal à l'avantage mathématique.

Pour le faire voir, reprenons la formule (a), et donnons-lui cette forme,

$$kp.\int\int.dx.d\varepsilon.c^{-\left(1+\frac{\varepsilon}{r}\right).x}.\left(p.c^{-\frac{\varepsilon x}{r}}+1-p\right)^{r-1}+\log.h;\quad(a')$$

l'intégrale relative à x étant prise depuis x nul jusqu'à x infini. Dans cet intervalle, le coefficient de dx sous les signes $\int\int$, n'a ni *maximum* ni *minimum;* car sa différentielle, prise par rapport à x, est

$$-c^{-\left(1+\frac{\varepsilon}{r}\right).x}.dx.\left(p.c^{-\frac{\varepsilon x}{r}}+1-p\right)^{r-2}.\left[p.(1+\varepsilon).c^{-\frac{\varepsilon x}{r}}+(1-p).\left(1+\frac{\varepsilon}{r}\right)\right];$$

cette différentielle est constamment négative depuis $x=0$ jusqu'à x infini; ainsi le coefficient lui-même diminue constamment dans cet intervalle. C'est donc ici le cas de faire usage de la formule (A) du n° 22 du premier livre, pour avoir, par une approximation convergente, l'intégrale $\int y dx$, y étant égal à

$$c^{-\left(1+\frac{\varepsilon}{r}\right).x}.\left(p.c^{-\frac{\varepsilon x}{r}}+1-p\right)^{r-1}.$$

La quantité que nous avons nommée v dans le numéro cité devient alors

$$v=-\frac{y\,dx}{dy}=\frac{p.c^{-\frac{\varepsilon x}{r}}+1-p}{p.(1+\varepsilon).c^{-\frac{\varepsilon x}{r}}+(1-p).\left(1+\frac{\varepsilon}{r}\right)};$$

ce qui donne

$$U=\frac{1}{1+p\varepsilon+(1-p).\frac{\varepsilon}{r}},$$

$$\frac{dU}{dx}=\frac{p.(1-p).\varepsilon^2.\left(1-\frac{1}{r}\right)}{r.\left[1+p\varepsilon+(1-p).\frac{\varepsilon}{r}\right]^2},$$

etc.

U, $\frac{dU}{dx}$, etc. étant ce que deviennent v, $\frac{dv}{dx}$, etc. lorsque x est nul. Cela posé, la formule (A) citée donnera

$$\int dx.c^{-\left(1+\frac{\varepsilon}{r}\right)x}.\left(p.c^{-\frac{\varepsilon x}{r}}+1-p\right)^{r-1}$$
$$=\frac{1}{1+p\varepsilon+(1-p).\frac{\varepsilon}{r}}.\left\{1+\frac{p.(1-p).\varepsilon^2.\left(1-\frac{1}{r}\right)}{r.\left[1+p\varepsilon+(1-p).\frac{\varepsilon}{r}\right]}+\text{etc.}\right\}.$$

la formule (a') devient ainsi, à très-peu près, lorsque r est un grand nombre,

$$k.\int \frac{p\,d\varepsilon}{1+p\varepsilon} + \log. h,$$

ou

$$k \log.(1+p\varepsilon) + \log. h.$$

Maintenant, soit X la fortune physique correspondante à cette fortune morale; on a, par ce qui précède,

$$k. \log. X + \log. h,$$

pour la fortune morale correspondant à X; en comparant donc ces deux expressions, on aura

$$X = 1 + p\varepsilon.$$

Dans ce cas, l'avantage moral est $p\varepsilon$; il est donc égal à l'avantage mathématique.

Souvent l'avantage moral des individus est augmenté par le moyen des caisses d'assurance, en même temps que ces caisses produisent aux assureurs un bénéfice certain. Supposons, par exemple, qu'un négociant ait une partie ε de sa fortune sur un vaisseau dont la probabilité d'arrivée est p, et qu'il assure cette partie, en donnant une somme à la compagnie d'assurance. Pour l'égalité parfaite entre les sorts mathématiques de la compagnie et du négociant, celui-ci doit donner $(1-p).\varepsilon$ pour prix de l'assurance. En représentant par l'unité la fortune du négociant, indépendamment de son expectative ε, sa fortune morale sera, par ce qui précède,

$$kp. \log.(1+\varepsilon) + \log. h,$$

dans le cas où il n'assure pas; et dans le cas où il assure, elle sera

$$k. \log.(1+p\varepsilon) + \log. h;$$

or on a

$$\log.(1+p\varepsilon) > p. \log.(1+\varepsilon),$$

ou, ce qui revient au même,

$$\int \frac{p\,d\varepsilon}{1+p\varepsilon} > \int \frac{p\,d\varepsilon}{1+\varepsilon},$$

p étant moindre que l'unité; la fortune morale du négociant est donc augmentée au moyen de son assurance. Il peut ainsi faire à la compagnie d'assurance un sacrifice propre à subvenir aux frais de l'établissement et au bénéfice qu'elle doit faire. Si l'on nomme α ce sacrifice, c'est-à-dire, si l'on suppose que le négociant donne à la compagnie, pour prix de son assurance, la somme $(1-p).\varepsilon+\alpha$, on aura, dans le cas de l'égalité des fortunes morales, lorsque le négociant assure, et lorsqu'il n'assure point,

$$\log.(1-\alpha+p\varepsilon) = p. \log.(1+\varepsilon);$$

ce qui donne

$$\alpha = 1 + p\varepsilon - (1+\varepsilon)^p.$$

C'est tout ce que le négociant peut donner à la compagnie, sans désavantage moral; il aura donc un avantage moral, en faisant un sacrifice moindre que cette valeur de α, et, en même temps, la compagnie aura un bénéfice qui, comme on l'a vu, devient certain, quand ses relations sont très-nombreuses. On voit par là comment des établissements de ce genre, bien conçus et sagement administrés, peuvent s'assurer un bénéfice réel, en procurant des avantages aux personnes qui traitent avec eux : c'est en général le but de tous les échanges; mais ici, par une combinaison particulière, l'échange a lieu entre deux objets de même nature, dont l'un n'est que probable, tandis que l'autre est certain.

42. Le principe dont nous venons de faire usage pour calculer l'espérance morale a été proposé par Daniel Bernoulli, pour expliquer la différence entre le résultat du calcul des probabilités et

l'indication du sens commun, dans le problème suivant. Deux joueurs A et B jouent à *croix* et *pile*, avec la condition que A paye à B deux francs, si *croix* arrive au premier coup; quatre francs, s'il arrive au second coup; huit francs, s'il arrive au troisième coup, et ainsi de suite jusqu'au $n^{ième}$ coup. On demande ce que B doit donner à A, en commençant le jeu.

Il est visible que l'avantage de B, relatif au premier coup, est un franc; car il a $\frac{1}{2}$ de probabilité de gagner deux francs à ce coup. Son avantage relatif au second coup est pareillement un franc, car il a $\frac{1}{4}$ de probabilité de gagner quatre francs à ce coup, et ainsi de suite; en sorte que la somme de tous ses avantages relatifs aux n coups est n francs. Il doit donc, pour l'égalité mathématique du jeu, donner à A cette somme, qui devient infinie, si l'on suppose que le jeu continue à l'infini.

Cependant personne, à ce jeu, ne risquera avec prudence une somme même assez modique, telle que cent francs. Pour peu que l'on réfléchisse à cette espèce de contradiction entre le calcul, et ce qu'indique le sens commun, on voit facilement qu'elle tient à ce que si l'on suppose, par exemple, $n=50$, ce qui donne 2^{50} pour la somme que B peut espérer au cinquantième coup, cette somme immense ne produit point à B un avantage moral proportionnel à sa grandeur; de manière qu'il y a pour lui un désavantage moral à exposer un franc pour l'obtenir, avec la probabilité excessivement petite $\frac{1}{2^{50}}$ de réussir. Mais l'avantage moral que peut procurer une somme espérée dépend d'une infinité de circonstances propres à chaque individu, et qu'il est impossible d'évaluer. La seule considération générale que l'on puisse employer à cet égard est que plus on est riche, moins une somme très-petite peut être avantageuse, toutes choses égales d'ailleurs. Ainsi, la supposition la plus naturelle que l'on puisse faire est celle d'un avantage moral réciproque, au bien de la personne intéressée.

C'est à cela que se réduit le principe de Daniel Bernoulli, principe qui, comme on vient de le voir, fait coïncider les résultats du calcul avec les indications du sens commun, et qui donne le moyen d'apprécier avec quelque exactitude, ces indications toujours vagues. Son application au problème dont on vient de parler va en fournir un nouvel exemple.

Nommons a la fortune de B avant le jeu, et x ce qu'il donne au joueur A. Sa fortune devient $a-x+2$, si *croix* arrive au premier coup; elle devient $a-x+2^2$, si *croix* arrive au second coup, et ainsi de suite jusqu'au coup n, où elle devient $a-x+2^n$, si *croix* n'arrive qu'au coup $n^{ième}$. La fortune de B devient $a-x$, si *croix* n'arrive point dans les n coups, après lesquels la partie est supposée finir; mais la probabilité de ce dernier événement est $\frac{1}{2^n}$. En multipliant les logarithmes de ces diverses fortunes par leurs probabilités respectives et par k, on aura, par ce qui précède, la fortune morale de B, en vertu des conditions du jeu, égale à

$$\frac{1}{2}.k.\log.(a-x+2)+\frac{1}{2^2}.k.\log.(a-x+2^2)\ldots\ldots$$

$$\ldots+\frac{1}{2^n}.k.\log.(a-x+2^n)+\frac{1}{2^n}.k.\log.(a-x)+\log.h.$$

Mais avant le jeu, sa fortune morale était $k.\log.a+\log.h$; en égalant donc ces deux fortunes, pour que B conserve toujours la même fortune morale, et repassant des logarithmes aux nombres, on aura, $a-x$ étant supposé égal à a', et faisant $\frac{1}{a'}=\alpha$,

$$1+\alpha x=(1+2\alpha)^{\frac{1}{2}}.(1+2^2.\alpha)^{\frac{1}{2^2}}\ldots.(1+2^n.\alpha)^{\frac{1}{2^n}}; \qquad (o)$$

les facteurs $(1+2\alpha)^{\frac{1}{2}}$, $(1+2^2\alpha)^{\frac{1}{2^2}}$ vont en diminuant sans cesse, et leur limite est l'unité; car on a

$$(1+2^i.\alpha)^{\frac{1}{2^i}}>(1+2^{i+1}.\alpha)^{\frac{1}{2^{i+1}}}.$$

En effet, si l'on élève à la puissance 2^{i+1}, les deux membres de cette inégalité, elle devient

$$1+2^{i+1}.\alpha+2^{2i}.\alpha^2 > 1+2^{i+1}.\alpha,$$

et sous cette forme, l'inégalité devient évidente. De plus, le logarithme de $(1+2^i.\alpha)^{\frac{1}{2^i}}$ est égal à $\frac{i.\log.2}{2^i}+\frac{1}{2^i}.\log.\left(\alpha+\frac{1}{2^i}\right)$; et il est visible que cette fonction est nulle dans le cas de i infini, ce qui exige que dans ce cas, $(1+2^i.\alpha)^{\frac{1}{2^i}}$ soit l'unité.

Si l'on suppose n infini dans l'équation (o), on a le cas où la partie peut se prolonger à l'infini, ce qui est le cas le plus avantageux à B. a', et par conséquent α étant supposés connus, on prendra la somme des logarithmes tabulaires d'un assez grand nombre $i-1$, des premiers facteurs du second membre, pour que $2^i\alpha$ soit au moins égal à dix. La somme des logarithmes tabulaires des facteurs suivants, jusqu'à l'infini, sera à très-peu près égale à

$$\frac{\log.\alpha}{2^{i-1}}+\frac{(i+1).\log.2}{2^{i-1}}+\frac{0,4342945}{3\alpha.2^{i-2}}.$$

L'addition de ces deux sommes donnera le logarithme tabulaire de $a'+x$ ou de a. Ainsi, on aura pour une fortune physique a, supposée à B avant le jeu, la valeur de x qu'il doit donner à A au commencement du jeu, pour conserver la même fortune morale. En supposant, par exemple, a' égal à cent, on trouve $a=107^f,89$; d'où il suit que la fortune physique de B étant primitivement $107^f,89$, il ne doit alors risquer prudemment à ce jeu que $7^f,89$, au lieu de la somme infinie que le résultat du calcul indique, lorsqu'on fait abstraction de toutes considérations morales. Ayant ainsi la valeur de a relative à $a'=100$, il est facile d'en conclure de la manière suivante, sa valeur relative à $a'=200$; en effet, on a, dans ce dernier cas,

$$a=(200+2)^{\frac{1}{2}}.(200+2^2)^{\frac{1}{4}}.\text{etc.}=2.(100+1)^{\frac{1}{2}}.(100+2)^{\frac{1}{4}}.(100+4)^{\frac{1}{8}}.\text{ etc.}$$

Mais on vient de trouver

$$(100+2)^{\frac{1}{4}}.(100+4)^{\frac{1}{8}}.\text{ etc.}=(107,89)^{\frac{1}{2}};$$

donc

$$a=2.\sqrt{101.107,89}=208,78.$$

Ainsi la fortune physique de B étant primitivement 208,78, il ne peut risquer prudemment à ce jeu au delà de $8^f,78$.

43. Nous allons maintenant étendre le principe exposé ci-dessus, aux choses dont l'existence est éloignée et incertaine. Pour cela, considérons deux personnes A et B, qui veulent placer chacune, en viager, un capital q. Elles peuvent le faire séparément : elles peuvent s'associer et constituer une rente viagère sur leurs têtes, de manière que la rente soit reversible à celle qui survit à l'autre. Examinons quel est le parti le plus avantageux.

Supposons les deux personnes du même âge, et ayant la même fortune annuelle que nous représenterons par l'unité, indépendamment du capital qu'elles veulent placer. Soit β la rente viagère que ce capital leur produirait à chacune, si elles plaçaient leurs capitaux séparément, en sorte que leur fortune annuelle devienne $1+\beta$. Nous exprimerons, conformément au principe dont il s'agit, leur fortune morale annuelle correspondante, par $k.\log.(1+\beta)+\log.h$. Mais cette fortune n'aura lieu que probablement à la $x^{ième}$ année; ainsi en désignant par y_x la probabilité que A vivra à la fin de la $x^{ième}$ année, on doit multiplier sa fortune morale annuelle relative à cette année, par y_x; en ajoutant donc tous ces produits, leur somme, que nous désignerons par $[k.\log.(1+\beta)+\log.h].\Sigma.y_x$, sera ce que je nomme ici, *fortune morale viagère*.

Supposons maintenant que A et B placent la somme $2q$ de leurs capitaux sur leurs têtes, et que cela produise une rente viagère β', reversible au survivant. Tant que A et B vivront, chacun d'eux ne touchera que $\frac{1}{2}\beta'$ de rente viagère, et leur fortune morale annuelle sera $k.\log.(1+\frac{1}{2}\text{б}')+\log.h$. En la multipliant par la probabilité qu'ils vivront tous deux à la fin de l'année x, probabilité égale à $(y_x)^2$; la somme de ces produits pour toutes les valeurs de x sera la fortune morale viagère de A, relative à la supposition de leur existence simultanée; cette fortune est donc

$$\left[k.\log.\left(1+\frac{\text{б}'}{2}\right)+\log.h\right].\Sigma.(y_x)^2.$$

La probabilité que A existe seul à la fin de la $x^{\text{ième}}$ année est $y_x-(y_x)^2$; sa fortune morale viagère relative à son existence après la mort de B, qui rend sa fortune morale annuelle égale à $1+\text{б}'$, est donc

$$[k.\log.(1+\text{б}')+\log.h].\Sigma.[y_x-(y_x)^2].$$

La somme de ces deux fonctions,

$$k.\log.\left(1+\frac{\text{б}'}{2}\right).\Sigma.(y_x)^2+k.\log.(1+\text{б}').[\Sigma.y_x-\Sigma.(y_x)^2]+\log.h.\Sigma.y_x,$$

sera la fortune morale viagère de A dans l'hypothèse où A et B placent conjointement leurs capitaux.

Si l'on compare cette fortune à celle que nous venons de trouver dans le cas où ils placent séparément leurs capitaux, on voit qu'il y aura pour A de l'avantage ou du désavantage à placer conjointement, suivant que

$$\log.\left(1+\frac{\text{б}'}{2}\right).\Sigma.(y_x)^2+\log.(1+\text{б}').[\Sigma.y_x-\Sigma.(y_x)^2]$$

sera plus grand ou moindre que $\log.(1+\text{б}).\Sigma.y_x$. Pour le savoir, il faut déterminer le rapport $\text{б}'$ à б; or on a, par le n° 40,

$$q=\text{б}.\Sigma.p^x y_x,$$

$\frac{1-p}{p}$ étant l'intérêt annuel de l'argent: on a ensuite, par le même numéro,

$$2q=\text{б}'.\Sigma.p^x.[2y_x-(y_x)^2];$$

on a donc

$$\text{б}'=\frac{2\text{б}.\Sigma.p^x y_x}{\Sigma.p^x.[2y_x-(y_x)^2]}.$$

Les tables de mortalité donneront les valeurs de $\Sigma.y_x$, $\Sigma.(y_x)^2$, $\Sigma.p^x y_x$, $\Sigma.p^x(y_x)^2$; on pourra ainsi juger lequel des deux placements dont il s'agit est le plus avantageux.

Supposons б et $\text{б}'$ de très-petites fractions; la quantité $\log.(1+\text{б}).\Sigma.y_x$ devient à très-peu près $\text{б}.\Sigma.y_x$. La quantité

$$\log.\left(1+\frac{\text{б}'}{2}\right).\Sigma.(y_x)^2+\log.(1+\text{б}').[\Sigma.y_x-\Sigma.(y_x)^2]$$

devient

$$\frac{\text{б}'}{2}.[2.\Sigma.y_x-\Sigma.(y_x)^2],$$

et en substituant pour $\text{б}'$ sa valeur précédente, elle devient

$$\text{б}.\frac{[2.\Sigma.y_x-\Sigma.(y_x)^2].\Sigma.p^x y_x}{2\Sigma.p^x y_x-\Sigma.p^x(y_x)^2};$$

il y a donc de l'avantage à placer conjointement, si

$$[2\Sigma.y_x-\Sigma.(y_x)^2].\Sigma.p^x y_x$$

l'emporte sur

$$[2\Sigma.p^x y_x-\Sigma.p^x(y_x)^2].\Sigma.y_x,$$

ou si l'on a

$$\frac{\Sigma.p^x(y_x)^2}{\Sigma.p^x y_x}>\frac{\Sigma.(y_x)^2}{\Sigma.y_x};$$

c'est en effet ce qui a lieu généralement, p étant plus petit que l'unité.

L'avantage de placer conjointement les capitaux s'accroît par la considération que l'augmentation $\frac{\text{б}'}{2}$ de revenu arrive au survivant à un âge ordinairement avancé, dans lequel de plus grands besoins qui se font sentir la rendent beaucoup plus utile. Cet avantage s'accroît encore de toutes les affections qui peuvent attacher les deux individus l'un à l'autre, et qui leur font désirer le bien-être de celui qui doit survivre. Les établissements dans lesquels on peut ainsi placer ses capitaux, et, par un léger sacrifice de son revenu, assurer l'existence de sa famille pour un temps où l'on doit craindre de ne plus suffire à ses besoins, sont donc très-avantageux aux mœurs, en favorisant les plus doux penchants de la nature. Ils n'offrent point l'inconvénient que nous avons remarqué dans les jeux même les plus équitables, celui de rendre la perte plus sensible que le gain; puisqu'au contraire ils offrent les moyens d'échanger le superflu contre des ressources assurées dans l'avenir. Le Gouvernement doit donc encourager ces établissements, et les respecter dans ses vicissitudes; car les espérances qu'ils présentent, portant sur un avenir éloigné, ils ne peuvent prospérer qu'à l'abri de toute inquiétude sur leur durée.

CHAPITRE XI.

DE LA PROBABILITÉ DES TÉMOIGNAGES.

44. Je vais d'abord considérer un seul témoin. La probabilité de son témoignage se compose de sa véracité, de la possibilité de son erreur, et de la possibilité du fait en lui-même. Pour fixer les idées, concevons que l'on ait extrait un numéro d'une urne qui en renferme le nombre n, et qu'un témoin du tirage annonce que le n° i est sorti. L'événement observé est ici le témoin annonçant la sortie du n° i. Soit p la véracité du témoin, ou la probabilité qu'il ne cherche point à tromper; soit encore r la probabilité qu'il ne se trompe point.

Cela posé, on peut former les quatre hypothèses suivantes : ou le témoin ne trompe point et ne se trompe point; ou il ne trompe point et se trompe; ou il trompe et ne se trompe point; enfin, ou il trompe et se trompe à la fois. Voyons quelle est, *à priori*, dans chacune de ces hypothèses, la probabilité que le témoin annoncera la sortie du n° i.

Si le témoin ne trompe point et ne se trompe point, le n° i sera sorti; mais la probabilité de cette sortie est *à priori*, $\frac{1}{n}$; en la multipliant par la probabilité pr de l'hypothèse, on aura $\frac{pr}{n}$ pour la probabilité entière de l'événement observé, dans cette première hypothèse.

Si le témoin ne trompe point et se trompe, le n° i ne doit point être sorti, pour qu'il annonce sa sortie; la probabilité de cela est $\frac{n-1}{n}$. Mais l'erreur du témoin doit porter sur l'un des numéros du nombre cherché (voir fig. 3). Le triangle de Pascal est obtenu en plaçant à la ligne i et à la colonne j le coefficient du binôme $\binom{i}{j}$. Les entiers i, j peuvent, ici, être nuls, en posant, par convention, $0\,! = 1$.

1					
1	1				
1	2	1			
1	3	3	1		
1	4	6	4	1	
1	5	10	10	5	1

Fig. 3. – Le triangle de Pascal

Ce qui nous intéresse ici est que, dans ce calcul, apparaît pour la première fois *l'idée de gain probable, ou d'espérance de gain.* On affecte, en effet, aux joueurs *la part de la mise qu'ils peuvent espérer gagner,* compte tenu de leurs probabilités de gain respectives.

Chapitre III

ESPÉRANCE DE GAIN DANS UN JEU DE HASARD ; LOI DES GRANDS NOMBRES DE BERNOULLI

Nous avons vu apparaître, au chapitre II, dans le *problème de partage des mises* un cas particulier très simple de calcul d'*espérance de gain.* Dans un jeu de hasard plus général, dont les différentes possibilités sont de : gagner S_1 avec probabilité p_1, S_2 avec probabilité p_2, ..., S_k avec probabilité p_k, avec :

$$p_1 + \ldots + p_k = 1,$$

le gain obtenu étant symbolisé par X est un *nombre aléatoire* (on dira une *variable aléatoire*), dont on définira l'*espérance* (ou *espérance mathématique),* notée E(X), par la somme pondérée :

$$E(X) = p_1\ S_1 + \ldots + p_k\ S_k.$$

Ainsi, dans le problème du chevalier de Méré, le joueur (A) reçoit son *espérance de gain,* égale à :

$$0 \times (1 - P(A)) + S \times P(A) = S \times P(A).$$

La notion d'*espérance* a été initialement introduite par Huygens en 1657, dans son traité *De Ratiociniis in Aleae Ludo (De la logique du jeu de dé).* Le nom d'*espérance* y apparaît en latin sous le nom de « *expectatio* », avec l'interprétation d'être « le juste prix auquel un joueur accepterait de céder sa place dans une partie ».

L'interprétation de Huygens est très suggestive pour comprendre la notion d'espérance, mais elle recèle bien des pièges. En fait, il vaut mieux considérer l'*espérance* comme un opérateur mathématique, plutôt que comme le calcul, obtenu logiquement à partir de l'idée empirique de la « valeur estimée » ou « escomptée » d'un gain. Pour ces raisons, l'espérance, telle que nous l'avons définie, et par opposition à d'autres définitions possibles, est appelée *espérance mathématique*.

Il est, en effet, très facile de construire des exemples où l'espérance mathématique ne coïncide pas avec la valeur intuitivement escomptée du gain. Considérons, par exemple, deux joueurs (A) et (B), jouant un jeu où (A) mise 1 million d'euros avec une probabilité de 1/1 000 de perdre, et (B) mise 1 000 €, avec une probabilité de 999/1 000 de perdre. Ces deux joueurs ont des espérances mathématiques de gain quasiment identiques (soit 1 000 € pour (B) et 999 € pour (A)). En termes d'espérance, la situation est donc sensiblement en faveur de (B). Il n'en demeure pas moins que la position de (A), qui est pratiquement assuré de gagner la partie (avec une probabilité de 99,9 %), paraît nettement plus favorable que celle de (B).

Un des exemples les plus célèbres de cas où la notion d'espérance s'adapte mal à l'idée de gain escomptée est le *paradoxe de Saint-Pétersbourg*, posé par Nicolas Bernoulli à Montmort en 1713 et repris par Daniel Bernoulli en 1738. Ce paradoxe a ensuite été étudié par Laplace (1749-1827), Poisson (1781-1840), Bertrand (1822-1900), Borel (1871-1956), et plus récemment, par Feller (1950) et Martin-Löf (1985) (*J. Appl. Probab., 22,* 634-646). Le jeu de Saint-Pétersbourg se joue entre deux adversaires (A) et (B). Le joueur (A) reçoit initialement une somme forfaitaire de S unités de compte, tandis que le gain de (B) dépend du résultat d'une suite de jeux de pile ou face : si (B) gagne les n premières parties mais perd la $(n + 1)$-ième, il reçoit 2^n unités de compte. Ainsi (B) misant sur pile et (A) sur face, si les jeux successifs sont :

F	B reçoit	1 unité,	A reçoit S ;
PF	B reçoit	2 unités,	A reçoit S ;
PPF	B reçoit	4 unités,	A reçoit S ;
PPPF	B reçoit	8 unités,	A reçoit S ;
PPPPF	B reçoit	16 unités,	A reçoit S.

...

Le problème est alors de *déterminer le forfait* S *de* (A) *pour que le jeu soit équitable.*

La solution logique voudrait que les espérances de gain des deux joueurs soient rendues égales. Or, ceci est impossible

puisque *l'espérance du gain de* (B) *est infinie* (car égale à $1 \times (1/2) + 2 \times (1/4) + \ldots + 2^n \times 2^{-n-1} + \ldots$). *Le forfait* S *de* (A) *devrait donc être infini pour que le jeu soit équilibré.*

Ceci n'est cependant pas proprement *équitable*. Par exemple si on fixe le forfait de (A) à $S = 1\,024 = 2^{10}$ (B) ne peut gagner une somme supérieure au forfait S de (A) qu'avec une probabilité de :

$$2^{-10} = 1/1\,024,$$

ce qui fait que (A) est pratiquement assuré de réaliser un gain meilleur que celui de (B), si la partie ne se déroule qu'une seule fois. Il y a donc bien, en apparence, un paradoxe, puisque la solution logiquement équitable favorise à l'excès le joueur (A) par rapport au joueur (B).

Une façon très simple de résoudre ce « paradoxe » serait de dire qu'il est tout bonnement impossible de rendre le jeu mathématiquement équitable. Ce n'est pourtant pas satisfaisant, car, partant du principe que toute chose a un prix, on se doute bien qu'il doit y avoir un autre moyen d'évaluer le forfait S que par un calcul d'espérance.

La meilleure solution « humaine » du paradoxe de Saint-Pétersbourg a, sans doute, été donnée, dès 1738, par Daniel Bernoulli (dans son traité : *Specimen Theoriae Novae de Mensura Sortis : Exposé d'une théorie nouvelle de l'évaluation du risque*). Daniel Bernoulli part du principe que l'*utilité du gain d'un joueur n'a de sens que rapportée à sa fortune*. Ainsi un joueur disposant d'une fortune α apprécie un gain $d\alpha$ proportionnellement au rapport $(d\alpha)/\alpha$ et non, proportionnellement à $d\alpha$.

Ainsi, si un joueur (C) dispose d'une fortune initiale α et se livre à un jeu où son gain (ou sa perte comptée négativement) est un nombre aléatoire X, Daniel Bernoulli définit *l'utilité moyenne* du jeu par l'expression (où log u désigne le logarithme néperien (en base e) de u) :

$$U(\alpha, X) = \alpha E(\log((X + \alpha)/\alpha)),$$

et où E(.) désigne l'espérance mathématique. Lorsque α devient très grand, $\alpha \log((X + \alpha)/\alpha)$ tend vers X, l'utilité moyenne $U(\alpha, X)$ devient alors voisine de l'espérance mathématique.

Prolongeant ce raisonnement, Daniel Bernoulli définit *l'espérance morale* $E_M(X, \alpha)$ du jeu, comme la valeur définie par l'identité :

$$\alpha \log((E_M(X, \alpha) + \alpha)/\alpha) = U(\alpha, X) = \alpha E(\log(X + \alpha)/\alpha).$$

Dans le cas où $X = S_1$ avec probabilité p_1, $X = S_2$ avec probabilité p_2, ..., et ainsi de suite, ces formules mènent à définir l'espérance morale du jeu par la formule :

$$E_M(X, A) = \left\{\alpha\left(1+\frac{S_1}{\alpha}\right)^{p_1} \times \ldots \times \left(1+\frac{S_k}{\alpha}\right)^{p_k} \times \ldots\right\} - \alpha.$$

Dans le jeu de Saint-Pétersbourg, le joueur (B), disposant d'une fortune β a une espérance morale égale à :

$$E_M(X, B) = P(\beta) - \beta$$
$$= \{(\beta + 1)^{1/2} \times (\beta + 2)^{1/4} \times \ldots \times (\beta + 2^p)^{2^{-p}} \times \ldots) - \beta.$$

Pour rendre le jeu équitable, en offrant à (A) un forfait égal à l'espérance morale de (B), il faudrait donc que ce forfait s'établisse à $P(\beta) - \beta$ unités. Par exemple, pour $\beta = 0$, $P(\beta) - \beta = 2$, pour $\beta = 10$, $P(\beta) - \beta \sim 3{,}04$, pour $\beta = 100$, $P(\beta) - \beta \sim 4{,}38$, pour $\beta = 1\,000$, $P(\beta) - \beta \sim 5{,}97$.

Malheureusement, si le raisonnement paraît excellent pour (B), il semble beaucoup moins en faveur de (A), puisqu'on constate que, si la fortune α de (A) est finie, sa ruine est possible avec une probabilité non nulle, alors que son espérance morale se trouve être infinie négative. On peut ainsi reprocher à la notion d'espérance morale, *qui ne respecte plus la symétrie entre les joueurs*. En fait, la notion d'espérance morale, aujourd'hui, est plus ou moins abandonnée au profit de celle de l'espérance mathématique, malgré son intérêt réel socio-économique.

Il existe d'autres approches, plus modernes, du jeu de Saint-Pétersbourg que nous évoquons brièvement. Au cours de n parties identiques, le gain S_n de (B) vérifie la loi limite $S_n/(n \text{ Log } n) \to 1$ (en probabilité lorsque $n \to \infty$). Ici, $\text{Log } n = (\log n)/\log 2$ désigne le logarithme en base 2 de n (voir W. Feller (1950)). On peut alors se servir de cette propriété pour obtenir une autre évaluation du forfait de (B) (voir S. Csörgö et R. Dodunekova (1991), *Sums, Trimmed Sums and Extremes*, Bâle, Birkhäuser).

Des « paradoxes » comme celui du jeu de Saint-Pétersbourg ne sont, en fait, paradoxaux, que dans la mesure où ils mélangent, parfois de manière inextricable, des notions mathématiques rigoureuses avec des interprétations physiques hasardeuses.

Daniel Bernoulli (1700-1782) fut membre d'une famille de savants dont les plus connus ont été, outre lui-même, son père Jean Bernoulli (1667-1748) (inventeur de la lemniscate), et son oncle Jacques Bernoulli (1645-1705) (inventeur des nombres de Bernoulli).

Une étape majeure fut franchie, dans le cours de l'évolution du calcul des probabilités, par la publication posthume en 1713 de l'œuvre de Jacques Bernoulli, *Ars Conjectandi*, ou *L'art de la conjecture*. Dans ce livre était exposée la première démonstration rigoureuse de la *loi faible des grands nombres*. On y considère une suite $X_1, \ldots, X_n, \ldots$, de variables aléatoires indépendantes, pouvant chacune prendre la valeur 1 avec probabilité p, et la valeur 0 avec probabilité $1 - p$. La somme partielle de ces observations,

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

s'interprète comme le nombre de fois que X_i prend la valeur 1 lorsque i varie de 1 à n. La loi faible des grands nombres exprime alors que, pour tout nombre $\varepsilon > 0$ fixé, la probabilité $P(|S_n/n - p| > \varepsilon)$ tend vers 0 lorsque n tend vers l'infini.

Cette loi exprime que la fréquence S_n/n d'apparition d'un événement dans une répétition d'épreuves tend vers la probabilité p d'observer cet événement dans une seule parmi ces épreuves.

On remarquera que cette version de la loi des grands nombres est plus faible que celle que nous avions évoquée, exprimant que, sous les hypothèses précédentes, $S_n/n \to p$ avec probabilité 1. Cette dernière version est appelée *loi forte des grands nombres*, tandis que celle de Bernoulli est connue sous le nom de *loi faible des grands nombres*.

Une des preuves du caractère non trivial de ces deux résultats peut être trouvée dans le fait qu'il s'est écoulé, environ, 215 ans entre la démonstration des deux lois, *faible* et *forte* des grands nombres. La loi forte des grands nombres n'a été, en effet, prouvée, qu'en 1930, sous sa forme finale, par A. N. Kolmogorov (1903-1987).

La loi des grands nombres de J. Bernoulli ne fournit pas de renseignement précis sur la vitesse de convergence de $P(|S_n/n - p| > \varepsilon)$ vers 0. Ce problème n'a été, en fait, résolu qu'en 1718 par Abraham de Moivre (1667-1754) dans son traité : *La Doctrine des Chances*, et repris, en 1730, dans son ouvrage : *Miscellanea Analytica de Seriebus et Quadraturis (Analyses Diverses de Séries et de Quadratures)*. Les démonstrations de A. de Moivre furent reprises et quelque peu améliorées par Pierre-Simon de Laplace (1749-1827) dans son traité : *Théorie Analytique des Probabilités*, publié en 1812, où il omettait de faire référence à de Moivre (1667-1754). Ceci fait que les résultats correspondants sont, le plus souvent attribués à Laplace, au lieu de l'être à leur inventeur initial.

De Moivre avait, dans sa démonstration, utilisé un développement asymptotique de $n! = n(n-1)\dots 2\cdot 1$ pour évaluer la probabilité $P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Compte tenu de sa formule, il pouvait déduire la première version du *théorème central limite* prouvant que, pour tout $-\infty < a < b < \infty$ fixés (avec $p \neq 0$ ou 1) :

$$\lim_{n\to\infty} P\left(a < \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\, dt.$$

Le résultat obtenu ainsi par de Moivre-Laplace implique la loi des grands nombres de J. Bernoulli et précise la loi de probabilité limite du rapport $(S_n - np)/\sqrt{np(1-p)}$.

On observe, ici, la fameuse *loi de Gauss*, ou *loi normale*, ou *loi de Laplace-Gauss*. Cette loi, correspondant à une densité communément appelée « courbe en cloche » (fig. 4) est, communément, attribuée à Gauss plutôt qu'à Laplace, car Gauss (1777-1855) prouva en 1821 (dans son traité : *Theoria Combinationis Observationum Erroribus Minimus Obnoxiae*, Théorie de la combinaison d'erreurs de faible amplitude) qu'elle était, dans un cadre plus général, la limite de la loi de probabilité d'une somme d'erreurs aléatoires indépendantes de faible amplitude, lorsque leur nombre devenait infiniment grand. Il apparaissait ainsi que les résultats dû à de Moivre et à Laplace n'étaient qu'un cas particulier d'une convergence plus générale, où la *loi normale* jouait un rôle universel.

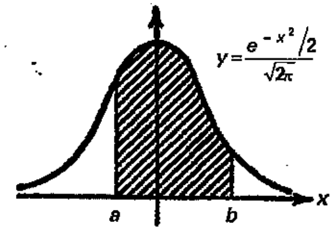

Fig. 4. – La courbe de Gauss (en hachuré $\int_a^b y dx$)

La loi des grands nombres de J. Bernoulli, précisée ci-dessus par le *théorème central limite*, exprime la convergence dite *faible* ou *en probabilité* de la suite S_n/n vers la valeur commune de l'espérance $E(X_i) = p$, pour $i = 1, 2, \dots$, des variables de la sommation.

Considérons plus généralement, au lieu de variables aléatoires X_i prenant les valeurs 1 ou 0, des variables X_i prenant la valeur x_1 avec probabilité $p_1, \dots, x_k$ avec probabilité p_k, où $p_1, \dots, p_k$ sont des nombres positifs ou nuls, tels que $p_1 + \dots + p_k = 1$.

En traitant, successivement, les fréquences des observations de $x_1, \dots, x_k$, il est possible de déduire de la loi des grands nombres de J. Bernoulli, précédemment énoncée pour le jeu de « pile ou face », qu'on a, plus généralement :

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E(X)\right| > \varepsilon\right) = 0, \text{ pour tout } \varepsilon > 0,$$

où $E(X) = p_1 x_1 + \dots + p_k x_k$ désigne l'*espérance mathématique* de la variable aléatoire $X := X_1$ (on note $A := B$, lorsque le membre de gauche A est défini par le membre de droite B).

Ce résultat n'est encore qu'un cas particulier, énoncé sous une forme affaiblie, de la *loi forte des grands nombres* due à Kolmogorov (1930) (A. N. Kolmogorov (1930), *Sur la loi forte des grands nombres*, Comptes rendus de l'Académie des sciences, Paris, 191, 910-912), dont nous présentons maintenant l'énoncé le plus général.

Soit $X_1, X_2, \dots$, *une suite infinie de variables aléatoires réelles indépendantes et de même loi de probabilité, définie par sa fonc-*

tion de répartition $F(x) = P(X_1 \leq x)$. *Alors, la limite de la moyenne d'ordre n, définie par*

$$\ell = \lim_{n\to\infty} \left\{ \frac{1}{n} \sum_{i=1}^{n} X_i \right\},$$

existe avec probabilité 1, si et seulement si l'espérance de la valeur absolue de $X := X_1$, *soit*

$$E(|X|) = \int_{-\infty}^{\infty} |x| dF(x),$$

est finie. Dans ce cas, on a, nécessairement,

$$\ell = \lim_{n\to\infty} \left\{ \frac{1}{n} \sum_{i=1}^{n} X_i \right\} = E(X) = \int_{-\infty}^{\infty} x dF(x).$$

L'expression ci-dessus, qui définit l'espérance mathématique d'une variable aléatoire X, généralise les exemples simples décrits antérieurement, en lui donnant un sens, en tant qu'*intégrale de Lebesgue-Stieltjes* (notions dues à Henri-Léon Lebesgue (1875-1941), et Thomas-Joannes Stieltjes (1856-1894)). Une façon explicite de définir une telle intégrale est donnée par la formule :

$$E(X) = \int_{-\infty}^{\infty} x dF(x) = \lim_{\substack{M\to\infty \\ N\to\infty}} \left\{ \lim_{n\to\infty} \sum_{k=-nM}^{nN} \frac{k}{n} (F((k+1)/n) - F(k/n)) \right\}.$$

Il ne s'agit que d'une généralisation aux variables aléatoires quelconques de la définition de l'espérance pour les variables aléatoires prenant leurs valeurs dans un ensemble dénombrable $\{x_i, i = 1, 2, \ldots\}$, et qui aurait été définie dans ce cas par :

$$E(X) = \sum_{i \geq 1} x_i P(X = x_i).$$

L'espérance mathématique d'une variable aléatoire X *est donc la limite avec probabilité 1 de moyennes de répliques indépendantes de même loi que* X, *lorsque leur nombre tend vers l'infini.*

L'énoncé, ci-dessus, de loi des grands nombres permet de comprendre les interprétations erronées dont a fait l'objet la notion d'espérance mathématique. Dans un jeu, l'espérance du gain n'est rien d'autre que la limite de la moyenne des gains observés « sur un grand nombre de parties ». Ce que signifie « un grand nombre de parties » est évidemment le point essentiel. Dans l'exemple que nous avions évoqué, d'un jeu où (A) mise

1 million d'euros avec probabilité de perdre 1/1 000, et (B) mise 1 000 €, avec probabilité de perdre 999/1 000, l'espérance des gains de (A) est 999 €, tandis que celle de (B) est de 1 000 €. La situation est donc, « à très long terme », favorable à (B), puisque, en supposant qu'on a répété un très grand nombre de parties identiques, (B) serait certain de gagner ; cependant, il est bien évident que, si on se limite à un petit nombre de parties, la situation est nettement plus en faveur de (A).

Il serait injuste, au sujet de tels résultats, de ne pas mentionner Siméon Denis Poisson (1781-1840) qui fut l'auteur du nom de « loi des grands nombres », et qui en obtint dans son traité, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile* (1837), une version très remarquable, citée ci-dessous ;

Au lieu de considérer une suite X_n de variables prenant les valeurs 1 ou 0 avec probabilités p et $1-p$ (on qualifie souvent d'ailleurs la loi de telles variables du nom de loi de Bernoulli), Poisson considéra des variables indépendantes, mais *de lois pas nécessairement identiques*, telles que, pour $n = 1, 2, \ldots,$

$$P(X_n = 1) = p_n \text{ et } P(X_n = 0) = 1 - p_n.$$

Poisson prouva alors que la conclusion de la loi faible des grands nombres subsistait en remplaçant, dans son énoncé, p par la moyenne $\bar{p}_n = (p1 + \ldots + p_n)$ des « probabilités de succès ». En d'autres termes, la version de Poisson de la loi des grands nombres montre que, pour tout $\varepsilon > 0$, on a :

$$\lim_{n\to\infty} P\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \bar{p}_n \right| > \varepsilon \right) = 0, \text{ pour tout } \varepsilon > 0.$$

On pourrait croire naïvement que le fait de rendre *hétérogène* la distribution des éléments de la suite est apte à rendre inexacte la loi des grands nombres. En fait, il n'en est rien, et, aussi bizarre que cela puisse paraître, c'est le contraire qui se passe. Cette propriété est rendue explicite par un résultat de Wassily Hoeffding (W. Hoeffding (1956), *Annals of Mathematical Statistics, 27,* 713-721) montrant que la probabilité de déviation :

$$P\left(\frac{1}{n} \sum_{i=1}^{n} X_i - \bar{p}_n > \varepsilon \right), \text{ pour } \varepsilon > 1/n,$$

est *maximale*, pour $\bar{p}_n$, fixé, lorsque les p_i sont égaux. En faisant jouer des rôles symétriques aux variables, et en changeant X_i en $1 - X_i$, on obtiendrait de même que :

$$P\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \bar{p}_n < \varepsilon\right), \text{ pour } \varepsilon > 1/n,$$

est *maximale*, pour fixé, lorsque les p_i sont égaux.

En d'autres termes *une fréquence de succès est d'autant plus concentrée autour de son espérance que les probabilités de succès des différentes variables (indépendantes) observées sont hétérogènes.*

Cette propriété inattendue est contraire à l'intuition. Elle montre la puissance de la loi des grands nombres, dont la validité dépasse largement le cas de jeux homogènes. Cette loi est donc peu affectée lorsque l'on y traite de fréquences de phénomènes ayant varié durant l'expérience.

Il importe, au passage, de noter que l'espérance d'une variable aléatoire, *n'est pas toujours définie.* En termes de lois des grands nombres, cette remarque a les conséquences suivantes :

1 / *Si l'espérance d'une variable est infinie, alors la moyenne d'une suite d'événements répétés indépendants converge vers l'infini avec probabilité 1.*

2 / *Si l'espérance d'une variable n'est pas définie (en excluant les cas où elle est infinie), alors la moyenne d'une suite d'événements répétés indépendants ne converge pas vers une limite, lorsque leur nombre croît, et oscille indéfiniment avec probabilité 1.*

Donnons quelques exemples :

1 / *Une martingale classique :* on considère deux joueurs (A) et (B) dans un jeu de pile ou face déséquilibré : (A) gagne avec probabilité $p \leq 1/2$, et (B) gagne avec probabilité $1 - p \geq 1/2$. Ceci se produit lorsque, par exemple (A) est un joueur misant sur pair (ou impair), et son adversaire (B) est le casino. Un « système » permettant, en apparence, d'assurer des gains certains à (A) consiste, pour lui, à miser initialement 1, puis à doubler, à chaque fois sa mise en cas de perte, et en arrêtant finalement la partie au premier gain enregistré (il s'agit de la *martingale, de Hawks*).

On peut schématiser les diverses possibilités du jeu dans le tableau page suivante.

(A) gagne au bout de	*Probabilité*	*Mise nécessaire*	*Gain*
1 jeu	p	1	1
2 jeux	$p(1-p)$	$3 = 1 + 2$	1
3 jeux	$p(1-p)^2$	$7 = 1 + 2 + 4$	1
4 jeux	$p(1-p)^3$	$15 = 1 + 2 + 4 + 8$	1
5 jeux	$p(1-p)^4$	$31 = 1 + 2 + 4 + 8 + 16$	1
:	:	:	:
n jeux	$p(1-p)^{n-1}$	$2^n - 1 = 1 + \ldots + 2^{n-1}$	1
:	:	:	:

On remarquera que la mise requise pour le joueur (A) croît extrêmement vite avec le nombre de jeux nécessaires avant qu'il puisse enregistrer un gain. De ce fait, l'utilisation effective de cette stratégie gagnante est rendue impossible dans la pratique. En effet, la partie s'arrête prématurément, soit parce que le casino impose des mises maximales et minimales, interdisant à (A) de poursuivre sa martingale au-delà d'un certain niveau d'enjeux, soit parce que (A) n'est plus en mesure de couvrir des mises excédant la fortune dont il dispose.

[Nous avons utilisé ici le terme « martingale » pour désigner, selon le vocabulaire des joueurs, un *système de jeu garantissant le gain.* Selon le vocabulaire mathématique, une *martingale* désigne une suite aléatoire X_n dont l'espérance conditionnelle, sachant les éléments précédents X_{n-1}, X_{n-2}, ..., de la suite, est égale à X_{n-1} ; si cette quantité était supérieure à X_{n-1}, on appellerait la suite *sous-martingale,* et inversement, *sur-martingale* si elle était inférieure à X_{n-1}. Ainsi, selon le vocabulaire scientifique, les gains ou pertes cumulés X_n d'un joueur (A) forment une martingale lorsque le jeu est équilibré (dans l'exemple précédent, $p = 1 - p = 1/2$), une sous-martingale si le jeu est en faveur de (A) (soit $p > 1/2$), et d'une surmartingale si le jeu est en défaveur de (A) (soit $p < 1/2$).]

L'espérance des mises cumulées nécessaires avant d'enregistrer un gain est, en effet, donnée par la formule

$$E(\text{Mise nécessaire}) = \sum_{n=1}^{\infty} (2^n - 1)\, p(1-p)^n.$$

Cette expression est *toujours infinie lorsque* $0 \leq p \leq 1/2$. Ainsi, *même pour un jeu équilibré* ($p = 1/2$), un joueur assez imprudent pour tenter, de manière répétée, de mettre en œuvre un tel système, verrait, par la loi des grands nombres, la moyenne des mises, requises pour continuer le jeu, croître vers l'infini, lui assurant, sur le long terme, une ruine certaine, indépendamment de l'importance initiale de sa fortune.

2 / *La loi de Cauchy :* nous avons décrit la façon logique selon laquelle la loi normale (ou de Laplace-Gauss-de Moivre) apparaissait dans la nature, comme loi de probabilité d'une grandeur aléatoire, s'interprétant comme la somme d'un très grand nombre de petits aléas indépendants. Ce résultat n'est cependant pas toujours vrai. Il suppose en effet que les « petits aléas » vérifient certaines conditions, comme, par exemple, d'être tels que leurs carrés soient d'espérance finie (cette dernière condition peut d'ailleurs encore être sensiblement affaiblie).

On peut se poser plus généralement le problème suivant : étant donné une suite $X_1, \ldots, X_n, \ldots$ de variables aléatoires indépendantes de même loi, quelles sont les distributions limites possibles de $(X_1 + \ldots + X_n - b_n)/a_n$, où $a_n > 0$ et b_n sont des coefficients non aléatoires convenablement choisis ? On reconnaît là une version générale du problème de Gauss, lorsque les X_i désignent les « petits aléas ».

Cette question a son importance, car les distributions ainsi obtenues pourront être considérées comme « naturelles », le procédé permettant de les construire se retrouvant dans toutes sortes de phénomènes physiques.

La solution de ce problème fut obtenue par Paul Lévy en 1937, par l'invention des *lois stables.* Son résultat montre que, si $E(X_i^2)$ est finie, la seule loi limite possible est la loi *normale,* tandis que, autrement, on obtient des lois, dites *stables,* d'expression plus complexe, et dont l'exemple le plus simple est la *loi de Cauchy.* Sous forme standard, X suit une loi de Cauchy, si la probabilité que X soit compris entre a et b est donnée par

$$P(a < X \leq b) = \int_a^b \frac{dx}{\pi(1+x^2)} = \frac{\operatorname{Arc\,tg} b - \operatorname{Arc\,tg} a}{\pi}.$$

Dans le cas général, l'expression de la densité d'une loi stable est fort complexe. Sous forme réduite (avec une échelle

convenablement choisie) la densité $f(t)$ d'une loi stable dépend de deux paramètres α et γ, vérifiant $0 \leq \alpha \leq 2$ ($\alpha = 2$ donne la loi normale, $\alpha = 1$ donne la loi de Cauchy comme cas particulier), et $|\gamma| \leq 1 - |1 - \alpha|$. On a, comme ci-dessus,

$$P(a < X \leq b) = \int_a^b f(t)\,dt,$$

où, posant $\Gamma(r) = \int_0^\infty t^{r-1} e^{-t}\,dt$ pour $r > 0$ (il s'agit de la *fonction gamma d'Euler,* vérifiant, entre autres, $\Gamma(n) = (n-1)!$ pour n entier), la densité $f(x)$ est donnée par :

– pour $0 < \alpha < 1$:

$$f(x) = \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}\,\Gamma(\alpha k+1)}{k!} |x|^{-\alpha k} \sin\left\{\frac{k\pi}{2}\left(\gamma + \alpha - \frac{2\alpha}{\pi}\operatorname{Arg}(x)\right)\right\},$$

– pour $1 < \alpha < 2$:

$$f(x) = \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}\,\Gamma(k/\alpha+1)}{k!} |x|^{k} \sin\left\{\frac{k\pi}{2\alpha}\left(\gamma + \alpha - \frac{2\alpha}{\pi}\operatorname{Arg}(x)\right)\right\},$$

et où, ici, on pose $\operatorname{Arg}(x) = \pi$ si $x > 0$, et $\operatorname{Arg}(x) = 0$ si $x < 0$.

Ces formules ont été obtenues par H. Bergström et W. Feller en 1952. Elles ne constituent pas d'ailleurs les expressions les plus simples qu'on puisse obtenir de ces lois, puisque leurs transformées de Fourier s'expriment de façon relativement moins complexe.

Ces lois de probabilité ont de nombreuses applications. Il y a d'excellentes raisons de croire qu'elles décrivent effectivement tout un ensemble de phénomènes naturels. Parmi ceux-ci, on peut citer aussi bien des grandeurs météorologiques que des fluctuations de cours de bourse. Les modèles apparentés aux lois stables ont été abondamment étudiés dans la littérature scientifique des dernières décennies à la suite, notamment des travaux de Benoît Mandelbrot, qui ont montré leurs liens intimes avec les caractéristiques géométriques d'objets de dimension *non entière* (les *fractals*). Nous renvoyons le lecteur à Mandelbrot (1975), Glieck (1988) et Falconer (1990) pour des détails et références sur le sujet.

Pour les propriétés générales des lois stables, on se référera à Lukács (1970). En ce qui concerne leur analyse statistique, on consultera S. Csörgö, *Adaptative Estimation of the Parameters of Stable Laws,* coll. Math. J. Bolyái, 36, North Holland (1984).

Le caractère *naturel* des lois stables montre également qu'il est *naturel* pour certaines grandeurs aléatoires de ne pas posséder d'espérance mathématique. Le seul cas de loi stable non normale pour lequel la densité s'exprime de manière simple, à savoir, la loi de Cauchy, tombe dans cette catégorie. On constate, en effet, que celle-ci n'a pas d'espérance, puisqu'il n'est pas possible d'attribuer de signification à l'intégrale divergente :

$$\int_{-\infty}^{\infty} \frac{xdx}{\pi(1 + x^2)}$$

Il est relativement facile de *simuler* artificiellement la suite des moyennes d'ordre n d'une suite de variables indépendantes de Cauchy. On observe alors que, conformément à la loi des grands nombres de Kolmogorov, ces moyennes ne convergent pas vers une limite lorsque $n \to \infty$. Leur comportement se trouve être assez déroutant, puisque les valeurs observées, au cours du temps, semblent parfois se stabiliser, pour ensuite, d'un seul coup faire des sauts, dans un sens puis dans un autre, par oscillations brutales et inattendues.

Ce type de variations provoque régulièrement la surprise de ceux qui tombent, pour la première fois, sur l'observation d'un phénomène de ce type. On s'attendrait, naturellement, à devoir observer, dans tous les cas, une convergence des moyennes vers une limite finie. Toujours est-il que, bien que cette convergence soit le cas le plus fréquent, il existe de nombreux exemples répertoriés où celle-ci n'a structurellement pas lieu. Il convient alors de pouvoir interpréter correctement le phénomène, sans chercher, de manière erronée, à attribuer ces fluctuations à des facteurs extérieurs aux observations.

Chapitre IV

FONDEMENTS LOGIQUES DU CALCUL DES PROBABILITÉS. EINSTEIN ET LE MOUVEMENT BROWNIEN, LE MODÈLE DE KOLMOGOROV

Dans ce qui précède, nous avons donné des explications très sommaires de la notion de probabilité. Il convient maintenant d'être un peu plus précis sur le modèle mathématique utilisé dans sa description.

Peut-être faudrait-il également faire une distinction précise entre *les probabilités,* en tant que théorie mathématique, et *les probabilités,* en tant qu'outil descriptif du Monde où nous vivons.

D'innombrables discussions ont été menées sur l'existence ou la non-existence des probabilités dans la nature. Il existe, à ce sujet, des « objectivistes », affirmant que la probabilité *existe naturellement,* pour des raisons de symétrie physique, même si on ne peut pas toujours la déterminer de façon évidente. Ainsi, les jeux de : pile ou face, dés, cartes, fournissent des exemples de probabilités définies de manière objective. D'un autre point de vue, les « subjectivistes » ne veulent pas voir une probabilité associée à chaque événement possible d'un phénomène aléatoire, mais, plutôt, un « degré de croyance » lié à l'observateur, qui exprime simplement par un chiffre son évaluation des chances d'occurrence d'un phénomène. Derrière tout cela se greffent, le plus souvent, des idées religieuses ou philosophiques sur le déterminisme ou le libre arbitre, la toute-puissance de Dieu, et bien d'autres choses encore.

De notre point de vue, il s'agit là d'un affreux mélange, propre à engendrer confusions et querelles, bien loin de discussions scientifiques sérieuses. Le *calcul des probabilités* est une science mathématique tout aussi rigoureuse que la géométrie,