NOUVELLES ANNALES DE MATHÉMATIQUES

EUGÈNE ROUCHÉ

La théorie des chances

Nouvelles annales de mathématiques 3e série, tome 7 (1888), p. 553-588

<http://www.numdam.org/item?id=NAM_1888_3_7__553_0>

LA THÉORIE DES CHANCES.

Calcul des Probabilités; par M. *J. Bertrand,* Membre de l'Académie française, Secrétaire perpétuel de l'Académie des Sciences. Paris, Gauthier-Villars et Fils; 1889.

Nous n'avons pas la prétention de juger le nouvel Ouvrage de M. Bertrand; nous avons pensé que le meilleur hommage à rendre à notre illustre Maître consistait à présenter une analyse substantielle de son Livre. Aussi bien, ceux de nos lecteurs qui sont étrangers à la doctrine des hasards nous sauront gré sans doute de leur offrir ici un résumé, dégagé de toute démonstration, des principes de cette théorie et de ses applications les plus importantes.

I. — L'énumération des chances.

La nature d'un événement aléatoire détermine un certain nombre de cas, tous également possibles, mais les uns favorables, les autres défavorables à l'arrivée de cet événement. Ce qu'il importe surtout de connaître, c'est le rapport du nombre des cas favorables au nombre total des cas, et c'est ce rapport qui a reçu le nom de *probabilité.*

On jette deux dés n fois de suite; quelle est la probabilité d'amener au moins une fois *sonnez,* c'est-à-dire deux 6? Le nombre total des cas est 36^n; celui des combinaisons défavorables est 35^n; la probabilité est donc

$$1-\left(\frac{35}{36}\right)^n;$$

d'où l'on voit qu'il y a infériorité de chances pour gagner en 24 coups et supériorité de chances pour gagner en 25 coups.

Tel est le problème qui, proposé à Pascal par le chevalier de Méré, a donné naissance aux premières recherches sur le calcul des chances.

On peut résoudre un certain nombre de problèmes à l'aide de la théorie des combinaisons et de la seule définition de la probabilité.

Au jeu de *passe dix,* l'un des joueurs jetait trois dés et gagnait si la somme des points amenés surpassait 10. Quelles étaient les probabilités des deux adversaires? Elles étaient égales. Si l'on jouait avec cinq dés au lieu de trois, quel nombre faudrait-il substituer à 10 pour laisser les chances égales? La théorie répond 17.

Deux candidats X et Z sont soumis à un scrutin de ballottage; l'urne renferme $n+k$ billets au nom de X, et n seulement au nom de Z; X sera donc élu. Mais quelle est la probabilité pour que, pendant toute la durée du dépouillement, X ne cesse de conserver l'avantage? Une analyse, d'ailleurs assez délicate, répond

$$\frac{k}{2n+k}.$$

Au jeu de la rencontre, on tire successivement d'une urne les m boules qu'elle contient, sans les remettre dans l'urne. Ces boules portent les numéros 1, 2, 3, ..., m. Quelle est la probabilité pour qu'une boule au moins sorte au rang marqué par le numéro qu'elle porte? Dès que m est égal ou supérieur à 9, cette probabilité diffère très peu de $1 - \frac{1}{e} = 0,63$.

Voilà quelques-uns des nombreux exemples traités dans le premier Chapitre. L'auteur insiste, en outre, sur les erreurs qui peuvent provenir d'une constatation imparfaite ou de l'inégale possibilité des divers cas. Le type des erreurs de ce genre est assurément le dilemme de d'Alembert : « l'événement arrive ou n'arrive pas; de là deux cas, dont un favorable; la probabilité de tout événement est donc $\frac{1}{2}$ ».

Il faut aussi se garder d'introduire, sans explication, l'infini dans les raisonnements. Choisir au hasard entre un nombre infini de cas possibles n'est pas une indication suffisante. Quelle est la probabilité pour qu'un plan, choisi au hasard dans l'espace, fasse avec l'horizon un angle moindre que 45°? La question étant mal posée peut donner lieu à des réponses contradictoires. On peut dire que la probabilité est $\frac{1}{2}$, puisque tous les angles entre 0° et 90° sont possibles, et que ceux entre 0° et 45° sont seuls favorables. On pourrait dire aussi que la probabilité est 0,28; car, si par le centre d'une sphère on imagine le rayon perpendiculaire au plan, il faut, pour que l'angle en question soit inférieur à 45°, que l'extrémité du rayon tombe dans une zone dont le rapport à la demi-sphère est précisément 0,28.

II. — La probabilité totale et la probabilité composée.

L'évaluation directe du nombre des combinaisons est le plus souvent impraticable, tant ce nombre croît rapidement pour peu qu'augmente celui des objets combinés. Demande-t-on, par exemple, de répartir 459 personnes en 8 groupes de 51 membres chacun? Le nombre des distributions possibles ne renferme pas moins de 429 chiffres! D'ailleurs, si l'on pouvait toujours faire directement le compte des cas favorables et celui de tous les cas possibles, le Calcul des probabilités ne serait pas une science. A toute science, il faut des principes. Voici les deux plus simples parmi ceux qui servent de base au Calcul des probabilités :

Si l'on partage les cas favorables en plusieurs groupes, la probabilité de l'événement est la somme des probabilités pour qu'il appartienne à chacun des groupes. C'est le principe de la *probabilité totale.*

Quand un événement consiste dans le concours de plusieurs autres, sa probabilité est le produit des probabilités des événements simples dont il exige la réunion. C'est le principe de la *probabilité composée.*

Dans l'application du premier principe, il faut veiller, en formant les groupes, à ce qu'aucun des cas possibles ne soit omis ni répété. Dans l'application du second, il importe de regarder si les événements simples sont ou non indépendants les uns des autres; et, s'il y a dépendance, on doit calculer la probabilité de chaque événement simple, en tenant compte de ce que ceux qui le précèdent sont arrivés.

Ainsi la probabilité d'amener avec deux dés le point 3 ou le point 4 n'est pas la somme des probabilités pour amener 3 et pour amener 4, puisqu'on peut d'un même coup amener à la fois ces deux points.

Si, de deux météorologistes, dont les prédictions ont pour probabilités respectives p et p', l'un dit qu'il pleuvra demain et l'autre qu'il ne pleuvra pas, la probabilité pour qu'ils disent juste tous les deux n'est pas pp' : elle est nulle.

Les principes sont justes; ils s'appliquent à toutes les combinaisons des probabilités simples évaluées en nombre. Mais, si l'évaluation est imparfaite, les conséquences ne méritent aucun crédit.

C'est donc avec raison que M. Bertrand s'étend longuement dans le Chapitre II sur les précautions à prendre dans la mise en œuvre des deux principes fondamentaux. Après avoir donné des exemples remarquables de leur fausse interprétation, il traite un grand nombre de problèmes dont le choix est bien propre à captiver le lecteur en l'initiant à ce genre de recherches. Nous signalerons particulièrement les questions suivantes :

Quelle est la probabilité des brelans au jeu de la bouillotte?

Quelle est la probabilité de la chance favorable réservée au banquier dans le jeu de trente et quarante?

Au baccarat, est-il avantageux au ponte de demander une carte lorsqu'il a le point 5?

Les réponses aux deux premières questions résultent de formules ou de tableaux que nous ne saurions rapporter ici. Quant à la troisième, on peut dire : « Si, sans jouer au plus fin, dès le début de la partie, le ponte déclare franchement ses habitudes, il doit tirer à 5. Si les conventions du jeu permettent la ruse, il doit se tenir à 5, en faisant croire au banquier, s'il le peut, qu'il a l'habitude de tirer. »

III. — L'espérance mathématique.

On appelle *espérance mathématique* de celui qui prétend à une somme éventuelle la valeur probable de cette somme, c'est-à-dire le produit de la somme par la probabilité que le prétendant a de l'obtenir.

Dans tout jeu équitable, la mise de chaque joueur doit être égale à son espérance mathématique; et, par suite, *les mises des divers joueurs sont proportionnelles à leurs probabilités respectives de gagner.* C'est la règle des *paris.*

Lorsque des joueurs se séparent avant que la partie soit terminée, ils doivent partager l'enjeu proportionnellement aux probabilités respectives qu'ils auraient alors de gagner. C'est la règle des *partis* (compositio sortis).

Pierre a trois pièces de 5^{fr}, Paul en a deux. Chacun doit jeter ses pièces, et celui qui obtiendra le plus grand nombre de faces prendra les cinq pièces. Le jeu est-il équitable? Nullement, il est avantageux pour Pierre, dont l'espérance mathématique, $\frac{200}{11}$, est supérieure à sa mise.

Plusieurs joueurs ayant déposé chacun 1^{fr} jettent tour à tour μ dés. L'enjeu total appartiendra à celui qui amènera la plus grande somme de points ou sera partagé entre ceux qui amèneront une même somme de points supérieure à celle obtenue par les autres joueurs. Pierre joue le premier, il amène k points; quel est le nombre d'adversaires le plus favorable à ses intérêts, c'est-à-dire le nombre d'adversaires qui lui laisse la plus grande espérance mathématique? La solution dépend d'une équation aisée à résoudre avec une table de logarithmes. Par exemple, si le nombre des dés est six et si Pierre a obtenu cinq 6 et un 5, le nombre d'adversaires le plus avantageux est 15144.

Parmi les questions que la considération de l'espérance mathématique permet de résoudre avec élégance et facilité, il convient encore de citer le problème de l'aiguille et celui de Pétersbourg.

Des lignes parallèles et équidistantes sont tracées sur un plan indéfini; on lance au hasard sur ce plan une aiguille. Quelle est la probabilité pour que l'aiguille, parvenue au repos, rencontre l'une des parallèles? La théorie répond

$$\frac{2l}{a\pi},$$

l étant la longueur de l'aiguille et a la distance de deux parallèles. De là résulterait un moyen, assurément plus curieux qu'utile, de calculer approximativement le rapport π de la circonférence au diamètre.

Le problème de Pétersbourg est resté célèbre à cause des controverses singulières qu'il a suscitées.

Pierre et Paul jouent à pile ou face aux conditions suivantes : Pierre payera à Paul 1^{fr} s'il amène pile au premier coup, 2^{fr} s'il n'amène pile qu'au deuxième coup, 4^{fr} s'il n'amène pile qu'au troisième coup, ..., 2^{n-1} francs s'il n'amène pile qu'au $n^{ième}$ coup, et ainsi de suite, de manière que la partie ne se termine que lorsque Pierre aura amené pile. Quelle est l'espérance mathématique de Paul? Le calcul apprend que cette espérance est infinie. On a voulu voir là un paradoxe. Un tel jeu est déraisonnable, d'accord; il pourrait même devenir impraticable si l'on introduisait la condition de déposer la mise chaque fois avant de jeter la pièce. Mais l'assertion du calcul est exacte, et rien ne saurait prévaloir contre elle, ni la doctrine de l'espérance morale de Daniel Bernoulli, ni les dissertations éloquentes de Buffon.

« Un ingénieur calcule la charge capable d'abaisser de $0^m,50$ le tablier d'un pont. L'épreuve est inutile, imprudente, dangereuse; le poids calculé est-il moins juste? Il est mauvais de trop charger un pont, mauvais aussi de jouer trop gros jeu. Cela ne change ni la théorie du jeu, ni celle de l'élasticité. »

IV. — Les épreuves répétées et le théorème de Bernoulli.

La théorie des chances acquiert une grande importance, lorsque, au lieu de l'appliquer à la recherche de la probabilité d'un événement qui ne doit plus se reproduire, on considère des épreuves répétées un grand nombre de fois dans des circonstances identiques.

Soient p et q les probabilités de deux événements contraires A et B, en sorte que $p+q=1$. Désignons par la notation $(A_m B_{\mu-m})$ l'événement composé qui consiste dans m apparitions de A et $\mu - m$ apparitions de B, sur μ épreuves.

La probabilité de l'événement $(A_m B_{\mu-m})$ *est égale au* $(m+1)^{ième}$ *terme du développement du binôme* $(p+q)^m$ *ordonné suivant les puissances croissantes de* p.

Par exemple, si, dans une urne renfermant 4 boules blanches et 5 noires, on fait trois tirages successifs en remettant dans l'urne la boule sortie, la probabilité d'amener 1 blanche et 2 rouges sera

$$3 \cdot \frac{5}{9}\left(\frac{4}{9}\right)^2 = \frac{80}{243}.$$

Tel est le principe des *épreuves répétées*. On en déduit ce corollaire important :

Si le nombre μ *des épreuves est très considérable, l'événement composé le plus probable est* $(A_{\mu p}, B_{\mu q})$, c'est-à-dire *celui où les événements simples* A *et* B *entrent proportionnellement à leurs probabilités*. Cet événement composé, quoique étant le plus probable, n'a cependant qu'une probabilité très faible si μ est, comme nous le supposons, considérable.

On regarde ordinairement la valeur la plus probable μp du nombre des apparitions de l'événement A comme une valeur normale à laquelle on rapporte les autres nombres d'arrivées de cet événement. Ainsi, l'on pose $m = \mu p - z$, z étant d'ailleurs positif ou négatif; il en résulte $\mu - m = \mu q + z$, en sorte que l'événement composé quelconque $(A_m B_{\mu-m})$; se trouve désigné par $(A_{\mu p-z} B_{\mu q+z})$ la quantité z est ce qu'on nomme l'*écart* relatif à cet événement. Si, par exemple, à pile ou face, sur 10000 épreuves, face se présente 5017 fois, on aura $\mu = 10000$, $p = \frac{1}{2}$, d'où $50000 - z = 5017$; l'écart sera donc -17.

Lorsque μ est un grand nombre, le calcul des termes du développement du binôme devient impraticable, même avec une Table de logarithmes; il faut avoir recours à la formule d'approximation de Stirling

$$1.2.3\ldots x = e^{-x} x^x \sqrt{2\pi x},$$

dont M. Bertrand donne une démonstration simple et naturelle.

A l'aide de cette formule, on trouve

$$(1) \qquad \frac{1}{\sqrt{2\pi\mu pq}}$$

pour la valeur approchée de la probabilité de la combinaison $(A_{\mu p}, B_{\mu q})$ la plus probable.

On démontre en outre que, *si l'on pose*

$$(2) \qquad \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2} dx = \Theta(t),$$

on a

$$(3) \qquad P = \Theta\left(\frac{\alpha}{\sqrt{2\mu pq}}\right)$$

pour la probabilité que l'écart z soit compris entre α et $-\alpha$, c'est-à-dire pour que le nombre d'arrivées de l'événement A soit compris entre $\mu p - \alpha$ et $\mu p + \alpha$.

Enfin de cette formule remarquable, qu'on nomme *formule des écarts*, résulte aisément la proposition suivante :

Soient p la probabilité de l'événement A *et m le nombre des apparitions de cet événement sur μ épreuves; on peut prendre μ assez grand pour qu'il y ait une probabilité* *aussi voisine que l'on voudra de la certitude que l'écart relatif*

$$p - \frac{m}{\mu}$$

devienne inférieur à toute quantité donnée.

C'est le théorème de Jacques Bernoulli.

Les géomètres du XVII^e siècle, Pascal, Fermat, Leibnitz, Huygens n'avaient guère en vue que le problème des partis, et le Calcul des probabilités n'était à leurs yeux qu'une science purement spéculative. C'est le théorème de Bernoulli qui, publié en 1713 dans l'*Ars conjectandi,* a donné à la théorie des chances sa valeur objective.

Outre la démonstration qui dérive de la formule des écarts, M. Bertrand donne du théorème de Bernoulli deux preuves plus élémentaires fondées sur la considération de la valeur probable du carré ou de la valeur absolue de l'écart. C'est sans doute par inadvertance que, dans le premier des deux Chapitres consacrés à ce sujet, le théorème n'a pas été énoncé en termes formels. « Le hasard corrige le hasard » est une image heureuse, mais qui devrait être immédiatement suivie d'un énoncé mathématique.

Le théorème de Bernoulli est susceptible de nombreuses applications : mais encore y a-t-il certaines précautions à prendre. Deux conditions sont indispensables, et il faut préalablement s'assurer qu'elles sont remplies : la probabilité doit rester constante pendant les épreuves; elle doit en outre avoir une valeur *objective*, c'est-à-dire bien déterminée indépendamment des renseignements connus, la même en un mot pour tous ceux qui l'évaluent sans se tromper. Par exemple, la probabilité pour qu'il pleuve demain, la probabilité pour qu'un homme âgé de 40 ans vive dans dix ans, sont subjectives et ne comportent pas l'application du théorème de Bernoulli. M. Bertrand se livre sur ce sujet à une dissertation un peu étendue, mais fort intéressante.

V. — La ruine des joueurs.

Pierre joue à un jeu équitable une série de parties. Sa mise à chaque partie est a et celle de son adversaire est b ; en sorte que, à chaque partie, Pierre expose a francs avec une probabilité égale à $\frac{a}{a+b}$ d'en gagner b. Plus b est grand, a restant le même, plus Pierre joue *gros jeu;* c'est une définition.

Or la formule des écarts montre que la probabilité P, pour qu'après μ parties la somme perdue ou gagnée par Pierre soit inférieure à une somme donnée S, a pour expression

$$P = \Theta\left(\frac{S}{\sqrt{2\mu ab}}\right).$$

La situation faite à Pierre par les conditions du jeu et par le nombre μ des parties dépend donc uniquement du produit μab. Plus Pierre joue gros jeu, c'est-à-dire plus b est grand a restant le même, moindre est le nombre μ des parties qu'il faut jouer pour amener la même situation, c'est-à-dire le même gain final ou la même perte; μ est inversement proportionnel à b. Ainsi supposons que, dans un cas, Pierre expose à chaque partie 10^{fr} pour en gagner autant, et que dans une seconde série il expose encore 10^{fr} pour en gagner 350, le jeu étant toujours, bien entendu, supposé équitable; il faudra dans ce second cas faire un nombre de parties 35 fois moindre que dans la première, 751 par exemple au lieu de 20000.

La quantité $1 - P$ représente la probabilité pour que la perte de Pierre, dans μ parties, aux conditions indiquées, soit plus grande que S. Or les Tables de la fonction Θ donnent

$$\Theta(0,09) = 0,10;$$

si donc on pose

$$\frac{S}{\sqrt{2\mu ab}} = 0,09,$$

et si l'on fait $a = b = 1$, la formule

$$\mu = (62,4)S^2$$

indiquera le nombre des parties nécessaires pour que la probabilité d'une perte supérieure à S soit égale à $\frac{9}{10}$. Ainsi, à 1^{fr} la partie, pour avoir une probabilité égale à $\frac{9}{10}$ de perdre 100^{fr}, 1000^{fr}, 10000^{fr}, 100000^{fr}, ..., il faudrait un nombre de parties égal à 624 mille, 62 millions 400 mille, 6240 millions, 624 milliards, etc.

La chance de perte est donc loin d'être effrayante quand le nombre des parties est fixé d'avance, le règlement ayant lieu à la fin.

Cette application intéressante de la formule des écarts nous ramène à la théorie du jeu, qui se mêle, bon gré malgré, à toute exposition des principes du Calcul des probabilités.

La question de la ruine des joueurs et de la durée du jeu a préoccupé les géomètres les plus distingués, Huygens, Moivre, Lagrange, Laplace, Ampère, etc.; elle est cependant loin d'être épuisée, tant on peut changer les conditions et varier le point de vue.

Voici les principales questions traitées dans le Livre de M. Bertrand.

Pierre joue contre le public, c'est-à-dire contre tout venant, ou encore contre un adversaire infiniment riche. Le jeu est équitable ou non, mais les conditions sont invariables d'une partie à l'autre. Quelle est la probabilité pour qu'il finisse par se ruiner?

Désignons par p la probabilité que Pierre a de gagner à chaque partie, par a sa mise, par b celle de l'adversaire et par ε la quantité

$$(a+b)p-a,$$

que nous nommerons l'*avantage* de Pierre à chaque coup; avantage qui est nul quand le jeu est équitable et qui constitue en réalité une défaveur lorsqu'il est négatif.

Cela posé, le calcul montre que la ruine de Pierre est certaine si ε est négatif ou nul; si ε est positif, la certitude de ruine disparaît; c'est le cas du banquier dans les jeux publics. Par exemple, à la roulette ordinaire, la probabilité p du banquier à chaque partie est $\frac{19}{37}$, et, les mises étant égales, on a

$$\varepsilon = a(2p-1) = \frac{a}{37}.$$

La chance de ruine du banquier est

$$\left(\frac{18}{19}\right)^n,$$

n étant le rapport de sa fortune à la mise totale de l'un des coups. Cette chance est très faible; pour $n = 1000$, elle est plus petite que

$$0,00000\,00000\,00000\,00000\,0007.$$

Connaissant la fortune m de Pierre, on peut se demander quelle est la probabilité pour que Pierre, jouant contre le public, dans des conditions équitables, soit ruiné juste à la fin de la $\mu^{\text{ième}}$ partie. L'enjeu étant de 1^{fr} la partie, on trouve pour cette probabilité la valeur approchée très simple

$$\frac{m}{\mu}\sqrt{\frac{2}{\mu\varpi}}\,e^{-\frac{m^2}{2\mu}},$$

d'où résulte la probabilité

$$1-\Theta\left(\frac{m}{\sqrt{2\mu}}\right)$$

pour que Pierre soit ruiné avant le $\mu^{\text{ième}}$ coup. Par exemple, si $m = 100^{\text{fr}}$, la probabilité pour que la ruine de Pierre précède la dix-millième partie est 0,3154. De la première formule on déduit encore que le nombre μ des parties pour lequel la probabilité de voir Pierre ruiné juste au $\mu^{\text{ième}}$ coup a la valeur maxima est

$$\mu = \frac{1}{3}m^2;$$

ce nombre est égal à 3333 quand m est égal à 100, et la probabilité maxima est

$$0{,}0000925.$$

Pierre et Paul, dont les fortunes sont respectivement A et B, font un nombre illimité de parties. Quelle est pour Pierre la probabilité P de ruiner Paul? Si le jeu est équitable, on a

$$P = \frac{A}{A+B}.$$

Quand le jeu n'est pas équitable, la réponse est moins simple; en désignant par a et b les mises à chaque coup, par p et $q = 1 - p$ les probabilités respectives de Pierre et de Paul pour gagner chaque partie, on a

$$P = \frac{\alpha^{A}-1}{\alpha^{A+B}-1},$$

α désignant la racine positive, autre que 1, de l'équation trinôme

$$pX^{a+b} - X^{a} + q = 0.$$

Le résultat se traduit aisément en langage ordinaire dans le cas particulier où $a = b = 1$ et $A = B$. Les chances de ruine pour les deux joueurs sont dans le rapport de p^A à q^A.

Pierre et Paul jouent l'un contre l'autre jusqu'à la ruine de l'un d'eux; A et B sont leurs fortunes primitives, a et b leurs mises à chaque partie, p et $q = 1 - p$ leurs probabilités respectives de gagner l'une quelconque des parties. Quelle est la valeur probable V du nombre des parties qui seront jouées? En d'autres termes, si l'on promet à une troisième personne, Jean, qui ne participe pas au jeu, 1fr par partie jouée, quelle est l'espérance mathématique de Jean? Lorsque le jeu est équitable, on a

$$V = \frac{AB}{ab};$$

le nombre probable des parties est proportionnel au produit des fortunes. Mais qu'arrive-t-il quand le jeu n'est pas équitable et quelle proposition faut-il substituer au théorème si élégant de M. Bertrand? L'étude de ce cas nous a conduit à une proposition qui se recommande aussi par sa simplicité; la formule est alors

$$V = \frac{(A+B)P - A}{(a+b)p - a},$$

où P désigne la probabilité, calculée dans l'alinéa précédent, pour que Pierre ruine Paul. En donnant au mot *avantage* le sens précis que nous avons indiqué au commencement de ce paragraphe, on peut dire : le nombre probable des parties est égal au rapport de l'avantage total de l'un quelconque des deux joueurs à l'avantage du même joueur dans chaque partie.

A propos de ce théorème, M. Bertrand signale une tentative de démonstration *a priori* qu'il propose comme un nouvel exemple de ces raisonnements en apparence fort plausibles, mais dépourvus de rigueur, auxquels on est particulièrement exposé dans la théorie des chances. Qu'on nous permette ici une réflexion. Les vues *a priori* procurent à l'esprit une satisfaction incontestable. Mais la plupart de ces démonstrations ne sont-elles pas suggérées par la connaissance préalable du résultat? Lorsque Poinsot disait : « les formules ne donnent que ce qu'on y a mis », ne plaidait-il pas un peu sa propre cause? Pris à la lettre, cet aphorisme de l'éminent géomètre ne serait rien moins que la négation de l'Algèbre. Les transformations d'une équation algébrique ne sont au fond que des transformations de l'idée qu'elle exprime; mais, faites d'une main sûre d'après des règles connues, elles conduisent naturellement à des rapprochements qui, sans le secours des formules, resteraient souvent inaperçus.

Pierre et Paul jouent l'un contre l'autre avec des probabilités égales. Ils possèdent chacun n francs avant d'entrer au jeu; à chaque partie le perdant donne 1^{fr} au gagnant, et le jeu ne cesse que lorsque l'un quelconque des joueurs est ruiné. Quelle est la probabilité pour que le jeu se termine précisément à la fin de la $\mu^{\text{ième}}$ partie? Cette probabilité est nulle si $\mu - n$ est impair; mais, si $\mu - n$ est pair, l'expression de la probabilité, sans être compliquée, ne saurait être traduite simplement en langage ordinaire; elle contient en facteur un déterminant composé, chose assez curieuse, avec les coefficients de la fonction bien connue V_n que l'on rencontre dans la théorie de la division du cercle en parties égales. Pour $n = 2$ ou $n = 3$, ce déterminant se réduit à son terme principal et l'on obtient les résultats suivants : si la fortune primitive de chacun des joueurs est de 2^{fr}, la probabilité pour que le jeu cesse au $2\mu^{\text{ième}}$ coup est $\left(\frac{1}{2}\right)^{\mu}$, tandis que, si la fortune primitive est égale à 3^{fr}, la probabilité pour que le jeu se termine à la $(2\mu+1)^{\text{ième}}$ partie est $\frac{1}{4}\left(\frac{3}{4}\right)^{\mu-1}$.

Les joueurs systématiques espèrent accroître leurs chances de gain en observant une certaine progression dans leurs mises, en réglant d'une certaine façon leurs entrées au jeu et leurs sorties. Pierre, par exemple, entre au jeu, supposé équitable, avec la résolution de continuer tant qu'il gagnera et de se retirer dès sa première perte. Il se dit que le nombre des parties pouvant être illimité, son bénéfice peut être immense, tandis que sa perte ne peut dépasser sa mise pour une seule partie. Pierre se leurre d'un vain espoir; p étant sa probabilité et $q = 1 - p$ celle de son adversaire pour gagner l'une quelconque des parties, et a_i et b_i désignant les mises des deux joueurs à la partie de rang i, la valeur probable des sommes que Pierre espère toucher a pour expression

$$\begin{aligned} pq(a_1+b_1) &+ p^2q(a_1+b_1-a_2+b_2) \\ &+ p^3q(a_1+b_1-a_2-b_2+a_3+b_3)+\ldots \end{aligned}$$

tandis que la valeur probable de ses débours possibles est

$$qa_1 + pq(a_1 + a_2) + p^2q(a_1 + a_2 + a_3) + \ldots;$$

or ces deux valeurs sont égales, puisque, le jeu étant équitable, on a

$$p(a_i + b_i) = a_i.$$

Ce calcul n'était pas d'ailleurs indispensable. Toutes les combinaisons plus ou moins ingénieuses des joueurs systématiques sont illusoires; si le jeu est équitable, l'espérance mathématique du joueur est pour chaque partie égale à celle de son adversaire, et cette égalité subsiste quels que soient la progression des mises et le nombre des parties. Le joueur peut accroître la valeur possible de son gain, mais en diminuant proportionnellement la probabilité de l'obtenir.

VI. — La probabilité a posteriori.

Après cette digression sur la durée du jeu, revenons aux principes; il n'en reste plus qu'un à exposer : c'est le théorème sur la probabilité des causes ou la *probabilité a posteriori*, que Bayes a fait connaître en 1763 dans les *Transactions philosophiques*.

Il résulte des principes de la probabilité totale et de la probabilité composée que, lorsqu'un événement E peut provenir de plusieurs causes C_1, C_2, ..., C_n qui s'excluent mutuellement, sa probabilité a pour expression

$$p_1q_1 + p_2q_2 + \ldots + p_nq_n,$$

q_i désignant la probabilité d'action *a priori*, c'est-à-dire avant toute épreuve, de la cause C_i, et p_i désignant la probabilité que la cause C_i, quand elle agit, donne à l'événement.

Mais, après l'épreuve, quand l'événement E est arrivé, la probabilité Q_i pour qu'il soit dû à l'action de la cause C_i diffère de q_i; on lui donne le nom de *probabilité a posteriori* de la cause considérée.

Un exemple rendra la distinction très sensible : deux urnes, d'apparence semblable, renferment, la première, 6 boules blanches et 1 noire; la seconde, 2 boules blanches et 5 noires; l'événement consiste dans l'extraction d'une boule blanche. Les causes sont ici les deux urnes; leurs probabilités d'action *a priori* sont $q_1 = q_2 = \frac{1}{2}$, puisque les urnes sont semblables. Les probabilités que le choix de l'urne donne à l'événement sont $p_1 = \frac{6}{7}$ pour la première urne et $p_2 = \frac{2}{7}$ pour la seconde; en sorte que la probabilité de la sortie d'une boule blanche est

$$\frac{1}{2}\frac{6}{7} + \frac{1}{2}\frac{2}{7} = \frac{4}{7}.$$

Posons maintenant la question d'autre sorte. Supposons que l'événement ait eu lieu, que la boule tirée soit blanche. Si l'on demande quelle est alors la probabilité Q_1 pour que la boule provienne de la première urne, on sent bien que Q_1 n'est pas égal à $\frac{1}{2}$, mais lui est supérieur.

La règle de Bayes a pour objet la détermination de ces probabilités *a posteriori* $Q_1, Q_2, \ldots, Q_n$. Voici son énoncé :

Plusieurs causes $C_1, C_2, \ldots, C_n$ *peuvent produire un événement* E; $q_1, q_2, \ldots, q_n$ *sont les probabilités* a priori *de ces causes, et* $p_1, p_2, \ldots, p_n$ *sont les probabilités que chacune d'elles donne à l'événement. Quand l'événement* E *a eu lieu, la probabilité* Q_i *pour qu'il soit dû à la cause* C_i *est donnée par la formule*

$$(4)\qquad Q_i = \frac{p_i q_i}{p_1 q_1 + p_2 q_2 + \ldots + p_n q_n}.$$

En d'autres termes, les probabilités *a posteriori* $Q_1, Q_2, \ldots, Q_n$ sont proportionnelles aux produits de la probabilité *a priori* de chaque cause par la probabilité qu'elle donne a l'événement.

Dans l'exemple considéré ci-dessus, on a

$$Q_1 = \frac{3}{4} \qquad \text{et} \qquad Q_2 = \frac{1}{4}.$$

Une urne contient des boules blanches et des boules noires. Chaque boule porte un des n^{os} 1, 2, ..., n. On connaît pour l'un quelconque i de ces numéros le rapport q_i du nombre des boules marquées i au nombre total des boules, et aussi le rapport p_i du nombre des boules blanches marquées i au nombre total des boules qui portent le même numéro. On a tiré une boule, elle est blanche; quelle est la probabilité Q_i pour qu'elle porte le n° i? La solution est donnée par la formule (4) : ici ce sont les numéros qui représentent les causes.

Une urne contient μ boules, blanches ou noires, en proportion inconnue. On a fait n tirages, en remettant chaque fois dans l'urne la boule sortie. Les n tirages n'ont amené que des boules blanches. Quelle est la probabilité Q_μ pour que toutes les boules de l'urne soient blanches?

Toutes les hypothèses sur la composition de l'urne sont possibles : ce sont là les causes $C_0, C_1, C_2, \ldots, C_\mu$. La cause C_i, relative au cas où l'urne serait composée de i boules blanches et de $\mu - i$ boules noires, donne à l'événement la probabilité $p_i = \left(\frac{i}{\mu}\right)^n$. Mais les probabilités, *a priori*, $q_1, \ldots, q_\mu$ des diverses compositions de l'urne restent complètement inconnues. La probabilité demandée, c'est-à-dire la probabilité *a posteriori* de la cause C_μ, contient donc dans son expression

$$Q_\mu = \frac{q_\mu}{\left(\frac{1}{\mu}\right)^n q_1 + \left(\frac{2}{\mu}\right)^n q_2 + \ldots + \left(\frac{\mu}{\mu}\right)^n q_\mu}$$

les rapports inconnus de $q_1, q_2, \ldots, q_{\mu-1}$ à q_μ.

Le problème, tel qu'il est posé, est donc insoluble, les données étant insuffisantes. Il devient bien déterminé, si l'on admet que toutes les compositions de l'urne sont également possibles : les rapports en question sont alors égaux à l'unité et l'on a

$$Q_\mu = \frac{\mu^n}{1^n + 2^n + \ldots + \mu^n}.$$

Le problème serait encore déterminé, mais le résultat serait tout autre, si l'on admettait que la composition de l'urne a été faite au moyen d'un tirage au sort de la couleur blanche ou noire, avec probabilité $\frac{1}{2}$ pour chacune des couleurs, ou, ce qui revient au même, en tirant à pile ou face la couleur de chaque boule. Alors, les probabilités *a priori* $q_0, q_1, q_2, \ldots, q_\mu$ seraient les divers termes du développement de $\left(\frac{1}{2} + \frac{1}{2}\right)^\mu$, et l'on aurait

$$Q_\mu = \frac{\mu^n}{\mu + \frac{\mu(\mu-1)}{2} 2^n + \frac{\mu(\mu-1)(\mu-2)}{1.2.3} 3^n + \ldots + \mu^n}.$$

Une indétermination du même genre se rencontrerait dans le problème suivant :

Une urne contient μ boules, blanches ou noires, en proportion inconnue. On a fait $m+n$ tirages, en remettant chaque fois dans l'urne la boule sortie. On a ainsi obtenu m boules blanches et n noires. Quelle est la composition la plus probable de l'urne?

Toutes les hypothèses sur la composition de l'urne étaient possibles avant l'épreuve. Le nombre x des boules blanches peut prendre toutes les valeurs entières entre o et μ; chacune de ces hypothèses sur la valeur de x donne à l'événement, c'est-à-dire à la sortie de m blanches et de n noires, la probabilité

$$\frac{1.2\ldots(m+n)}{1.2\ldots m\,1.2\ldots n}\left(\frac{x}{\mu}\right)^m\left(1-\frac{x}{\mu}\right)^n;$$

mais la probabilité *a priori* de chacune de ces hypothèses sur la valeur de x reste inconnue, et le problème reste insoluble tant qu'on n'ajoute rien à l'énoncé.

Si l'on complète l'énoncé en regardant toutes les hypothèses, en nombre infini, comme également possibles, les quantités q qui figurent dans la formule de Bayes disparaissent, et les probabilités *a posteriori* Q deviennent simplement proportionnelles aux valeurs de p, c'est-à-dire aux valeurs de

$$x^m(\mu-x)^n.$$

La composition la plus probable de l'urne répondra donc à la valeur de x qui rend ce produit maximum ; et comme, d'après un théorème élémentaire, cette valeur de x est donnée par la relation

$$\frac{x}{m}=\frac{\mu-x}{n},$$

on voit que, dans la composition la plus probable, le rapport du nombre des boules blanches à celui des boules noires sera celui de m à n.

Si l'on complétait, au contraire, l'énoncé, en admettant, ce qui est plus rationnel, que la composition de l'urne a été fixée par le hasard, c'est-à-dire, comme nous l'avons expliqué au problème précédent, en tirant à pile ou face la couleur de chaque boule, le résultat serait bien différent. On trouve alors que, dans la composition la plus probable de l'urne, le rapport du nombre des boules blanches à celui des boules noires est, si μ est assez grand, égal au rapport de $\mu + 2m$ à $\mu + 2n$.

Il faut bien le dire : on a fait parfois de la règle de Bayes des applications bien peu judicieuses. Buffon a jeté 4040 fois une pièce de monnaie et a obtenu face 2048 fois. Quelle est la probabilité pour que la pièce fût imparfaite et eût une tendance à favoriser face? La question est insoluble, les données manquent, la probabilité *a priori* de telle ou telle imperfection de la pièce est inconnue. Poisson l'a cependant calculée. Mais son raisonnement suppose que, si l'on désigne respectivement par $\frac{1}{2} + z$ et $\frac{1}{2} - z$ les probabilités données par la pièce à l'arrivée de face et à l'arrivée de pile, toutes les valeurs de z entre $-\frac{1}{2}$ et $+\frac{1}{2}$ soient également possibles *a priori*. Il trouve 0,81 pour la probabilité que z soit positif. M. Bertrand fait voir que, si la pièce n'avait été jetée qu'une fois et qu'elle eût montré face, un calcul fondé sur le même principe donnerait 0,75 pour la probabilité que la pièce eût une tendance à favoriser face. Or, dans ce cas, l'événement ne saurait évidemment rien apprendre sur la qualité de la pièce. Que vaut donc le principe du calcul de Poisson?

La liste serait longue des erreurs de ce genre, que M. Bertrand se plaît à relever, saisissant ainsi l'occasion d'exercer sa verve étincelante, pour la grande joie de ses lecteurs.

Un corollaire important du théorème de Bayes est relatif à la recherche de la probabilité des événements futurs déduite des événements observés. Le voici :

Soient E *et* E′ *deux événements dépendant l'un et l'autre des causes* $C_1, C_2, \ldots, C_n$ *dont les probabilités* a priori *sont* $q_1, q_2, \ldots, q_n$; *soient encore* $p_1, p_2, \ldots, p_n$ *les probabilités que ces causes donnent à l'événement* E, *et* $p'_1, p'_2, \ldots, p'_n$ *les probabilités qu'elles donnent à l'événement* E′. *La probabilité pour que, l'événement* E *ayant eu lieu, l'événement* E′ *se produise, a pour expression*

$$(5) \qquad \frac{p_1 p'_1 q_1 + p_2 p'_2 q_2 + \ldots + p_n p'_n q_n}{p_1 q_1 + p_2 q_2 + \ldots + p_n q_n}.$$

Par exemple, une urne contient des boules blanches et des boules noires en proportion inconnue ; toutes les hypothèses sur la composition de l'urne sont supposées également possibles. On a fait $m+n$ tirages, en remettant chaque fois dans l'urne la boule sortie. On a amené de la sorte m boules blanches et n noires. Quelle est la probabilité pour qu'un nouveau tirage donne une boule blanche? On trouve

$$\frac{m+1}{m+n+2}.$$

Plus encore peut-être que le théorème de Bayes, cette formule a donné lieu à des applications ridicules. Condorcet n'est-il pas allé jusqu'à en déduire la probabilité pour que le Soleil se lève demain! Mais comment accepter l'assimilation entre la probabilité de voir le Soleil se lever et celle d'extraire une boule blanche? L'une des probabilités est subjective, l'autre est objective. D'après Condorcet, la probabilité de voir le Soleil se lever a donc été $\frac{2}{3}$ au lendemain de la création de l'homme, $\frac{3}{4}$ au surlendemain, $\frac{4}{5}$ au jour suivant, $\frac{366}{367}=0,997$ au bout d'une année ; elle serait

$$\frac{2191501}{2191502}=0,999\,999$$

après six mille ans.

Si la croyance au lever du Soleil s'est changée en certitude, « c'est par la découverte des lois astronomiques et non par le succès renouvelé d'un même jeu de hasard ».

VII. — La loi de probabilité des erreurs fortuites.

Parmi les applications de la doctrine des chances, celle qui offre le plus d'intérêt est relative à la combinaison des erreurs d'observation.

Il semblait qu'il n'y eût guère qu'à glaner dans un champ si habilement exploité par Gauss : on va voir cependant quelle moisson abondante M. Bertrand a su encore y recueillir. Mais, avant d'analyser les quatre Chapitres consacrés à ce sujet, nous devons préciser le problème à résoudre.

Les erreurs inhérentes à la mesure de toute grandeur physique sont dites *systématiques* ou *fortuites*, suivant que les causes qui les produisent sont permanentes et régulières ou irrégulières et accidentelles. Les erreurs provenant d'une imperfection spéciale de l'instrument employé, d'une tendance particulière de l'observation, d'une défectuosité de la théorie, etc., se rapportent à la première classe, tandis qu'il faut ranger dans la seconde les erreurs dues à l'imperfection générale de nos sens, aux ébranlements de l'air, aux trépidations des supports, etc. Nous excluons de la recherche actuelle les erreurs systématiques : c'est à l'observation qu'il appartient de rechercher les causes qui peuvent engendrer une erreur constante, pour apprécier leur effet et en purger les observations. Il en est tout autrement des erreurs fortuites; on est obligé de les tolérer, et c'est par une combinaison habile des résultats qu'on peut réduire leur influence.

La question se pose dès lors en ces termes :

On a déterminé, à l'aide d'observations faites avec le même soin, les valeurs

$$r_1, r_2, \ldots, r_n$$

de n fonctions,

$$f_1(\mathrm{U}, \mathrm{V}, \mathrm{W}, \ldots), \quad f_2(\mathrm{U}, \mathrm{V}, \mathrm{W}, \ldots), \quad \ldots, \quad f_n(\mathrm{U}, \mathrm{V}, \mathrm{W}, \ldots),$$

de m inconnues U, V, W, ...; n est plus grand que m. On demande de trouver les *meilleures* valeurs de U, V, W, ...

A chaque système u, v, w, ... de valeurs attribuées à U, V, W, ... répondent des différences

$$\begin{array}{l} f_1(u, v, w, \ldots) - r_1, \\ f_2(u, v, w, \ldots) - r_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_n(u, v, w, \ldots) - r_n \end{array}$$

entre les valeurs calculées et les valeurs observées. On donne à ces différences le nom de *résidus*.

L'idée qui s'offre la première serait de choisir les valeurs de u, v, w, ... qui rendent minimum la somme des valeurs absolues des résidus. Mais, comme l'Analyse mathématique se prête mal à cette condition, Legendre a été conduit à considérer une somme de puissances paires des résidus, et particulièrement la somme de leurs carrés, afin d'éviter la complication des calculs. La méthode à laquelle il a donné le nom de *méthode des* *moindres carrés* consiste donc dans le choix des valeurs de u, v, w, ... qui rendent les sommes des carrés des résidus minima.

Ces considérations n'ont, il est vrai, rien de rigoureux; mais il faut reconnaître aussi que la question a été posée d'une manière un peu vague. L'énoncé devient plus précis lorsque, au lieu de demander les valeurs les meilleures, on demande les valeurs les plus probables.

Seulement une nouvelle question surgit alors : quelle est la *loi de probabilité des erreurs fortuites?* Voici ce qu'il faut entendre par là :

D'après le principe de la probabilité totale, si un intervalle se compose de plusieurs autres, la probabilité pour qu'une erreur tombe dans l'intervalle total est la somme des probabilités pour qu'elle tombe dans chacun des intervalles partiels. La probabilité pour qu'une erreur tombe entre x et $x+dx$ est donc la différentielle $\varphi(x)\,dx$ de la fonction qui exprime la probabilité pour que l'erreur tombe entre o et x, et la probabilité pour qu'une erreur tombe entre a et b est exprimée par l'intégrale

$$\int_a^b \varphi(x)\,dx.$$

La recherche de la fonction $\varphi(x)$, qu'on nomme la *facilité* de l'erreur x, est donc la première à entreprendre.

Disons tout de suite avec Gauss que, à proprement parler, cette recherche est vaine : « *Vix, ac ne vix quidem, nunquam in praxi possibile erit, hanc functionem a priori assignare.* » On ne saurait découvrir ce qui n'existe pas; mais on peut chercher une expression analytique qui traduise fidèlement les caractères généraux des erreurs fortuites et que ses conséquences justifient.

Tout le monde s'accorde à reconnaître aux erreurs fortuites les trois caractères suivants : deux erreurs égales et de signes contraires ont la même probabilité; dans une même série d'observations, toutes les erreurs restent comprises entre deux limites très resserrées $-l$ et $+l$, dont la valeur absolue dépend du genre des observations; enfin, entre les limites $-l$ et $+l$, les erreurs sont réparties de telle façon que les plus petites soient notablement les plus probables.

Il résulte de là que $\varphi(x)$ doit être une fonction impaire.

prenant sa valeur maximum pour $x = 0$, puis décroissant rapidement quand x varie de o à $\pm l$, de manière à avoir des valeurs insensibles entre $-l$ et $-\infty$, ainsi qu'entre $+l$ et $+\infty$; on a d'ailleurs par là même

$$(6)\qquad \int_{-\infty}^{+\infty} \varphi(x)\,dx = 1, \qquad \int_{-\infty}^{\infty} x\,\varphi(x)\,dx = 0.$$

Ces diverses propriétés sont insuffisantes pour déterminer $\varphi(x)$. Mais il n'en est plus ainsi lorsque les observations sont assez bien faites pour que les carrés et les produits des erreurs soient négligeables; dans ce cas, comme le démontre M. Bertrand, on a

$$(7)\qquad \varphi(x) = \frac{k}{\sqrt{\pi}} e^{-k^2 x^2}.$$

La dernière condition énoncée est en général remplie dans la pratique; c'est ce qui explique pourquoi cette formule se trouve si bien justifiée par ses conséquences.

Cette loi de probabilité des erreurs fortuites a été donnée pour la première fois par Gauss, dans sa *Théorie du mouvement des corps célestes*. La méthode des moindres carrés en est un corollaire immédiat. Dans le cas particulier où il s'agit d'une même grandeur X dont on fait n mesures x_1, x_2, ..., x_n, la méthode donne la moyenne arithmétique

$$\frac{x_1 + x_2 + \ldots + x_n}{n},$$

pour la valeur la plus probable de X.

C'est en quelque sorte la marche inverse de la précédente que Gauss a suivie pour établir la formule (7). Il prend pour point de départ cette proposition regardée comme un axiome : la moyenne arithmétique entre plusieurs mesures est la valeur la plus probable.

Après avoir rapporté la démonstration de Gauss, M. Bertrand se pose une question intéressante et qu'il résout avec beaucoup de bonheur : si, au lieu de prendre la moyenne arithmétique pour valeur la plus probable, on attribuait cette propriété à une autre fonction

$$\psi(x_1, x_2, \ldots, x_n)$$

des n mesures, existerait-il une loi correspondante de probabilité des erreurs? La réponse est négative; elle s'exprime élégamment : l'existence d'une loi de probabilité impose à la fonction ψ une forme telle que, si l'on donne à toutes les mesures prises, $x_1, x_2, \ldots, x_n$, un même accroissement quelconque α, la fonction croisse elle-même de α.

Pour compléter l'étude de la loi de Gauss, il faut indiquer la signification du paramètre k, apprendre à calculer pour chaque série d'observations la valeur correspondante de ce paramètre, enfin montrer comment on a pu vérifier l'exactitude de la formule (7).

Quand on a effectué sur un angle deux séries de mesures, en changeant d'instrument d'une série à l'autre, si une erreur d'une minute dans le premier système est aussi probable qu'une erreur d'une seconde dans l'autre, chacun s'accordera à reconnaître que le second système de mesures est 60 fois plus précis que le premier. Nous dirons donc, d'après cela et d'une manière générale, que la précision P d'un système de mesures est α fois plus grande que la précision P′ d'un autre système, si la probabilité d'une erreur comprise entre z et $z + dz$ pour une mesure du premier système est égale à la probabilité d'une erreur comprise entre αz et $\alpha z + d.\alpha z$ dans une mesure du second. La condition

$$\frac{k}{\sqrt{\pi}} e^{-k^2 z^2} dz = \frac{k'}{\sqrt{\pi}} e^{-k'^2 \alpha^2 z^2} \alpha \, dz,$$

exige $k = \alpha k'$ et, par suite,

$$\frac{\mathrm{P}}{\mathrm{P}'} = \alpha = \frac{k}{k'}.$$

La précision d'un système d'observations sera donc mesurée par la valeur correspondante k du paramètre, si l'on prend pour unité de précision celle d'un système dont le paramètre serait égal à 1.

Voyons maintenant comment on peut *calculer la valeur de k qui répond à un système de mesures.*

Plaçons-nous d'abord dans le cas où les mesures portent sur une grandeur dont la valeur serait exactement connue.

On calculera alors en fonction de k l'expression $f(k)$ de la valeur probable d'une puissance quelconque x^i de l'erreur; et, comme, en vertu du théorème de Bernoulli, sur un grand nombre n d'épreuves, la valeur probable d'une quantité diffère peu de la moyenne arithmétique, si l'on désigne par

$$e_1, e_2, \ldots, e_n$$

les erreurs successivement commises dans les n observations, on aura la relation approximative

$$\frac{e_1^i + e_2^i + \ldots + e_n^i}{n} = f(k),$$

d'où l'on déduira k. Si, pour abréger l'écriture, on représente par S_i le numérateur du premier membre, on trouve pour k les expressions

$$\frac{n}{S_1\sqrt{\pi}}, \quad \sqrt{\frac{n}{2S_2}}, \quad \sqrt[3]{\frac{n}{S_3\sqrt{\pi}}}, \quad \sqrt[4]{\frac{3n}{4S_4}}, \quad \ldots,$$

dont la concordance constitue une véritable vérification de la loi de probabilité des erreurs. Les deux premières sont les plus simples, et l'on démontre qu'entre les deux c'est la seconde

$$(8) \qquad \frac{1}{2k^2} = \frac{S_2}{n}$$

qui mérite le plus de crédit. M. Bertrand fait connaître plusieurs autres formules pouvant remplacer les précédentes, et parmi lesquelles nous signalerons celle-ci :

$$\frac{1}{2k^2} = \frac{S_2}{n+2},$$

qui est analogue à (8) et probablement plus exacte.

Voici encore, pouvant servir au même objet, plusieurs résultats bien dignes de remarque à cause de leur simplicité. Si l'on groupe les observations deux à deux, en chargeant le hasard de les associer, et que dans chaque groupe on choisisse la plus grande des deux erreurs, M. Bertrand trouve, pour les valeurs probables de cette erreur et de son carré,

$$\frac{\sqrt{2}}{k\sqrt{\pi}} \quad \text{et} \quad \frac{1}{2k^2}\left(1 - \frac{2}{\pi}\right).$$

Si l'on groupait au hasard les erreurs trois par trois, la valeur probable du carré de la plus grande des trois serait

$$\frac{1}{2k^2}\left(1+\frac{2\sqrt{3}}{\pi}\right),$$

etc.

Les formules que nous venons d'indiquer pour le calcul de k renferment les erreurs vraies, puisque nous avons supposé que la grandeur considérée était exactement connue; dans la pratique, la grandeur considérée X n'est connue que par les mesures $x_1, x_2, \ldots, x_n$ qu'on en a prises. Comment alors calculer k? Tout simplement en substituant, dans l'une des formules précédentes, aux erreurs vraies qu'on ne connaît pas, les erreurs présumées, c'est-à-dire les différences entre les mesures x_1, $x_2, \ldots, x_n$ et leur moyenne arithmétique ξ qui est la valeur la plus probable de X. Ainsi, à la formule

$$(8)\qquad \frac{1}{2k^2}=\frac{(X-x_1)^2+\ldots+(X-x_n)^2}{n},$$

on substituera la suivante

$$(9)\qquad \frac{1}{2k_1^2}=\frac{(\xi-x_1)^2+\ldots+(\xi-x_n)^2}{n},$$

qui déterminera approximativement le paramètre.

Mais on peut faire un peu mieux. On déduit de (8) et (9), par un calcul facile, la relation

$$\frac{1}{2k_1^2}=\frac{1}{2k^2}\,\frac{n-1}{n}-\frac{2}{n}\sum e_i e_{i'},$$

où e_i représente l'erreur vraie $X-x_i$ et d'où l'on déduit, en négligeant le dernier terme, dont la valeur probable est nulle,

$$\frac{1}{2k_1^2}=\frac{1}{2k^2}\,\frac{n-1}{n}.$$

D'après cela, on déterminera finalement le paramètre k par la formule

$$(10)\qquad \frac{1}{2k^2}=\frac{(\xi-x_1)^2+\ldots+(\xi-x_n)^2}{n-1},$$

qui, depuis Gauss, est adoptée.

M. Bertrand ne pouvait manquer de signaler le point faible de la démonstration. « Une grandeur », dit-il, « dont la valeur probable est petite, est certainement petite elle-même quand elle est essentiellement positive; mais, quand elle peut, comme ici, changer de signe, la valeur probable étant nulle, cela prouve seulement que les valeurs positives ont même probabilité que les négatives. » Cela est parfaitement vrai; mais il n'en résulte pas non plus qu'elles soient fort grandes. C'est à l'expérience à décider; elle donne raison à Gauss et à ses adeptes. D'ailleurs, la formule (8) n'est pas mathématiquement rigoureuse, et si l'on compare la formule (10) à la suivante

$$(11) \qquad \frac{1}{2k^2} = \frac{(\xi - x_1)^2 + \ldots + (\xi - x_n)^2}{n},$$

que l'on déduit immédiatement de (8) en substituant les erreurs présumées aux erreurs vraies, on voit que la formule (11) fournira pour k une valeur plus grande que la formule (10); on risque donc moins en choisissant la relation (10), qui attribue au résultat une précision moindre. Enfin, dans le cas extrême où l'on ne mesurerait la grandeur qu'une fois, la formule (11), ayant alors son numérateur nul et son dénominateur égal à 1, est absurde : elle accuse une précision infinie alors qu'on n'a aucune donnée sur la précision; la formule (10) est plus logique, elle donne $\frac{0}{0}$.

Nous avons indiqué comment la comparaison des diverses expressions que l'on peut employer pour calculer k permettait de vérifier la loi de Gauss. La vérification directe a été entreprise, la première fois, par Bessel; elle est fondée sur ce que, dans une série suffisamment nombreuse d'observations, si l'on partage les erreurs en groupes bien définis, chaque groupe se présentera, en vertu du théorème de Bernoulli, un nombre de fois à peu près proportionnel à sa probabilité. Bessel a étudié 470 observations de Bradley portant sur les coordonnées, aujourd'hui bien connues, d'une même étoile. Après avoir réduit soigneusement les mesures de ce « modèle des observateurs », Bessel a classé les erreurs par ordre de grandeur; il a compté le nombre des erreurs, tant positives que négatives, qui étaient renfermées dans chacun des intervalles $0^s,0$ et $0^s,1$, $0^s,1$ et $0^s,2$, ..., et il a comparé les nombres ainsi obtenus avec les nombres probables que donne, pour les erreurs de chaque groupe, la formule de Gauss. L'accord est plus satisfaisant qu'on ne l'eût espéré. L'épreuve a été depuis renouvelée bien des fois.

et toujours avec le même succès, pour diverses séries d'observations préalablement purgées de toute erreur systématique.

La loi de probabilité des erreurs permet de prévoir, lorsque les épreuves sont nombreuses, tous les détails relatifs à la série des erreurs commises. M. Bertrand en donne des exemples intéressants.

Nous signalerons particulièrement celui-ci :

On a mesuré n fois une grandeur : tout est régulier; après avoir déterminé k, on calcule l'erreur λ telle que la probabilité pour qu'une erreur soit inférieure à λ ait une valeur donnée p; puis on rejette toutes les mesures dont les différences avec la moyenne sont moindres que λ, et l'on prend pour valeur approchée de la grandeur, au lieu de la moyenne générale, la moyenne des mesures réservées, dont le nombre m diffère peu de np. Y a-t-il avantage à opérer de la sorte? La réponse est affirmative en supposant, bien entendu, que m soit un grand nombre. M. Bertrand calcule la valeur probable du carré de l'erreur, et il trouve que, dans le second cas, elle est multipliée par un facteur qui décroît sans limite quand λ diminue.

VIII. — Le tir a la cible.

A la loi de probabilité des erreurs fortuites se rattache la formule de Bravais sur les *erreurs de situation d'un point*.

La probabilité d'une erreur comprise entre u et $u+du$ pour l'abscisse et entre v et $v+dv$ pour l'ordonnée d'un point a pour expression

$$(12) \qquad P = \frac{\sqrt{k^2k'^2 - \lambda^2}}{\pi} e^{-(k^2u^2+2\lambda uv+k'^2v^2)} du\, dv.$$

On est amené à cette formule par l'application de la loi de Gauss, suivie de quelques transformations analytiques. On en déduit cette proposition, énoncée en 1709 par Cotes : si plusieurs positions ont été obtenues, qui méritent la même confiance, la position la plus probable du point sera le centre des moyennes distances des positions obtenues.

On peut, inversement, prendre pour point de départ le théorème de Cotes considéré comme un postulatum et remonter à la formule (12).

Cette formule s'applique à la probabilité des écarts dans le tir à la cible.

La première question à résoudre est alors la détermination des valeurs des constantes k, k', λ pour un certain tireur et pour une arme donnée. On y parvient de la manière suivante : désignons par u et v les coordonnées du point où frappe la balle par rapport à deux axes passant par le centre des moyennes distances de tous les points frappés (u_1, v_1), (u_2, v_2), ..., (u_n, v_n), supposés en très grand nombre ; on exprimera en fonction des constantes k, k', λ les valeurs probables de u^2, v^2 et de uv ; puis on égalera ces expressions aux quantités

$$A = \frac{1}{n}(u_1^2 + \ldots + u_n^2),$$

$$B = \frac{1}{n}(v_1^2 + \ldots + v_n^2),$$

$$C = \frac{1}{n}(u_1 v_1 + \ldots + u_n v_n);$$

on aura de la sorte, en vertu du théorème de Bernoulli, des relations suffisamment approximatives, d'où l'on tirera les valeurs

$$\frac{k^2}{B} = \frac{k'^2}{A} = -\frac{\lambda}{C} = \frac{1}{2(AC - B^2)}$$

des constantes.

Cela fait, si l'on pose

$$k^2 u^2 + 2\lambda uv + u'^2 v^2 = H,$$

la probabilité pour que, pour l'arme et le tireur considérés, une balle tombe entre les deux ellipses correspondant à H et $H + dH$ aura pour expression

$$e^{-H^2} dH.$$

Par suite, la probabilité pour que la balle frappe en dehors de l'ellipse correspondant à une valeur H_1 du paramètre H sera e^{-H_1}.

Les neuf ellipses semblables, correspondant aux neuf valeurs

0,10536	0,22315	0,35669	0,51082
0,69315	0,91329	1,20677	1,60944
2,30359			

du paramètre H, partageront le plan en 10 régions, contenant très vraisemblablement chacune le dixième du nombre des balles tirées. La première de ces régions est la plus petite de ces ellipses; la dernière est la portion indéfinie du plan située au delà de la plus grande ellipse.

M. Bertrand a appliqué les résultats précédents à l'examen de 1000 coups tirés par des tireurs habiles, à 200^m de distance, avec dix armes de même modèle, chaque tireur tirant 10 coups avec chaque arme. Les nombres de balles, qui devaient être théoriquement tous égaux à 100 pour chacune des dix régions, ont été trouvés égaux à

$$99, 106, 100, 108, 100, 115, 89, 94, 90, 97.$$

L'accord, comme on le voit, est assez satisfaisant pour confirmer la théorie.

IX. — La méthode des moindres carrés.

Dans le cas général, au lieu de n mesures d'une même grandeur, on a à considérer les valeurs $r_1, r_2, \ldots, r_n$, fournies par l'observation, de n fonctions

$$f_1(\mathrm{U}, \mathrm{V}, \mathrm{W}, \ldots), \quad f_2(\mathrm{U}, \mathrm{V}, \mathrm{W}, \ldots), \quad \ldots, \quad f_n(\mathrm{U}, \mathrm{V}, \mathrm{W}, \ldots),$$

de m inconnues U, V, W, ..., m étant plus grand que n. L'application immédiate du principe de la méthode des moindres carrés, c'est-à-dire la recherche des valeurs $u, v, w, \ldots$, qui rendent minimum la somme des carrés des résidus

$$\begin{gathered} f_1(u, v, w, \ldots) - r_1, \\ f_2(u, v, w, \ldots) - r_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n(u, v, w, \ldots) - r_n, \end{gathered}$$

conduirait à des calculs inextricables sans une circonstance heureuse qui se présente toujours dans la pratique, et qui consiste en ce que l'on connaît d'avance des valeurs approchées A, B, C, ... des inconnues U, V, W, ..., telles que l'on soit en droit de négliger les carrés et les produits des différences U — A, V — B, W — C, ..., que nous désignerons respectivement par X, Y, Z, ...; on a alors, en négligeant x^2, y^2, ..., xy, xz, ...,

$$f_i(U, V, W, \ldots) - f_i(A, B, C) = X\frac{\partial f_i}{\partial A} + Y\frac{\partial f_i}{\partial B} + Z\frac{\partial f_i}{\partial C} + \ldots,$$

et comme $f_i(U, V, W, \ldots)$ est donné par l'observation et que $f_i(A, B, C, \ldots)$ et ses dérivées sont immédiatement calculables, le problème est ramené à la recherche des valeurs de x, y, z, ... qui rendent minimum la somme des carrés des n fonctions linéaires

$$(13)\qquad \left\{\begin{array}{l} a_1x + b_1y + c_1z + \ldots - l_1, \\ a_2x + b_2y + c_2z + \ldots - l_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_nx + b_ny + c_nz + \ldots - l_n, \end{array}\right.$$

où les coefficients a, b, c, ... sont exactement connus et où les quantités l résultent de l'observation.

Une règle bien connue du Calcul différentiel donne, pour déterminer ces valeurs, le système des m équations à m inconnues

$$(14)\qquad \left\{\begin{array}{l} [aa]x + [ab]y + [ac]z + \ldots = [al], \\ [ab]x + [bb]y + [bc]z + \ldots = [bl], \\ [ac]x + [bc]y + [cc]z + \ldots = [cl], \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

que l'on nomme *équations normales* et dans lesquelles la notation $[uv]$, empruntée à Gauss, désigne la somme

$$u_1v_1 + u_2v_2 + \ldots + u_nv_n.$$

La règle pour former les équations normales est fort simple : on égale à zéro la somme des produits obtenus en multipliant chacune des fonctions (13) par le coefficient de x dans cette fonction; on opère de même pour y, z, ...

Il ne suffit pas de résoudre les équations normales, il faut encore déterminer le degré de confiance à accorder aux valeurs ainsi obtenues.

Nous avons vu qu'à toute série de mesures répondait une valeur de paramètre h qui figure dans la loi des erreurs fortuites et que la précision d'une mesure du système considéré était proportionnelle à h. Si h est proportionnel au degré de confiance, son inverse est au contraire proportionnel à l'erreur à craindre; mais nous savons aussi que la valeur probable du carré de l'erreur a pour expression

$$\frac{1}{2h^2}.$$

On est ainsi conduit à donner le nom d'*erreur à craindre* à la racine carrée $\frac{1}{h\sqrt{2}}$ de la valeur probable du carré de l'erreur; c'est l'erreur à craindre sur chacune des valeurs trouvées pour les inconnues, qu'il convient de calculer.

On s'appuie pour cela sur un théorème dû à Gauss et qui, à cause de son emploi fréquent, mérite d'être rapporté, bien qu'il ne soit pas énoncé dans le livre de M. Bertrand.

Si U, V, W, ... sont des grandeurs indépendantes entre elles et si u, v, w, ... désignent respectivement les erreurs à craindre sur chacune d'elles, l'erreur à craindre E sur une fonction quelconque f(A, B, C, ...) de ces quantités a pour expression

$$(15)\qquad E = \sqrt{\left(\frac{\partial f}{\partial U}\right)^2 u^2 + \left(\frac{\partial f}{\partial V}\right)^2 v^2 + \left(\frac{\partial f}{\partial W}\right)^2 w^2 + \ldots}.$$

En appliquant cette règle à l'expression de x en fonction de l_1, l_2, ..., l_n que donne la résolution des équations normales, on trouve, pour l'erreur à craindre E_x sur l'inconnue x,

$$E_x = \varepsilon\sqrt{x'}.$$

Dans cette formule, ε désigne l'erreur à craindre sur les mesures l_1, l_2, ..., l_n, et x' est la valeur que donnerait pour x la résolution du système obtenu en remplaçant respectivement, dans les équations normales, le second membre de la première par 1 et le second membre de chacune des autres par zéro. Pour y, c'est dans la deuxième équation qu'il faut remplacer le second membre par 1; pour z, c'est dans la troisième, etc.

Quant à ε, il peut se faire qu'il soit donné *a priori* par la connaissance de l'instrument employé; mais souvent la précision des mesures l_1, ..., l_n est inconnue, et l'on est condamné à l'évaluer approximativement *a posteriori* d'après les valeurs fournies par la méthode des moindres carrés pour les inconnues. On adopte alors pour ε la valeur donnée par la formule

$$(16)\qquad \varepsilon = \sqrt{\frac{[\Delta\Delta]}{n-m}}.$$

où $[\Delta\Delta]$ représente la somme des carrés des erreurs présumées des observations, c'est-à-dire la somme des carrés des valeurs que prennent les fonctions linéaires (13), lorsqu'on y remplace $x, y, z, \ldots$ par les valeurs fournies par la méthode des moindres carrés.

Cette formule (16) est la généralisation de la formule (10) relative au cas où il n'y a qu'une inconnue. Sa démonstration comporte une objection déjà indiquée à propos de la formule (10). La critique est fondée : M. Bertrand montre même qu'un raisonnement analogue permettrait d'obtenir pour le même objet une infinité d'autres formules infiniment différentes; mais il ne faut pas confondre l'exactitude de la démonstration avec celle de la formule, et l'on peut sans témérité donner la préférence à la relation (16), proposée par Gauss et adoptée depuis par les calculateurs les plus compétents.

Dans un Mémoire dont certaines parties sont restées classiques et qui a pour titre : *Theoria combinationis observationum erroribus minimi obnoxiæ*, Gauss, délaissant le point de vue auquel il s'était placé dans le *Theoria motus,* a voulu affranchir l'exposition de la méthode des moindres carrés de la connaissance de la loi de probabilité des erreurs fortuites.

Sans rien préjuger sur la forme de la fonction $\varphi(z)$ qui représente la facilité de l'erreur z, Gauss donne le nom d'*erreur à craindre* à la quantité E définie par la relation

$$E^2 = \int_{-\infty}^{+\infty} z^2\,\varphi(z)\,dz,$$

c'est-à-dire à la racine carrée de la valeur probable du carré de l'erreur z. Il est clair que, si, dans un système d'observations, la valeur de E est petite, toute erreur qui n'est pas petite en valeur absolue a une probabilité très faible, sans quoi la valeur probable du carré de l'erreur ne serait pas très petite. Si, au contraire, E est grand, l'existence de grandes erreurs peut être supposée, leur probabilité n'est pas petite. On peut donc regarder *a priori* la valeur de E comme donnant la mesure du degré de confiance à accorder au système d'observations considéré.

En partant de cette idée, Gauss a été conduit à regarder comme le meilleur, non plus le système de valeurs de x, y, z, ... dont la probabilité est maxima, mais le système de valeurs de x, y, z, ... pour lesquelles l'erreur à craindre est minima. C'est un postulatum fort plausible, mais c'est un postulatum ; Gauss d'ailleurs ne le dissimule pas : « Quod si quis hanc rationem pro arbitrio, nulla cogente necessitate, electam esse objiciat, lubenter assentiemus. Quippe questio hæc per rei naturam aliquid vagi implicat, quod limitibus circonscribi nisi per principium aliquatenus arbitrarium nequet. »

La seconde théorie s'accorde d'ailleurs avec la première; elle conduit à la méthode des moindres carrés et, en particulier, à la règle de la moyenne arithmétique dans le cas des mesures réitérées d'une même grandeur. Les calculs de Gauss sont d'une élégance extrême; les principes de l'Algèbre élémentaire sont seuls mis à contribution et toutes les formules sont préparées d'une manière explicite pour le calculateur, qui n'a plus qu'à les mettre en nombres sans leur faire subir aucune transformation.

Il nous reste à dire encore quelques mots sur la moyenne arithmétique. La formule (15) montre que l'erreur à craindre E sur la moyenne arithmétique de n mesures est égale à l'erreur à craindre ε sur chacune d'elles divisée par $\sqrt{n}$. Il résulte d'ailleurs de la formule (10) que ε^2 est approximativement égal à la somme $[\Delta\Delta]$ des carrés des erreurs présumées divisées par $n-1$. On a donc finalement

$$E = \frac{[\Delta\Delta]}{\sqrt{n(n-1)}}.$$

Lorsqu'une même grandeur X a été mesurée par des procédés différents ou par divers observateurs munis d'instruments d'inégale perfection, on ne doit plus prendre la moyenne arithmétique. Soient x_1, x_2, ..., x_n les mesures obtenues, et ε_1, ε_2, ..., ε_n les erreurs à craindre correspondantes. Si l'on cherche parmi les expressions de la forme

$$X = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n,$$

où les λ doivent avoir une somme égale à l'unité, sans quoi la valeur à adopter ne serait pas exacte dans le cas où toutes les mesures le seraient, si l'on cherche, disons-nous, celle pour laquelle l'erreur à craindre est minimum, on trouve pour la valeur à adopter

$$X = \frac{\frac{x_1}{\varepsilon_1^2} + \frac{x_2}{\varepsilon_2^2} + \ldots + \frac{x_n}{\varepsilon_n^2}}{\frac{1}{\varepsilon_1^2} + \frac{1}{\varepsilon_2^2} + \ldots + \frac{1}{\varepsilon_n^2}},$$

c'est la moyenne, prise en comptant les mesures $x_1, \ldots, x_n$ comme répétées respectivement un nombre de fois égal à l'inverse du carré de l'erreur à craindre correspondante.

Le facteur $\frac{1}{\varepsilon^2}$ se nomme le *poids* de l'observation correspondante x. On voit par la formule ci-dessus que r observations semblables équivalent à une seule observation dont le poids serait r fois plus grand.

Cette notion de *poids*, telle que nous venons de l'établir, est indépendante de la forme de la loi des erreurs fortuites. Quand on considère la loi de Gauss, le poids est proportionnel à k^2. La précision ne saurait, au contraire, être définie avec rigueur que si la loi de probabilité revêt une forme spéciale; M. Bertrand démontre que la facilité de l'erreur $\varphi(z)$ doit avoir pour expression

$$C e^{-az^{\mu+1}},$$

où $\mu + 1$ est pair. On tombe sur la loi de Gauss pour $\mu = 1$.

X. — Les fausses applications du Calcul des probabilités. Conclusions.

Les deux derniers Chapitres se rapportent à la Statistique et aux probabilités des décisions judiciaires. Ils sont relativement fort courts; leur véritable objet est d'achever de mettre en pleine lumière cette pensée originale qui est l'idée maîtresse de l'œuvre : l'application du Calcul des probabilités n'est légitime que pour les événements fortuits assimilables à une série de tirages dans une urne; autant vaut l'assimilation, autant valent les conséquences.

C'est pour avoir méconnu ou plutôt ignoré cette vérité fondamentale que tant de bons esprits se sont livrés à des recherches chimériques et se sont engagés dans des routes sans issue.

Une première condition, évidemment nécessaire pour l'assimilation, est l'invariabilité approchée du rapport entre le nombre des événements et le nombre total des épreuves; mais elle ne suffit pas. Ce rapport constant ne fait connaître que la composition de l'urne; il faut encore que les probabilités d'écart soient les mêmes.

Prenons un exemple :

Les Tables de mortalité montrent que sur 10000 individus âgés de 30 ans, 500 environ survivent 35 ans après. D'autre part, si dans une urne qui contient une boule blanche et une boule noire, on effectue 10000 tirages en remettant chaque fois la boule sortie, on amènera environ 5000 fois la boule blanche. Y a-t-il assimilation? Non certes. Les nombres comparés diffèrent peu de 5000, voilà tout. Mais, dans le premier cas, l'écart est complètement inconnu; dans l'autre, il est soumis à des lois précises. On peut affirmer que, si l'on répète un grand nombre de fois la série des 10000 tirages, la valeur moyenne de l'écart sera 40; mais, pour un grand nombre de groupes de 10000 hommes, la moyenne générale des décès étant égale à 5000, qui oserait affirmer que la moyenne des écarts n'atteindra pas 100?

L'assimilation la plus téméraire qu'on ait jamais faite est assurément celle des jurés à des urnes. Ni les principes ni les conclusions de Condorcet ne sont acceptables. Aussi voit-on tour à tour Laplace rejeter les résultats de Condorcet, Poisson écarter ceux de Laplace, tandis que Arago ne craint pas d'affirmer à la Chambre des Députés, à propos d'une loi sur le jury, que les nombres de Laplace sont aussi certains que la parallaxe du Soleil! Cournot va même jusqu'à calculer, pour un tribunal de trois juges, la probabilité pour chacun d'eux de ne pas se tromper dans une cause qui leur est soumise!

M. Bertrand fait bonne justice de ces étranges tentatives. Le récit est piquant; mais on ne voit pas sans tristesse des hommes éminents, des savants distingués, se livrer ainsi à des calculs stériles dont Stuart Mill a pu dire, non sans raison, qu'ils étaient le scandale des Mathématiques.

Notre tâche est terminée. Nous l'aurions remplie à notre satisfaction, si nous étions parvenu à inspirer à nos lecteurs le désir de mieux connaître ce Livre si remarquable par l'élévation des idées, par la finesse des aperçus, par une critique ferme et incisive.

Quelques passages eussent gagné à être un peu condensés. Mais, si c'est là un défaut, il se trouve largement compensé par les termes charmants en lesquels ces choses-là sont mises. L'agrément du style n'est pas la moindre qualité de l'Ouvrage, et nous ne saurions résister au plaisir d'ajouter, en finissant, quelques extraits :

« S'il pleut un jour entier sur la place du Carrousel, tous les pavés seront également mouillés. Sous une forme simplifiée, mais sans en rien retrancher, c'est là le théorème de Bernoulli. Il pourrait se faire assurément, lorsque tout autour la pluie tombe à torrents, qu'un certain pavé restât sec. Aucune goutte n'a pour lui de destination précise; le hasard les disperse, il peut les porter toutes sur les pavés voisins, personne ne le supposera sérieusement. Le hasard a des caprises; jamais on ne lui vit d'habitudes. Si mille gouttes tombent sur mille pavés, chaque pavé n'aura pas la sienne; s'il en tombe mille millions, chaque pavé recevra son million ou bien peu s'en faudra.

....

» Les grands nombres régularisent tout. La moyenne de tous les écarts peut être prédite avec confiance. La même certitude s'attache à la moyenne des carrés des écarts, de leurs cubes, de leurs quatrièmes puissances... ».

« Beaucoup de joueurs, préoccupés de cette régularité nécessaire dans les moyennes, cherchent, dans les coups qui précèdent celui qu'ils vont jouer, une indication et un conseil. Ce n'est pas bien entendre les principes. La Science, à ces chimères, ne reste pas sans réponse. Mais la décision du bon sens suffit; elle est nette et claire: à quoi bon la traduire en Algèbre? Le préjugé est opiniâtre; les géomètres perdraient, à le combattre, leur temps et leurs formules.

» L'illusion repose sur un sophisme; on allègue la loi de Bernoulli comme certaine; elle n'est que probable. Sur 20000 épreuves, dit-on, à la roulette, la noire ne peut pas sortir plus de 10500 fois; l'assertion de la Science est formelle. Si les 10000 premières parties ont donné 6000 noires, les 10000 suivantes ont donc contracté une dette envers la rouge. On fait trop d'honneur à la roulette; elle n'a ni conscience ni mémoire. En supposant qu'à une rencontre inouïe succédera, pour la séparer, un nouvel écart de la règle, on n'efface pas l'invraisemblance, on la redouble. »

EUGÈNE ROUCHÉ.