

Toute reproduction pour publication ou à des fins commerciales, de la totalité ou d'une partie de l'article, devra impérativement faire l'objet d'un accord préalable avec l'éditeur. Toute reproduction à des fins privées, ou strictement pédagogiques dans le cadre limité d'une formation, de la totalité ou d'une partie de l'article, est autorisée sous réserve de la mention explicite des références éditoriales de l'article (titre, auteur, éditeur, copyright, pages extraites) et éventuellement, de la déclaration au Centre français d'exploitation du droit de copie (www.cfcopies.com) conformément à la législation en vigueur.

LA LOI DES GRANDS NOMBRES

Le théorème de De Moivre-Laplace

Denis Lanier

Didier Trotoux

IREM de Basse-Normandie

"Pour beaucoup d'esprits, une probabilité calculée à sept ou huit décimales près est beaucoup plus convaincante qu'un argument fondé sur des considérations qualitatives. Ces esprits oublient que si le calcul en question est fondé sur des éléments statistiques, qui ne sont donc pas numériquement précis, le nombre de décimales est une pure illusion. Ce nombre fait croire à la précision alors qu'elle n'existe pas. De ce point de vue la théorie des probabilités est fondamentalement une imposture."

René Thom¹

Le couple fréquence-probabilité, ainsi que la théorie instituant ce rapport qu'on peut appeler schématiquement "loi des grands nombres", est un leitmotiv de la période classique de l'histoire du calcul des probabilités. Il est au coeur du développement de la théorie et des préoccupations des probabilistes, comme de ses utilisateurs. Son sort ne paraît pas entièrement réglé de nos jours, comme le montre la citation ci-dessus de René Thom. D'autre part des problèmes pédagogiques s'ajoutent aux problèmes théoriques. Les récents programmes des lycées imposent de prendre une approche fréquentiste pour définir une probabilité. Cela pose le problème du statut de ces énoncés que l'on rassemble sous le nom de "loi des grands nombres". Peu de propositions mathématiques portent ce titre de "loi". Est-ce un

¹ Entretien avec René Thom, *Le Monde*, 22-23 janvier 1995.

théorème, comme il est utilisé habituellement pour le "théorème de De Moivre-Laplace ? Est-ce un énoncé extra-mathématique, admis comme prémisse à toute théorie scientifique ? C'est Poisson qui utilise pour la première fois l'expression "loi des grands nombres" dans ses *Recherches sur la probabilité des jugements* (1837) :

"Les choses de toutes natures sont soumises à une loi universelle qu'on peut appeler *la loi des grands nombres*."

Mais le même Poisson, après avoir précisé en quoi consiste cette loi, s'empresse de la "démontrer" ! D'une certaine manière, pour pouvoir faire des probabilités (et en particulier les appliquer), tout se passe comme si on devait admettre que la nature suit (au moins localement) des lois, qui permettent l'élaboration d'une théorie, à l'intérieur de laquelle on démontre ensuite cette loi. Le rapprochement avec la situation du physicien, pour embarrassante qu'elle soit pour certains mathématiciens, est flagrante.

Il nous semble, d'un point de vue historique et aussi d'un point de vue pédagogique, qu'il est un peu vain de s'en tenir aux différentes définitions du concept de probabilité, en recherchant la meilleure. En revanche l'étude historique de sa mise en oeuvre et de ses problématiques nous paraît riche d'enseignements. En ce qui concerne la "loi des grands nombres", depuis son énoncé initial par Jacques Bernoulli jusqu'à la résolution du problème limite classique par Kolmogorov, le travail des probabilistes consiste à chercher à comprendre l'énoncé, à le préciser, à évaluer la portée, à tenter de simplifier la démonstration. Du point de vue de l'histoire interne des mathématiques, il y a là, pendant deux siècles, un travail sur le texte lui-même qui est tout à fait fascinant. On verra dans la suite que ces questions ne se posaient que si l'on voulait appliquer les probabilités à d'autres domaines que les jeux de hasard. On verra aussi qu'elles recoupaient initialement des préoccupations métaphysiques - si la nature suit la loi des grands nombres, c'est qu'il y a quelqu'un qui a fabriqué le dé, d'où l'existence de Dieu - qu'on se garde bien d'aborder maintenant dans nos classes.

L'histoire de la loi des grands nombres est jalonnée par quelques grands textes sans oublier que l'histoire ne s'arrête pas là.

Les trois premiers pas sont marqués par Jacques Bernoulli et *l'Ars Conjectandi* (1713), Abraham De Moivre et la *Doctrine of Chances* (1756), Pierre-Simon de Laplace et la *Théorie Analytique des Probabilités* (1812). Nous avons choisi de centrer le travail de l'atelier sur le texte, moins connu, de De Moivre et sa reprise par Laplace. On trouvera en annexe des extraits des ouvrages de De Moivre et Laplace.

*
* *

I. BERNOULLI.

L'énoncé et la démonstration du théorème de Jacques Bernoulli ont été présentés de façon très complète par Norbert Meusnier². Nous nous contenterons d'en faire ici un rapide résumé.

On sait que le Calcul des Probabilités s'est constitué dans la deuxième moitié du 17^{ème} siècle autour de questions de jeux de hasard et de "l'espérance de gain" qu'on peut y obtenir (Pascal, Fermat, Huygens)³. Huygens pose le premier - si l'on excepte le "pari" de Pascal - le problème de l'application du nouveau calcul à d'autres domaines que celui des jeux de hasard. Il travaille ainsi sur l'espérance de vie humaine, à partir de données récoltées à Londres, en liaison avec des problèmes de rentes et annuités. Mais c'est Jacques Bernoulli qui va explicitement vouloir fonder cette application du calcul des probabilités aux affaires civiles, morales et économiques.

L'Ars Conjectandi paraît à Bâle en août 1713, huit ans après la mort de son auteur. Ce dernier y avait travaillé depuis 1685, à la suite de sa lecture du traité de Huygens. Pendant ces vingt ans de recherche, Jacques Bernoulli opère un glissement progressif du calcul des espérances de Pascal et Huygens vers celui des probabilités. Il finit par définir la probabilité comme le degré de certitude avec lequel un événement futur peut éventuellement se produire. *L'Ars Conjectandi* comporte, dans sa forme publiée, quatre parties. La première est une reprise annotée et commentée du traité de Huygens de 1656-7. La deuxième partie est une compilation systématique de résultats sur les permutations et combinaisons. La troisième partie applique les résultats précédents aux jeux de hasard. Enfin la quatrième partie envisage l'extension de la théorie à d'autres domaines, extension fondée sur l'énoncé et la démonstration du fameux théorème. L'importance de ce résultat n'échappe pas à Bernoulli qui écrit :

"J'estime cette invention bien davantage que si j'avais livré la quadrature même du cercle, car si celle-ci était effectivement trouvée son utilité serait peu considérable."⁴

² Norbert Meusnier, *Argumentation et démonstration : à quoi sert la démonstration de la "loi des grands nombres" de Jacques Bernoulli*, in Actes du colloque Inter-Irem de Besançon, La démonstration mathématique dans l'histoire, Irem de Besançon, 1990.

³ Voir par exemple : Denis Lanier, *La Géométrie du hasard, L'espérance du Hollandais*, in SCHOLIES, Actes du séminaire interdisciplinaire d'histoire des Sciences du Lycée Malherbe de Caen, n°16, juin 1995

⁴ cité par N. Meusnier in Actes du colloque Inter-Irem de Besançon, La démonstration mathématique dans l'histoire, Irem de Besançon, 1990.

L'objectif de Bernoulli est d'étudier les arguments que l'on peut utiliser dans l'art de conjecturer, comment estimer leurs poids pour supputer les probabilités. Il est ainsi amené à distinguer dans le chapitre IV⁵ de cette 4ème partie deux types de situations aléatoires.

La première situation est celle des jeux de hasard, où les probabilités sont connues "a priori" par les données plus ou moins explicites de l'expérience aléatoire. Bernoulli décrit un modèle permettant de décrire ce type de situation. Il s'agit du modèle de l'urne, valable pour toutes les situations des plus simples (lancés de dés, tirage de cartes, ...) aux plus complexes. En supposant - le plus souvent implicitement - l'équiprobabilité, on se ramène alors à un problème de dénombrement de cas.

La seconde situation est celle des expériences aléatoires où la probabilité sera définie "a posteriori". Ce sont les expériences menées en grand nombre qui permettent d'estimer la probabilité. Avant de prendre des exemples tirés d'autres domaines (maladies, ...), Bernoulli reprend le modèle de l'urne, mais en supposant qu'on n'en connaît pas la composition. Bernoulli sous-entend donc que cet "autre chemin" pour définir les probabilités suppose des conditions sur l'expérience aléatoire réalisée. Cela revient en fait à se ramener à la première situation, mais dans le cas où l'on ignore la composition de l'urne. Dans ces conditions, on peut alors énoncer et démontrer le théorème qui affirme la convergence probable des fréquences. L'énoncé que propose Bernoulli est tout à fait remarquable, en ce sens qu'il invente l'idée d'intervalle de confiance, ne se contentant pas d'une vague formulation sur la convergence des fréquences.

"Mais pour que cela ne soit pas compris autrement qu'il ne convient, il faut bien noter ce qui suit ; je voudrais que le rapport entre les nombres de cas, que nous entreprenons de déterminer expérimentalement, ne fût pas pris de façon nette et sans partage (car ainsi c'est tout le contraire qui arriverait et il deviendrait d'autant moins probable de découvrir le vrai rapport qu'on ferait de plus nombreuses observations), mais je voudrais que le rapport fût admis avec une certaine latitude, c'est-à-dire compris entre une paire de limites, pouvant être prises aussi rapprochées qu'on voudra. [.../...] On montrera que l'on peut arriver à ce que le rapport trouvé grâce à des expériences recommencées de nombreuses fois tombe entre ces limites [.../...] plus probablement, que toute probabilité donnée, qu'en dehors."⁶

⁵ Ce chapitre IV est intitulé *La double manière de rechercher les nombres de cas. Ce qu'il faut penser de celui qui est établi par des expériences. Problème particulier proposé à ce propos. Etc.*

⁶ Jacques Bernoulli, *Ars Conjectandi*, 4ème partie, traduction de Norbert Meusnier, IREM de Rouen, 1987.

Il peut être intéressant de comparer ce texte avec les extraits du traité de Siméon-Denis Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, 1837. Poisson, reprenant les travaux de Laplace, élargit par ailleurs considérablement le champ d'application de la "loi des grands nombres", en considérant une suite d'expériences différentes avec des probabilités de "réussite" variables. Mais le propos de Poisson est ici de faire comprendre l'énoncé de Bernoulli, de le formuler de façon plus algébrique, et de donner les grandes lignes de la démonstration de Bernoulli. Il donne ensuite sa propre démonstration.

"Les choses de toutes natures sont soumises à une loi universelle qu'on peut appeler *la loi des grands nombres*. Elle consiste en ce que, si l'on observe des nombres très considérables d'événements d'une même nature, dépendants de causes constantes et de causes qui varient irrégulièrement, tantôt dans un sens, tantôt dans l'autre, [...] on trouvera entre ces nombres, des rapports à très peu près constants. [...] Et d'après les nombres des faits constatés, et la grandeur des écarts qui resteront encore entre leurs rapports, le calcul fournira des règles certaines pour déterminer la probabilité que la valeur spéciale vers laquelle ces rapports convergent est comprise entre des limites aussi resserrées qu'on voudra."

"Cette loi des grands nombres s'observe dans les événements que nous attribuons à un aveugle hasard, faute d'en connaître les causes, ou parce qu'elles sont trop compliquées. Ainsi, dans les jeux où les circonstances qui déterminent l'arrivée d'une carte ou d'un dé, varient à l'infini et ne peuvent être soumises à aucun calcul, les différents coups se présentent cependant suivant des rapports constants, lorsque la série des épreuves a été longtemps prolongée. De plus, lorsqu'on aura pu calculer d'après les règles d'un jeu, les probabilités respectives des coups qui peuvent arriver, on vérifiera qu'elles sont égales à ces rapports constants, conformément au théorème connu de Jacques Bernoulli. Mais dans la plupart des questions d'éventualité, la détermination *à priori* des chances des divers événements est impossible, et ce sont, au contraire, les résultats observés qui les font connaître : on saurait, par exemple, calculer d'avance la probabilité de la perte d'un vaisseau dans un voyage de long cours [...]."

"Le théorème [...] est dû à Jacques Bernoulli, qui en avait médité la démonstration pendant vingt années. Celle qu'il a donnée se déduit de la formule du binôme au moyen des propositions suivantes.

Soient, à chaque épreuve, p et q les chances données des deux événements contraires E et F ; soient aussi g, h, k , des nombres entiers, tels que l'on ait

$$p = g/k, q = h/k, g+h = k, p+q = 1 ;$$

désignons par m, n, m , d'autres nombres entiers, liés à g, h, k , par les équations

$$m = gk, n = hk, m = m+n = (g+h)k,$$

de manière que les chances p et q soient entre elles comme les nombres m et n , que l'on pourra rendre aussi grands qu'on voudra en augmentant convenablement g, h, k , sans changer leur rapport. Cela posé :

1°. Dans le développement de $(p+q)^m$, le terme le plus grand sera celui qui répond au produit $p^m q^n$, et comme ce terme est la probabilité de l'arrivée de m fois E et de n fois F, il s'ensuit que cet événement composé, c'est-à-dire, l'arrivée des événements en raison directe de leurs chances

respectives, est le plus probable de tous les événements composés qui peuvent avoir lieu dans un nombre quelconque m d'épreuves.

2°. Si ce nombre m est très grand, le rapport du plus grand terme du développement de $(p+q)^m$ à la somme de tous les termes, ou à l'unité, sera une très petite fraction, qui diminuera indéfiniment à mesure que m augmentera encore davantage ; par conséquent, dans une longue série d'épreuves, l'événement composé le plus probable, le sera cependant très peu, et de moins en moins à mesure que les épreuves seront plus longtemps prolongées.

3°. Mais si l'on considère dans le développement de $(p+q)^m$, son plus grand terme, les l termes qui le suivent et les l termes qui le précèdent, et si l'on désigne par l la somme de ces $2l+1$ termes consécutifs, on pourra toujours, sans changer ni p ni q , prendre m assez grand pour que la fraction l diffère de l'unité, d'aussi peu qu'on voudra ; et à mesure que m augmentera encore davantage, l approchera de plus en plus d'être égal à un. On conclut de là que dans une longue série d'épreuves, il y a toujours une grande probabilité l que l'événement E arrivera un nombre de fois compris entre les limites $m \pm l$, et F un nombre de fois compris entre $n \pm l$; de sorte que sans changer l'intervalle $2l$ des limites de ces deux nombres, on pourra rendre le nombre m des épreuves assez grand pour que la probabilité l soit aussi approchante qu'on voudra de la certitude. Si l'on prend les rapports de ces limites au nombre m des épreuves, que l'on ait égard aux équations précédentes, et qu'on fasse

$$l/m = d, p \pm d = p', q \pm d = q',$$

ces rapports seront p' et q' ; et comme la fraction d diminuera indéfiniment à mesure que m augmentera, il s'ensuit que ces rapports, variables avec m , approcheront aussi indéfiniment, et avec une très grande probabilité, des chances p et q de E et F ; ce qui est l'énoncé du beau théorème de Jacques Bernoulli."⁷

*
* *

II. DE MOIVRE.

Abraham De Moivre est né en 1667 à Vitry-le-François, d'une famille protestante. Exilé en Angleterre après la révocation de l'Édit de Nantes en 1685, il passera le reste de sa vie à Londres, en donnant des leçons particulières et en publiant divers travaux mathématiques. Élu en 1697 à la Royal Society, c'est un ami estimé de Newton. De Moivre meurt à Londres en 1754. La première édition de *The Doctrine of Chances* date de 1718, elle est dédiée à Newton. La seconde est de 1738, la troisième de 1756 (deux ans après la mort de l'auteur)⁸. Ces deux dernières versions sont considérablement augmentées, entre autres à la suite du débat qui oppose De Moivre à Pierre de Montmort.

⁷ Siméon-Denis Poisson, *Recherches sur la probabilité des jugements...*, 1837, Préambule et Chapitre II, § 51.

⁸ Cette troisième édition a été réimprimée à New-York en 1967.

De Moivre se place dans la lignée de Huygens et de Jacques Bernoulli, mais en fait il suit des méthodes différentes. Son originalité consiste à investir dans le Calcul des Probabilités les idées du Calcul Infinitésimal, de Newton, Leibniz et des frères Bernoulli, qui obtient alors ses premiers succès.

The Doctrine of Chances est consacré en partie aux jeux de hasard et aux problèmes des annuités. Le traité contient des résultats remarquables, par exemple sur la durée de jeu, à l'aide d'approximations fonctionnelles et de séries infinies.

L'extrait étudié ici est une annexe de la troisième édition, situé à la fin du traité, après le problème LXIII qui en est l'avant-dernier. Ce texte, comme l'ensemble du traité, est en anglais et n'a, à notre connaissance, jamais été traduit en français⁹. Nous utilisons notre propre traduction¹⁰ qu'on trouvera en annexe.

Le texte débute par une courte introduction où De Moivre situe son travail par rapport à ses prédécesseurs Jacques et Nicolas Bernoulli. Il s'agit d'aller plus loin que les encadrements obtenus par Bernoulli dans l'étude de la somme de certains termes du binôme $(a+b)^n$. De Moivre souhaite évaluer correctement ce type de somme.

La première partie du texte concerne le cas particulier du binôme $(1+1)^n$ (Propositions I et II, Corollaires 1 et 2, Lemme 1, Corollaires 3, 4, 5 et 6). Dans une deuxième partie, De Moivre étendra ses résultats au cas général $(a+b)^n$ (Lemmes 2 et 3, Corollaires 8, 9 et 10). Dans une conclusion (Remarques I et II) De Moivre analyse la portée des résultats obtenus. Remarquons qu'il n'y a pas de Corollaire 7.

De Moivre rappelle d'abord des propriétés publiées par lui-même douze ans auparavant, et qu'il donne donc sans démonstration¹¹.

Le premier résultat consiste en l'évaluation du rapport du terme central du développement de $(1+1)^n$ à la somme totale 2^n , pour une très grande valeur de n . En termes modernes, De Moivre donne un équivalent au voisinage de l'infini de $\frac{C_n^{\frac{n}{2}}}{2^n}$

⁹ Montucla signale que Lagrange a eu le projet de faire une telle traduction en français.

¹⁰ réalisée avec l'aide de Marie-Jeanne Lorang.

¹¹ Ces résultats ont été publiés par De Moivre dans les *Miscellanea Analytica* de Seriebus et Quadratis en 1730.

sous la forme $\frac{2}{B\sqrt{n}}$. B étant une constante, dont De Moivre donne le développement en série du logarithme, à savoir :

$$-1 + \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680}.$$

De Moivre ajoute que son ami James Stirling lui a communiqué depuis que cette constante B était égale à la racine carrée de la circonférence d'un cercle de rayon unité, soit $B = \sqrt{2\pi}$. De Moivre, comme plus tard Laplace, manifeste son étonnement de voir figurer la circonférence d'un cercle dans un contexte qui lui semblait étranger.

Ce premier résultat est lié aux travaux de De Moivre et Stirling à propos de ce qu'on appelle désormais la "formule de Stirling" qui donne un équivalent de $n!$ au voisinage de l'infini¹².

D'un point de vue moderne, on peut retrouver facilement le résultat de De Moivre en admettant la formule de Stirling. Ainsi, comme $C_n^{\frac{n}{2}}$ est égal à $\frac{n!}{\left(\frac{n}{2}!\right)^2}$, il est

équivalent à $\frac{\sqrt{2\pi n} \cdot e^{-n} n^n}{\left(\frac{2\pi n}{2} \cdot e^{-n} \left(\frac{n}{2}\right)^n\right)} = \frac{2^n \cdot 2}{\sqrt{2\pi n}}$. Ce qui donne le résultat obtenu par De Moivre

pour $C_n^{\frac{n}{2}}$.

La deuxième proposition donne un équivalent du logarithme du rapport du terme central au terme qui lui en est distant d'un intervalle l . A savoir, en posant comme De Moivre $m = \frac{n}{2}$ et comme $C_n^{\frac{n}{2}+l} = C_n^{\frac{n}{2}-l}$, que le logarithme de $\frac{C_n^m}{C_n^{m+l}}$ est équivalent à :

$$(m + l - \frac{1}{2})\text{Log}(m + l - 1) + (m - l + \frac{1}{2})\text{Log}(m - l + 1) - 2m\text{Log}m + \text{Log} \frac{m + l}{[m]}.$$

Là aussi ce résultat peut être retrouvé facilement à l'aide de la formule de Stirling, en supposant, ce que De Moivre n'annonce pas pour le moment, que l est constant ou bien qu'il tend vers l'infini mais en étant négligeable devant n . En effet,

¹² Nous avons l'intention de présenter lors d'un prochain colloque Inter-Irem une étude sur les démonstrations de la formule dite de Stirling. Les outils utilisés par De Moivre sont semblables à ceux qu'il utilisera dans la démonstration du Corollaire 2 qui suit : développement en série de $\ln(1+x)$, polynômes et nombres de Bernoulli, obtention d'un équivalent de $\ln(n!)$, tout ceci étant effectué avec des séries dont certaines sont divergentes. La valeur de la constante trouvée par Stirling peut se retrouver en utilisant les intégrales de Wallis. On peut consulter à ce sujet l'article de J.Dutka, *The Early History of the Factorial Function*, in *Archive for History of Exact Sciences*, vol. 43, n° 3, 1991.

$\frac{C_n^m}{C_n^{m+l}} = \frac{(m+l)!(m-l)!}{(m!)^2}$, dont un équivalent est $\frac{(m-l)^{m-l+\frac{1}{2}}(m+l)^{m+l+\frac{1}{2}}}{m^{2m+1}}$. Il ne reste plus qu'à prendre les logarithmes (sans se préoccuper de la légitimité de cette opération!) pour trouver une expression équivalente à celle de De Moivre.

On arrive enfin au Corollaire 1, donné sans démonstration, où l'on peut voir maintenant la première apparition de la loi normale :

"COROLLAIRE 1.

Ceci étant admis, je conclus, que si m ou $1/2 n$ est une Quantité infiniment grande, alors le Logarithme du rapport du Terme distant du terme central d'un intervalle l au Terme central est $-2ll/n$."

Une démonstration possible utilise le développement limité à l'ordre 2 de $\ln(1+x)$ au voisinage de 0, soit :

$$\ln(1+x) = x - \frac{x^2}{2} + o(x^2).$$

En modifiant légèrement l'écriture du résultat précédent de De Moivre, on a :

$$\ln\left(\frac{C_n^{m+l}}{C_n^m}\right) \text{ équivalent à } -(m-l+\frac{1}{2})\ln(1-\frac{l}{m}) - (m+l+\frac{1}{2})\ln(1+\frac{l}{m}).$$

En supposant que l/m , c'est-à-dire l/n , tend vers 0 quand n tend vers l'infini, ce qui revient à la condition implicite pour le moment que l soit négligeable devant n , on peut utiliser le développement limité rappelé ci-dessus, qui donne, après calculs, $-\frac{2l^2}{n} + o(\frac{1}{n})$.

Le Corollaire 2 évalue ensuite la somme des termes compris entre le terme central et celui qui lui est distant de l , d'un côté ou de l'autre. En fait il s'agit du quotient de cette somme par la somme totale des termes 2^n . De Moivre utilise ici explicitement le développement en série de l'exponentielle :

$$e^x = 1 + x + x^2/2! + \dots + x^n/n! + \dots$$

Le résultat du Corollaire 1 donne un équivalent, par exponentiation, du terme $\frac{C_n^{m+l}}{2^n}$ comme $\frac{2}{\sqrt{2\pi n}} \cdot \exp(-\frac{2l^2}{n})$, que l'on peut développer en série comme précédemment.

Pour faire la somme S des $\frac{C_n^{m+p}}{2^n}$ pour p allant de 0 à l , on remplace chaque terme à l'aide de l'expression précédente, et en échangeant les deux sommations, on a :

S équivalent à $\frac{2}{\sqrt{2\pi n}}$ multiplié par la somme de la série de terme général $\frac{(-2)^k}{k!n^k} \cdot u_k$, où u_k est la somme des puissances $2k$ -ièmes des entiers de 0 à l . On sait en utilisant les polynômes de Bernoulli que, si l tend vers l'infini, cette somme u_k est équivalente à $\frac{l^{2k+1}}{2k+1}$. Ce qui donne, si l'on néglige les problèmes de convergence, un équivalent de S sous la forme de $\frac{2}{\sqrt{2\pi n}}$ multiplié par la somme de la série de terme général $\frac{(-2)^k l^{2k+1}}{k!n^k(2k+1)}$.

Ce que De Moivre exprime sous la forme :

"la Somme des Termes compris entre le terme central et celui dont la distance à celui-ci est notée l sera $\frac{2}{\sqrt{nc}}$ multiplié par $1 - \frac{2l^3}{1 \times 3n} + \frac{4l^5}{2 \times 5nn} - \frac{8l^7}{6 \times 7n^3} + \frac{16l^9}{24 \times 9n^4} - \frac{32l^{11}}{120 \times 11n^5} \&c$ "

(où De Moivre note c la circonférence du cercle de rayon unité).

De Moivre remplace ensuite l par $s\sqrt{n}$, ce qui est légitime au vu des conditions implicites posées précédemment, l tend vers l'infini et l est négligeable devant n , c'est même la situation la plus simple qui vienne à l'esprit pour y satisfaire, on ne peut y voir la prémonition de la valeur de l'écart-type. L'autre intérêt est que la série ne va plus comporter de n . En effet la somme S est alors équivalente à $\frac{2}{\sqrt{2\pi}}$ multiplié par la somme de la série de terme général $\frac{(-2)^k s^{2k+1}}{k!(2k+1)}$.

Pour avoir une série qui converge assez vite, et faire des calculs manuels simples, De Moivre prend $s = 1/2$, ce qui donne pour l'équivalent de S , $\frac{2}{\sqrt{2\pi}}$ multiplié par la somme de la série de terme général $\frac{(-1)^k}{k!(2k+1)2^{k+1}}$.

En prenant les 7 premiers termes de la suite, De Moivre obtient une somme égale à 0,427812. Puis à l'aide d'une table de logarithme, il obtient comme valeur pour la somme 0,341344, qui en est une valeur approchée acceptable à 10^{-6} près. De Moivre utilise à la fois le logarithme hyperbolique (notre logarithme népérien) et des tables de logarithmes analogues à celles du père Caramuel¹³.

¹³ Voir Naux, Histoire des Logarithmes, tome 2, pp. 59-61.

Le premier Lemme qui suit aborde l'interprétation probabiliste du résultat du Corollaire précédent. D'après les principes de la "Doctrine des Chances", si on répète n fois une expérience aléatoire à deux issues équiprobables, la probabilité que le nombre de succès soit compris entre $\frac{n}{2} - l$ et $\frac{n}{2} + l$ est égale à la somme des termes du type $C_n^{\frac{n}{2}+p}$, pour p allant de $-l$ à $+l$, divisée par 2^n .

Dans le cas où n est infini, et $l = \frac{\sqrt{n}}{2}$, on peut appliquer les résultats précédents, et conclure qu'alors la probabilité précédente (que le nombre de succès soit compris entre $\frac{n}{2} - \frac{\sqrt{n}}{2}$ et $\frac{n}{2} + \frac{\sqrt{n}}{2}$) est le double du nombre trouvé au corollaire 2, soit 0,682688.

Dans le texte de De Moivre :

"COROLLAIRE 3.

Et ainsi, s'il était possible de faire un nombre infini d'Expériences, la Probabilité qu'un Événement qui a un nombre égal de Chances de se produire ou non, n'ait lieu ni plus que $\frac{n}{2} - \frac{\sqrt{n}}{2}$ fois, ni moins que $\frac{n}{2} + \frac{\sqrt{n}}{2}$, sera exprimée par le double de la Somme donnant le nombre trouvé au second Corollaire, c'est-à-dire, par 0,682688."

De Moivre ne signale pas qu'en doublant la somme il a compté deux fois le terme central. Il est vrai qu'il a pris n infini et démontré avant que le terme central (toujours divisé par 2^n) tendait vers zéro. Il reste que la suite du texte va consister à utiliser le résultat quand n est grand et non infini. Laplace, comme on le verra plus loin, tiendra compte lui de ce doublement du terme central.

Comme c'était l'usage alors, De Moivre exprime plutôt le rapport de la probabilité précédente à celle de l'événement contraire, qu'il exprime comme proche de 28 à 13.

Les Corollaires 4 et 5 expriment l'usage que veut faire De Moivre de ce résultat, dans le cas où n est grand, sans être infini. Il indique ainsi que le résultat est valable pour $n = 3600$, puis pour $n = 900$, et même $n = 100$. La justification est la suivante :

"Cependant, on ne doit pas croire qu'il faille que le nombre n soit immensément grand ; car en le supposant ne pas aller au-delà de la 900ème Puissance, voire même de la 100ème, la Règle donnée ici sera passablement précise, ce que j'ai confirmé par des Essais."

La fin du Corollaire 5 étudie le problème du choix d'une autre valeur de s (plus petite que $1/2$), pour laquelle la probabilité de l'événement recherché sera $1/2$

(au lieu de 0,682688). En procédant sans doute par essais et erreurs, De Moivre donne comme valeur $s = \frac{\sqrt{2}}{4}$. Il conclut en reprenant l'exemple de $n = 3600$:

"ainsi dans le Cas mentionné plus haut, où était posé $= 3600$, $\frac{1}{4}\sqrt{2n}$ sera à peu près 21,2, qui n'est pas plus de la $1/169$ partie de 3600 : ainsi il y a une Chance à peu près égale, ou plutôt un petit peu plus, que pour 3600 Expériences, en chacune desquelles un Evénement peut aussi bien se produire que ne pas se produire, l'Excès des succès ou des échecs au delà de 1800 fois ne dépasse pas environ 21."

Le Corollaire 6 est intéressant puisqu'il s'intéresse au cas où s n'est pas plus petit que $1/2$ (auquel cas la série converge aussi vite), mais où il est plus grand. Et d'abord le cas où $s = 1$. De Moivre remarque que la série obtenue au Corollaire 2 en faisant $l = \sqrt{n}$, converge nettement moins rapidement et nécessiterait des instruments de calcul qu'il ne possède pas. Il revient donc à l'équivalence initiale obtenue avant le développement en série, soit $\frac{C_n^{\frac{n}{2}+l}}{2^n} = \frac{2}{\sqrt{2\pi n}} \cdot \exp\left(-\frac{2l^2}{n}\right)$. Il faut donc

évaluer la somme des termes $\frac{C_n^{\frac{n}{2}+l}}{2^n}$ pour p allant de 0 à \sqrt{n} . Comme on a déjà fait le calcul, par la série, pour p allant de 0 à $\frac{\sqrt{n}}{2}$, il reste à évaluer la somme des termes de $\frac{\sqrt{n}}{2}$ à \sqrt{n} . Au lieu de calculer une somme finie, De Moivre a l'idée de transformer cette somme en calcul d'aire, en intégrale. On voit ici que, même si la rigueur moderne n'y est pas encore, l'idée fondamentale est celle de l'approximation d'une loi de probabilité discrète (la loi binomiale) par une loi continue (la loi normale). En termes modernes, De Moivre écrit donc que la somme recherchée (notée ici T) peut s'écrire $\frac{2}{\sqrt{2\pi n}}$ multiplié par l'intégrale de $\frac{\sqrt{n}}{2}$ à \sqrt{n} de la fonction $\exp(-2x^2/n)$.

Comme il n'a pas de primitive simple de cette fonction, De Moivre utilise un calcul approché par une méthode qu'il attribue à Newton, Cotes et Stirling. Il s'agit de ce qu'on appelle la "règle des trois-huitièmes de Newton"¹⁴. L'idée en est de remplacer la fonction à intégrer par une fonction polynomiale du troisième degré prenant les mêmes valeurs que la fonction initiale en quatre points équidistants, d'ordonnées respectives A, B, C, D . L'intégrale en est alors $\frac{A+D+3(B+C)}{8}$ multiplié par la longueur de l'intervalle d'intégration. Il ne reste plus qu'à appliquer la formule en prenant pour abscisses $3\frac{\sqrt{n}}{6} = \frac{\sqrt{n}}{2}, 4\frac{\sqrt{n}}{6}, 5\frac{\sqrt{n}}{6}, 6\frac{\sqrt{n}}{6} = \sqrt{n}$. Les ordonnées correspondantes sont les images par la fonction exponentielle des nombres $-\frac{1}{2}, -\frac{8}{9}, -\frac{25}{18}, -2$. Le calcul se termine par la valeur (approchée) de T , qui est ensuite doublée (par symétrie) pour obtenir : 0,27160. On lui ajoute la somme trouvée avant, 0,682688.

"et la Somme 0,95428 montrera ce qu'est, après un nombre d'Essais noté n , la Probabilité que l'Evénement n'ait lieu ni plus que $\frac{1}{2}n + \sqrt{n}$ fois, ni moins que $\frac{1}{2}n - \sqrt{n}$, et donc que la Probabilité du

¹⁴ Voir sur cette méthode Histoire d'algorithmes, J-L Chabert & alii, Ed. Belin, 1994, pp. 397-398.

contraire sera 0,04572 : ce qui montre que le Rapport des Probabilités pour qu'un Evénement n'ait lieu ni plus ni moins que dans les Limites assignées est environ 21 contre 1."

De Moivre utilise ensuite la même méthode dans le cas où $l = \frac{3}{2}\sqrt{n}$, et annonce le résultat : 0,99874, ou bien un rapport de probabilités de 369 contre 1.

La fin de ce Corollaire indique les conditions dans lesquelles on peut utiliser la méthode décrite par De Moivre en insistant sur le fait que l'unité de mesure à prendre pour l'écart avec la moyenne est la racine carrée du nombre d'épreuves.

La deuxième partie du texte de De Moivre est consacrée à la généralisation des résultats précédents au cas où il n'y a pas équiprobabilité, c'est-à-dire à l'étude de sommes de termes du développement du binôme $(a+b)^n$. Les résultats sont donnés sans démonstration.

Le Lemme 2 indique que le plus grand terme du développement du binôme $(a+b)^n$ est celui dans lequel a et b sont élevés à des puissances respectives dans la même proportion que a et b eux-mêmes. Il s'agit donc du terme $C_n^q a^q b^{n-q}$, où l'on a posé, comme dans la suite, $q = \frac{an}{a+b}$.

Le Lemme 3 affirme que si on répète n fois une expérience aléatoire à deux issues dont les probabilités d'apparition sont dans le rapport a/b , la probabilité que le nombre de succès soit compris entre $q-l$ et $q+l$ est égale à la somme des termes $C_n^{p+q} a^{p+q} b^{n-p-q}$, pour p allant de $-l$ à $+l$, divisée par $(a+b)^n$. Cf. Lemme 1.

Le Corollaire 8 donne un équivalent du rapport du plus grand terme à la somme totale, quand n tend vers l'infini, à savoir $\frac{a+b}{\sqrt{2\pi abn}}$. Cf. Proposition I.

Le Corollaire 9 donne un équivalent du rapport d'un terme du développement au plus grand terme. A savoir $\frac{C_n^{q+1} a^{q+1} b^{n-q-1}}{C_n^q a^q b^{n-q}}$ peut s'écrire sous la forme $\exp\left(-\frac{(a+b)^2 l^2}{2abn}\right)$. De Moivre ne donne pas même l'ébauche d'une démonstration (semblable à celle réalisée au Corollaire 1, mais qui n'était pas explicite non plus), mais il indique en revanche les conditions que doit vérifier l :

"pourvu que le Rapport de l à n ne soit pas un Rapport fini, mais qu'on puisse le considérer comme celui d'un nombre donné p à \sqrt{n} ; de telle manière qu'on puisse écrire l comme $p\sqrt{n}$."

Ceci permet à De Moivre de conclure cette deuxième partie de son travail :

"COROLLAIRE 10.

Si les Probabilités de succès et d'échec sont dans un Rapport donné d'inégalité, les Problèmes relatifs à la Somme des Termes du Binôme $(a+b)^n$ seront résolus avec la même facilité que ceux dans lesquels les Probabilités de succès et d'échec sont dans un Rapport d'Egalité."

Le texte de De Moivre se termine par deux remarques qui veulent situer la portée des résultats précédents et leur interprétation. La première remarque porte sur la signification des résultats dans le sens direct, celui du calcul. C'est-à-dire dans le cas d'une probabilité définie à priori. Il s'agit d'abord d'une remarque d'ordre mathématique. De Moivre indique d'une certaine manière les deux passages à la limite possible dans son théorème. Plus n devient grand, pour un écart de $\frac{\sqrt{n}}{2}$, plus on se rapproche de la certitude. Mais il revient aussi sur le fait qu'il a considéré l'écart avec la moyenne comme multiple de \sqrt{n} . Plus ce multiple de \sqrt{n} sera grand, pour n fixé, plus on se rapproche aussi de la certitude. Sachant de plus que cet écart $-p\sqrt{n}$ - tend vers l'infini avec n , mais en étant négligeable par rapport à n :

"la raison de ceci sera facilement perçue par les Mathématiciens, qui savent, que la Racine carrée d'une Puissance est avec cette Puissance dans une proportion d'autant plus petite, que son Exposant est grand."

Dans ces conditions il peut donc affirmer :

"De ce qui a été dit, il suit, que le Hasard perturbe très peu les Evénements qui dans leur Institution naturelle sont amenés à se produire ou non, suivant une certaine Loi déterminée."

Et finalement :

"bien que le Hasard produise des Irrégularités, cependant les Rapports de Probabilités seront infiniment grands, et que avec l'avancement du Temps, ces Irrégularités n'auront aucune proportion avec le retour de l'Ordre qui résulte naturellement du DESSEIN ORIGINEL."

La deuxième remarque porte sur la situation inverse, où les expériences permettent de définir une probabilité a posteriori :

"inversement, si à partir d'Observations innombrables, nous trouvons que le Rapport des Evénements converge vers une quantité déterminée, comme le Rapport de P à Q ; alors nous concluons que ce Rapport exprime la Loi déterminée suivant laquelle l'Evénement se produit."

De Moivre donne même une démonstration de ce fait :

"Car supposons que cette Loi soit exprimée non par le rapport $P : Q$, mais par un autre $R : S$; alors le Rapport des Evénements convergerait vers ce dernier, non vers le premier : ce qui contredit notre *Hypothèse*. Et une Absurdité semblable, ou plus grande, découle, si nous supposons que l'Evénement ne se produise pas suivant une Loi, mais d'une manière à la fois décousue et incertaine ; car alors les Evénements ne convergeraient vers aucun Rapport fixé."

Ceci permet à De Moivre de justifier le "dessein originel" et son utilité pour conclure :

"Et ainsi, si nous ne nous aveuglons pas nous-mêmes avec de la poussière métaphysique, nous serons conduits, d'une manière rapide et évidente, à la reconnaissance du grand CREATEUR et MAITRE de toutes choses ; *Lui-même toute sagesse, toute puissance et bonté.*"

De Moivre prend alors l'exemple de la polémique qui a opposé Arbuthnot à Nicolas Bernoulli sur le nombre de naissances de garçons et de filles. On a trouvé qu'à Londres sur un temps assez long (82 ans), le nombre de naissances de garçons et celui de filles était dans un rapport de 18 à 17. De Moivre dit d'abord que si on suppose fixée initialement la proportion des probabilités d'apparition des garçons et des filles (comme s'il s'agissait de lancer un dé à 35 faces dont 18 blanches et 17 noires), alors il n'y a rien d'étonnant, d'après le théorème établi plus haut et d'après la première remarque, que les observations se rapprochent de ce rapport. Puis De Moivre indique que le fait qu'il y ait convergence probable des fréquences démontre l'existence d'un dé et donc d'un artiste l'ayant fabriqué. A la limite le fait d'avoir trouvé un rapport proche de l'égalité mais différent est encore plus convaincant que si l'on avait un rapport d'égalité qui pourrait laisser penser au "hasard" pur.

De Moivre conclut l'étude de cet exemple en s'en prenant aux "athées", atomistes, gassendistes et autres nominalistes, pour qui le mot "hasard" n'est qu'un son :

"Le *Hasard*, dans des écrits ou discours athées, est un son complètement dénué de sens : il n'implique de détermination à aucun *mode d'Existence* ; ni en fait à l'*Existence* elle-même, pas plus qu'à la *non-existence* ; il ne peut être ni défini ni compris ; aucune Proposition le concernant ne peut être ni affirmée ni réfutée, excepté celle-ci, "C'est simplement un mot"."

De Moivre a une autre conception du hasard :

"Le *Hasard*, tel que nous le comprenons, suppose l'*Existence* des choses, et leurs *Propriétés* généralement connues : par exemple, que si on lance un certain nombre de Dés, chacun d'eux s'établira sur l'une ou l'autre de ses Bases. Après quoi, la *Probabilité* d'une Chance donnée, c'est-à-dire d'une disposition particulière des Dés, devient un sujet d'Investigation aussi approprié que peut l'être une autre quantité ou Rapport."

Le texte se termine sur un plaidoyer en faveur des études probabilistes, sous l'autorité de Jacques Bernoulli.

"J'ajouterai seulement, Qu'on peut appliquer utilement cette méthode de Raisonnement à d'autres Recherches très intéressantes ; si ce n'est pour forcer l'Assentiment des autres par une Démonstration rigoureuse, au moins pour la Satisfaction du Chercheur lui-même : et je conclurai cette Remarque avec un passage tiré de l'*Ars Conjectandi* de Mr. *James Bernoulli*, Part IV, Chap. 4, où cet Auteur judicieux et précis introduit ainsi sa Solution du Problème pour *Trouver les Limites entre lesquelles, par la répétition des Expériences, la Probabilité d'un Evénement peut approcher indéfiniment une Probabilité donnée*, "Hoc igitur est illud Problema &c." Voici, dit-il, le Problème que je dois maintenant proposer au Public, après l'avoir gardé par devers moi pendant vingt ans, et dont la nouveauté d'une part, la grande utilité d'autre part, jointes à une égale difficulté, peuvent ajouter du poids et du prix à tous les autres Points de cette Doctrine. Cependant il y a des Auteurs, en fait d'une Classe très différente de celle de *James Bernoulli*, qui laissent entendre que la *Doctrine des Probabilités* ne pourrait trouver de place dans une Recherche sérieuse ; et que des Etudes de ce type, aussi simples et faciles soient-elles, disqualifient un homme pour raisonner sur tout autre sujet. Le Lecteur choisira."

*
* *

III. LAPLACE.

Pierre-Simon de Laplace (1749-1827) écrit dès 1773, année marquée par son entrée à l'Académie des Sciences, un premier mémoire sur les probabilités. À cette époque, Laplace s'était déjà beaucoup préoccupé du calcul différentiel. Il avait proposé une méthode générale pour intégrer les équations linéaires aux dérivées partielles et une théorie des équations aux différences finies partielles généralisant aux suites à plusieurs indices, la théorie des suites récurrentes développée par De Moivre. Il n'avait donc pas besoin du calcul des probabilités, sujet peu à la mode, pour se faire reconnaître du monde scientifique. Ce sujet lui fournissait des exemples permettant d'illustrer sa théorie toute neuve. Dans les mémoires, le calcul des probabilités intervient principalement par son aspect mathématique et les fondements philosophiques des probabilités sont peu abordés.

C'est dans *l'Essai philosophique sur les probabilités*, qui sert d'introduction à la troisième édition (1820) de la monumentale *Théorie analytique des probabilités* (première édition 1812), que Laplace développe sa profession de foi "déterministe" :

"Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre. Une intelligence qui pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ses données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir comme le passé, serait présent à ses yeux."

Il ne s'agit pas pour Laplace d'affirmer qu'une intelligence (supérieure ou non) pourrait calculer tous les effets des lois de la Nature mais son objectif est de développer la science des probabilités en donnant une consistance analytique à la méthode empirique pour avoir une connaissance plus exacte des lois de la nature. L'*Essai* présente "*sans le secours de l'analyse*" les principes et les résultats généraux de la *Théorie analytique des probabilités*. Laplace conclut par :

"J'ai rassemblé toutes ces méthodes dans ma Théorie analytique des probabilités, où je me suis proposé d'exposer de la manière la plus générale les principes et l'Analyse du Calcul des Probabilités, ainsi que les solutions des problèmes les plus intéressants et les plus difficiles que ce calcul présente. On voit par cet Essai que la théorie des probabilités n'est au fond que le bon sens réduit au calcul : elle fait apprécier avec exactitude, ce que les esprits justes sentent par une sorte d'instinct, sans qu'ils puissent souvent s'en rendre compte."

La *Théorie analytique des probabilités* est constituée de deux parties : le Livre Premier où sont développées la théorie des fonctions génératrices et la théorie des approximations des formules qui sont fonctions des grands nombres et le Livre Deuxième qui traite de la théorie générale des probabilités. Dans le Livre Premier sont traitées la théorie des développements asymptotiques et celle des transformations de Fourier et Laplace, en utilisant l'analogie entre les nombres et les

fonctions et en étendant, par induction, les méthodes du calcul algébrique au calcul différentiel. C'est ce qu'explique Laplace en écrivant :

"Mais une remarque très importante dans cette analyse, et qui permet de l'étendre aux fonctions que la théorie des probabilités présente le plus souvent, est que les séries auxquelles on parvient, ont également lieu dans le cas même où, par des changements de signe dans les coefficients de l'équation des limites, ses racines deviennent imaginaires. Ces passages du positif au négatif, du réel à l'imaginaire, [.../...], m'ont conduit dans ces Mémoires aux valeurs de plusieurs intégrales définies, qui offrent cela de remarquable, savoir, qu'elles dépendent de deux transcendentes, le rapport de la circonférence au diamètre, et le nombre dont le logarithme hyperbolique est l'unité. On peut donc considérer ces passages comme moyens de découvertes pareils à l'induction dont les géomètres font depuis longtemps usage. Mais ces moyens, quoique employés avec beaucoup de précautions et de réserve, laissent toujours à désirer des démonstrations de leurs résultats."

Les résultats sont énoncés de manière très générale et leur compréhension est difficile. Les preuves sont plutôt esquissées que détaillées. Pour mieux comprendre la *Théorie analytique des probabilités*, il est utile de lire les *Recherches sur la probabilité des jugements* de S.D. Poisson. Celui-ci a étudié avec soin le texte de Laplace et il présente les théories laplaciennes avec un souci pédagogique.

L'extrait étudié est le début du chapitre III du Livre Deuxième (pp.301-309 de l'édition des Œuvres Complètes de 1847) et consiste en une reprise, à la suite de De Moivre, du théorème de Bernoulli. Dans son *Essai*, Laplace écrit :

"Ce théorème indiqué par le bon sens était difficile à démontrer par l'Analyse. Aussi l'illustre géomètre Jacques Bernoulli, qui s'en est occupé le premier, attachait-il une grande importance à la démonstration qu'il en a donnée. Le calcul des fonctions génératrices appliqué à cet objet, non seulement démontre avec facilité ce théorème, mais de plus il donne la probabilité que le rapport des événements observés ne s'écarte que dans certaines limites du vrai rapport de leurs possibilités respectives."

Et un peu plus loin,

"Moivre a repris dans son ouvrage [The Doctrine of Chances] le théorème de Bernoulli sur la probabilité des résultats déterminés par un grand nombre d'observations. Il ne se contente pas de faire voir, comme Bernoulli, que le rapport des événements qui doivent arriver approche sans cesse de leurs possibilités respectives, il donne de plus une expression élégante et simple de la probabilité que la différence de ces deux rapports est contenues dans des limites données."

Le chapitre III est intitulé "Des lois de la probabilité qui résultent de la multiplication indéfinie des événements.". Dans l'introduction, Laplace indique que lorsque l'on répète un très grand nombre de fois une épreuve où deux événements simples sont possibles dont un seul doit arriver à chaque fois, chaque événement arrive proportionnellement à sa facilité. Il va démontrer analytiquement le théorème et précise qu'une utilisation "inverse" de ce théorème permettra de découvrir par l'expérience la facilité de l'événement.

Il suit une démarche analogue à celle de De Moivre avec des outils mathématiques différents ; ces outils sont développés dans le Livre Premier de la *Théorie Analytique des Probabilités*.

D'abord, il prouve que le plus grand terme du développement du binôme $(p + (1 - p))^{x+x'}$ s'obtient, quand x et x' sont de très grands nombres, lorsque le rapport de x à x' est voisin de celui de p à $(1-p)$. Il en déduit que la combinaison la plus probable, lorsque que l'on répète l'épreuve un très grand nombre de coups, est celle dans laquelle chaque événement est répété proportionnellement à sa probabilité.

Ensuite, il cherche à évaluer le $l^{\text{ième}}$ terme du binôme après le plus grand. Pour ce faire, il utilise la formule de Stirling obtenue dans le paragraphe n° 33 de la seconde partie du Livre Premier (et non pas n° 32 comme indiqué dans le texte). Celle-ci est obtenue à partir de la solution particulière d'une équation aux différences finies, sous forme intégrale qui est ensuite développée en série¹⁵. On peut remarquer la notation c pour désigner le nombre dont le logarithme hyperbolique est l'unité. Puis il utilise des développements en série pour obtenir des expressions de $(x - l)^{1-x-\frac{1}{2}}$ et $(x' + l)^{-x'-l-\frac{1}{2}}$, lorsque l est au plus de l'ordre de \sqrt{n} , ce qui lui permet d'obtenir :

$$\frac{1.2.3\dots n}{1.2.3\dots(x-l).1.2.3\dots(x'+l)} \cdot p^{x-l} \cdot (1-p)^{x'+l} = \frac{\sqrt{n} \cdot e^{-\frac{nl^2}{2xx'}}}{\sqrt{\pi} \cdot \sqrt{2xx'}} \cdot \left(1 + \frac{nl}{xx'} + \frac{l(x'-x)}{2xx'} - \frac{l^3}{6x^2} + \frac{l^3}{6x'^2} \right).$$

Il cherche ensuite à évaluer la somme des $(2l + 1)$ termes du binôme $(p + (1 - p))^{x+x'}$ compris entre les termes équidistants du plus grand terme de la "distance" l , en les groupant 2 par 2.

Laplace doit calculer la somme :

$$\sum_{k=0}^l \frac{2\sqrt{n}}{\sqrt{\pi}\sqrt{2x \cdot x'}} \cdot e^{-\frac{nl^2}{2x \cdot x'}} - \frac{\sqrt{n}}{\sqrt{\pi}\sqrt{2x \cdot x'}}.$$

Pour évaluer cette somme, il utilise une formule sommatoire obtenue au n°10 du Livre Premier, analogue à la formule sommatoire d'Euler-MacLaurin¹⁶. Celle-ci ramène la somme d'une série à celle d'une intégrale. Cette formule s'écrit :

$$\sum \cdot y = \frac{1}{e^{\frac{dy}{dl}} - 1} = \left(\frac{dy}{dl} \right)^{-1} - \frac{1}{2} \cdot \left(\frac{dy}{dl} \right)^0 + \frac{1}{12} \frac{dy}{dl} + \text{etc.}$$

¹⁵ Pour avoir une démonstration détaillée de cette formule, voir S.D. Poisson, *Recherches sur la probabilité des jugements*, ch III, pp 173-175

¹⁶ Voir pour cette formule *Histoire d'algorithmes*, J-L Chabert & alii, Ed. Belin, 1994, pp. 483-488.

où les coefficients des puissances successives de $\frac{dy}{dx}$ sont les nombres de Bernoulli.

Pour essayer de comprendre d'où vient cette formule, il faut écrire la formule de Taylor de manière formelle :

$$y_{x+h} = y_x + h \frac{d}{dx} y + \frac{h^2}{2!} \frac{d^2}{dx^2} y + \frac{h^3}{3!} \frac{d^3}{dx^3} y + \text{etc.}$$

$$y_{x+h} - y_x = h \frac{d}{dx} y + \frac{h^2}{2!} \frac{d^2}{dx^2} y + \frac{h^3}{3!} \frac{d^3}{dx^3} y + \text{etc.}$$

soit

$$\Delta_h y_x = \left(h \frac{d}{dx} + \frac{h^2}{2!} \frac{d^2}{dx^2} + \frac{h^3}{3!} \frac{d^3}{dx^3} + \text{etc.} \right) y_x = \left(e^{h \frac{d}{dx}} - 1 \right) y_x$$

Ensuite il faut utiliser l'analogie : l'opérateur Δ de différence finie étant l'inverse de l'opérateur Σ de somme finie, il suffit d'appliquer à y_x , l'opérateur inverse ! On obtient alors :

$$\Sigma_h y_x = \left(\frac{1}{e^{h \frac{d}{dx}} - 1} \right) y_x \quad \text{et pour } h = 1, \quad \Sigma_h y_x = \left(\frac{1}{e^{\frac{d}{dx}} - 1} \right) y_x.$$

Comme on applique la formule à $y = \frac{2\sqrt{n}}{\sqrt{\pi\sqrt{2x \cdot x'}}} \cdot e^{-\frac{nl^2}{2x \cdot x'}}$, il suffit de ne conserver que les deux premiers termes du développement. Les autres sont négligeables car l est au plus de l'ordre de \sqrt{n} .

$$\text{On a donc : } \sum_{k=0}^{l-1} y(k) = \int_0^1 y dk - \frac{1}{2} y(l) + \frac{1}{2} y(0)$$

$$\text{d'où } \sum_{k=0}^{l-1} y(k) - \frac{1}{2} y(0) = \int_0^1 y dk - \frac{1}{2} y(l) \quad \text{et} \quad \sum_{k=0}^1 y(k) - \frac{1}{2} y(0) = \int_0^1 y dk + \frac{1}{2} y(l)$$

en ajoutant $y(l)$ à chaque membre.

Si l'on fait le changement de variable $t = \frac{l\sqrt{n}}{\sqrt{2x \cdot x'}}$, la somme cherchée s'écrit :

$$\frac{2}{\sqrt{\pi}} \int e^{-t^2} dt + \frac{\sqrt{n}}{\sqrt{\pi \cdot \sqrt{2x \cdot x'}}} \cdot e^{-t^2}.$$

Ceci est la probabilité que $\left(\frac{x}{n} - p\right)$ soit compris entre les valeurs $\frac{z}{n} \pm \frac{t\sqrt{2x \cdot x'}}{n\sqrt{n}}$.

On a donc :

$$P \left[\frac{z}{n} - \frac{t\sqrt{2x \cdot x'}}{n\sqrt{n}} \leq \frac{x}{n} - p \leq \frac{z}{n} + \frac{t\sqrt{2x \cdot x'}}{n\sqrt{n}} \right] = \frac{2}{\sqrt{\pi}} \int_0^{\frac{l\sqrt{n}}{\sqrt{2x \cdot x'}}} e^{-t^2} dt + \frac{\sqrt{n}}{\sqrt{\pi \sqrt{2x \cdot x'}}} e^{-t^2}.$$

Comme z est un nombre compris entre 0 et 1, on voit que l'intervalle a une longueur de l'ordre de $\frac{1}{\sqrt{n}}$ car $\sqrt{2x \cdot x'}$ est de l'ordre de $n \cdot \sqrt{p(1-p)}$.

On peut donc écrire que pour n suffisamment grand :

$$P \left[\frac{z}{n} - \frac{t\sqrt{2p(1-p)}}{\sqrt{n}} \leq \frac{x}{n} - p \leq \frac{z}{n} + \frac{t\sqrt{2p(1-p)}}{\sqrt{n}} \right] = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt + \frac{1}{\sqrt{\pi}\sqrt{2n.p(1-p)}} e^{-t^2}.$$

À partir de cette formule on obtient deux types d'approximations :

1) Pour t fixé égal à T, on peut rendre n suffisamment grand de manière à rendre les bornes de l'intervalle aussi proches l'une de l'autre que l'on veut, tandis que la probabilité reste supérieure à $\frac{2}{\sqrt{\pi}} \int_0^T e^{-t^2} dt$.

2) L'intervalle étant fixé, lorsque n croît, t croît aussi, pratiquement comme \sqrt{n} , et la probabilité tend très rapidement vers l'unité comme le montre l'observation d'une table de loi normale.

Laplace prend, comme De Moivre, l'exemple de la polémique ayant opposé le Dr Arbuthnot à Nicolas Bernoulli sur le nombre de naissances de filles et de garçons à Londres. Il s'en sert pour calculer la probabilité que sur 14000 naissances, le nombre de garçons soit dans l'intervalle [7037 ; 7363], en supposant qu'il naît en moyenne 18 garçons pour 17 filles. La présentation par Laplace, athée légendaire, de cet exemple numérique est là uniquement pour montrer l'efficacité de la méthode analytique contrairement à De Moivre qui l'utilisait pour justifier l'existence de Dieu.

Enfin, il termine par une utilisation inverse de la formule. Il s'agit du cas où l'on ne connaît pas a priori la probabilité p. Si l'on connaît le nombre de fois i, où l'événement a s'est réalisé au cours de n épreuves et si l'on néglige certains termes, on obtient :

$$P \left[\frac{i}{n} - \frac{T\sqrt{2i(n-i)}}{n\sqrt{n}} \leq p \leq \frac{i}{n} + \frac{T\sqrt{2i(n-i)}}{n\sqrt{n}} \right] = \frac{2}{\sqrt{\pi}} \int_0^T e^{-t^2} dt + \frac{\sqrt{n}}{\sqrt{\pi}\sqrt{2i(n-i)}} e^{-T^2}.$$

ce qui exprime la probabilité que la probabilité p de l'événement a soit contenue dans certaines limites. Lorsque n augmente, ces limites se rapprochent et la probabilité que p soit compris entre ces limites tend vers la certitude, ce qui permet en multipliant le nombre d'épreuves de connaître la probabilité p. Ce point de vue sera développé par Laplace dans le chapitre VI, intitulé "De la probabilité des causes et des événements futurs tirée des événements observés".

*
* *

CONCLUSION.

Les textes de Bernoulli, De Moivre, Laplace et Poisson que nous avons présentés pendant cet atelier apportent plusieurs enseignements.

Du point de vue de l'histoire interne des mathématiques, on voit le travail considérable, après l'invention du théorème par Bernoulli, de recherche de précision dans l'énoncé même du théorème et de simplification dans sa démonstration. Simplification qui va chez Laplace jusqu'à l'établissement de propriétés beaucoup plus générales et compliquées. Si les démonstrations de De Moivre et Laplace nous paraissent parfois manquer de la rigueur de la mathématique moderne, elles sont caractéristiques du style des analystes au 18^{ème} siècle. Il reste que les idées forces sont en place : énoncé du théorème avec un intervalle de confiance ; approximation d'une loi discrète par une loi continue, prééminence de la loi normale, développement formel taylorien. Si le rôle de la moyenne d'une variable aléatoire est encore confondu avec son mode, et si l'importance de l'écart type n'est pas encore reconnu, il faut remarquer comment, sur ce point, on voit s'appliquer une branche des mathématiques - le calcul infinitésimal - sur une autre - le calcul des probabilités -.

On a vu aussi comment l'interprétation de la "loi des grands nombres" dans le sens d'une définition a posteriori de la probabilité était accompagnée de considérations métaphysiques, de la volonté apologétique de De Moivre au déterminisme laplacien. Dans tous les cas, cette interprétation était liée à l'étude d'exemples qui sortaient du cadre habituel des jeux de hasard, où la définition a priori ne gêne personne. On pourrait en tirer une conséquence à portée pédagogique sur l'importance de la définition d'une probabilité à partir de l'observation de la convergence des fréquences. Si l'on est dans le cadre d'une probabilité définie a priori, la "loi des grands nombres" est un théorème interne à la théorie, un énoncé scientifique susceptible de démonstration. Sinon, il s'agit d'une affirmation extra mathématique, d'une prise de position d'ordre métaphysique. La présenter comme allant de soi, même comme conséquence de multiples observations, n'est peut-être pas la meilleure façon d'initier les élèves à ce qu'est une démarche scientifique.

Pour conclure, on pourra relever la singularité du calcul des probabilités, muni de la "loi des grands nombres", dont l'objectif paraît être de faire disparaître son objet : le hasard. C'est peut-être ce que voulait dire Andreï Kolmogorov :

"La valeur épistémologique de la théorie des probabilités est fondée sur le fait que les phénomènes aléatoires engendrent à grande échelle une régularité stricte, où l'aléatoire a, d'une certaine façon, disparu."¹⁷

¹⁷ cité par Jean-Philippe Bouchaud, *La loi des grands nombres*, in *La Recherche*, n°278, 1995, p. 784.

Bibliographie :

- J. Bernoulli 1703 : *Ars Conjectandi*.
- A. De Moivre 1756 : *The doctrine of Chances*, 3^{ème} édition, Londres
- P. S. Laplace 1812 : *Théorie analytique des probabilités*, Paris.
- P. S. Laplace 1825 : *Essai philosophique des probabilités*. Editions Ch. Bourgois, Paris 1986.
- S.-D. Poisson 1837 : *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Bachelier, Paris.
- I. Todhunter 1865 : *A history of the mathematical theory of probability*. Chelsea, New York 1949.
- B. Vauquois 1969 : *Probabilités*. Hermann, Paris. Pour une version moderne des théorèmes de Bernoulli et de De Moivre - Laplace.
- N. Meusnier 1987 : Traduction de la 4^{ème} partie de l'*Ars Conjectandi*, IREM de Rouen.

ANNEXE I

Abraham De Moivre
(1667-1754)

THE DOCTRINE OF CHANCES
(3ème édition, 1756)

pages 243 à 254

Méthode pour approcher la Somme des termes du Binôme $(a+b)^n$ développé en Série, d'où sont déduites des Règles pratiques pour estimer le Degré d'Assentiment qu'on doit donner aux Expériences.

Bien que la Solution des Problèmes de Hasard demande souvent que plusieurs Termes du binôme $(a+b)^n$ soient additionnés ensemble, néanmoins pour les très grandes Puissances la chose apparaît si laborieuse et de si grande difficulté, que peu de gens ont entrepris cette Tâche ; car en dehors de *Jacques* et *Nicolas Bernoulli*, deux grands Mathématiciens, je ne connais personne qui l'ait entreprise ; en ceci, bien qu'ils aient montré une très grande adresse, et qu'ils aient la louange due à leur Ingéniosité, pourtant il y avait encore des choses à faire ; car ce qu'ils ont fait n'est pas plus qu'une Approximation comme de déterminer les bornes très éloignées, entre lesquelles ils démontrèrent que la Somme des Termes était contenue. Or la Méthode qu'ils ont suivie a été brièvement décrite dans mes *Miscellanea Analytica*, que le Lecteur peut consulter s'il lui plaît, à moins qu'il ne choisisse plutôt, ce qui serait peut-être le mieux, de consulter ce qu'ils ont eux-mêmes écrit sur ce sujet : pour ma part, ce qui m'a fait m'appliquer à cette Recherche n'était pas du tout que je pourrais surpasser d'autres personnes, en quoi toutefois j'aurais pu être pardonné ; mais ce que j'ai fait était en réponse au désir d'un Gentilhomme de grande valeur, et bon Mathématicien, qui m'y encouragea ; j'ajoute maintenant de nouvelles réflexions aux précédentes ; mais pour montrer plus clairement leur liaison, il m'est nécessaire de résumer un petit nombre de choses que j'ai communiquées il y a un certain temps.

I. Il y a maintenant une douzaine d'années ou plus que j'ai trouvé ce qui suit : Si le Binôme $1+1$ est élevé à une très grande Puissance notée n , le rapport du Terme central à la Somme de tous les Termes, c'est-à-dire, à 2^n , peut être exprimé par la Fraction $\frac{2A(n-1)^n}{n^n \sqrt{n-1}}$ où A représente le nombre dont le Logarithme Hyperbolique est $\frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \dots$

Mais comme la Quantité $\frac{(n-1)^n}{n^n}$ ou $(1-\frac{1}{n})^n$ est très près d'être donnée quand n est une grande puissance, ce qui n'est pas difficile à prouver, il suit que, pour une Puissance infinie, cette quantité sera absolument donnée, et représente le nombre dont le Logarithme Hyperbolique est -1 ; d'où il suit, que si B désigne le Nombre dont le Logarithme Hyperbolique est $-1 + \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \dots$ l'Expression écrite ci-dessus deviendra $\frac{2B}{\sqrt[n-1]{n}}$ ou à peu près $\frac{2B}{\sqrt[n]{n}}$: et que donc si nous changeons les Signes de cette Série, et supposons maintenant que B désigne le Nombre dont le Logarithme Hyperbolique est $1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \frac{1}{1680} \dots$, cette Expression sera changée en $\frac{2}{B\sqrt[n]{n}}$.

Quand j'abordai pour la première fois cette recherche, je me contentai de déterminer en gros la valeur de B , en additionnant quelques termes de la Série ci-dessus ; mais comme je m'aperçus qu'elle convergeait très lentement, et en voyant au même moment que ce que j'avais fait répondait assez bien à mon propos, je n'essayai pas d'aller plus loin jusqu'à ce que mon digne et savant Ami M. *James Stirling*, qui s'était lancé après moi dans cette recherche, ait trouvé que la Quantité B représentait en fait la Racine Carrée de la Circonférence d'un Cercle dont le Rayon est l'Unité, si bien que si cette circonférence est appelée c , le rapport du Terme central à la somme de tous les termes sera exprimée par $\frac{2}{\sqrt{nc}}$.

Mais bien qu'il ne soit pas nécessaire de savoir quelle relation le nombre B peut avoir avec la Circonférence du Cercle, pourvu que sa valeur soit obtenue, soit en continuant la Série Logarithmique mentionnée avant, soit par une autre voie ; pourtant je reconnais avec plaisir que cette découverte, outre sa concision, a donné une Elégance particulière à la Solution.

II. J'ai aussi découvert que le Logarithme du Rapport du Terme central pour une grande Puissance à un Terme quelconque qui lui est distant d'un intervalle noté l , pouvait être représenté avec une très bonne approximation, (en posant $m = 1/2 n$) par les Quantités

$$(m + l - \frac{1}{2})\text{Log}(m + l - 1) + (m - l + \frac{1}{2})\text{Log}(m - l + 1) - 2m\text{Log}m + \text{Log} \frac{m + l}{[m]} .$$

COROLLAIRE 1.

Ceci étant admis, je conclus, que si m ou $1/2 n$ est une Quantité infiniment grande, alors le Logarithme du rapport du Terme distant du terme central d'un intervalle l au Terme central est $-\frac{2ll}{n}$.

COROLLAIRE 2.

Le Nombre, qui donne $-\frac{2ll}{n}$ par le Logarithme Hyperbolique, étant $1 - \frac{2ll}{n} + \frac{4l^4}{2nn} - \frac{8l^6}{6n^3} + \frac{16l^8}{24n^4} - \frac{32l^{10}}{120n^5} + \frac{64l^{12}}{720n^6} \dots$, il suit, que la Somme des Termes compris entre le terme

central et celui dont la distance à celui-ci est notée l sera $\frac{2}{\sqrt{nc}}$ multiplié par $1 - \frac{2l^3}{1 \times 3n} + \frac{4l^5}{2 \times 5n^2} - \frac{8l^7}{6 \times 7n^3} + \frac{16l^9}{24 \times 9n^4} - \frac{32l^{11}}{120 \times 11n^5} \dots$

Supposons maintenant $l = s\sqrt{n}$, alors la dite Somme sera exprimée par la Série $\frac{2}{\sqrt{c}}$ multiplié par $s - \frac{2s^3}{3} + \frac{4s^5}{2 \times 5} - \frac{8s^7}{6 \times 7} + \frac{16s^9}{24 \times 9} - \frac{32s^{11}}{120 \times 11} \dots$

De plus, si $s = 1/2$, alors la Série deviendra $\frac{2}{\sqrt{c}}$ multiplié par $\frac{1}{2} - \frac{1}{3 \times 4} + \frac{1}{2 \times 5 \times 8} - \frac{1}{6 \times 7 \times 1[6]} + \frac{1}{24 \times 9 \times 32} - \frac{1}{120 \times 11 \times 64} \dots$ qui converge si vite, qu'avec l'aide de pas plus de 7 ou 8 Termes, la Somme cherchée sera obtenue avec 6 ou 7 Décimales : A savoir qu'on trouvera cette Somme égale à 0,427812, indépendamment du multiplicateur commun $\frac{2}{\sqrt{c}}$, et donc le logarithme tabulaire de 0,427812, qui est d'après la Table $\bar{9}.6312529$, étant ajouté au Logarithme de $\frac{2}{\sqrt{c}}$, c'est-à-dire $\bar{9}.9019400$, la Somme sera $\bar{19}.5331929$, à quoi correspond le nombre 0,341344.

LEMME [1].

Si un Evénement dépend du Hasard, de manière que les Probabilités qu'il se produise ou non soient égales, et qu'on fasse un certain nombre donné n d'Expériences pour observer combien de fois il a lieu ou non, et qu'on prenne aussi un autre nombre donné l , plus petit que $1/2 n$, alors la Probabilité qu'il ne se produise ni plus que $1/2 n + l$ fois, ni moins que $1/2 n - l$ fois, peut être trouvée comme suit.

Soit L et L les deux Termes également distants des deux côtés du Terme central du développement du binôme $(1+1)^n$, d'un intervalle égal à l ; soit aussi S la Somme des Termes compris entre L et L y compris les extrémités, alors la Probabilité cherchée sera exprimée de manière exacte par la fraction $\frac{S}{2^n}$; ce qui étant fondé sur les Principes communs de la Doctrine des Chances, ne nécessite pas de Démonstration ici.

COROLLAIRE 3.

Et ainsi, s'il était possible de faire un nombre infini d'Expériences, la Probabilité qu'un Evénement qui a un nombre égal de Chances de se produire ou non, n'ait lieu ni plus que $1/2 n + 1/2 \sqrt{n}$ fois, ni moins que $1/2 n - 1/2 \sqrt{n}$, sera exprimée par le double de la Somme donnant le nombre trouvé au second Corollaire, c'est-à-dire, par 0,682688, et en conséquence la Probabilité du contraire, qui est celle du succès moins fréquent ou plus rare que dans la proportion mentionnée plus haut, sera 0,317312, ces deux Probabilités rassemblées donnant l'Unité, qui est la mesure de la Certitude : Donc le rapport de ces Probabilités est pour simplifier très proche de 28 contre 13.

COROLLAIRE 4.

Mais bien qu'il ne soit pas possible de réaliser un nombre infini d'Expériences, pourtant les Conclusions précédentes peuvent très bien s'appliquer aux nombres finis, pourvu qu'ils soient grands : par Exemple, si on fait 3600 Expériences, on pose $n = 3600$, d'où $1/2 n$ sera 1800, et $1/2 \sqrt{n} = 30$, donc la Probabilité que l'événement ne se produise ni plus que 1830 fois, ni moins que 1770, sera 0,682688.

COROLLAIRE 5.

Et ainsi nous pouvons énoncer ceci comme une Maxime fondamentale, que pour les grandes Puissances, le Rapport, de la Somme des Termes compris entre deux Extrémités distantes des deux côtés du Terme central d'un Intervalle égal à $1/2 \sqrt{n}$, à la Somme de tous les Termes, sera correctement exprimé par le nombre Décimal 0,682688, ce qui est à peu près 28/41.

Cependant, on ne doit pas croire qu'il faille que le nombre n soit immensément grand ; car en le supposant ne pas aller au-delà de la 900ème Puissance, voire même de la 100ème, la Règle donnée ici sera passablement précise, ce que j'ai confirmé par des Essais.

Mais il vaut la peine d'observer, que comme $1/2 \sqrt{n}$ est petit par rapport à n , et d'autant plus petit par rapport à n quand n augmente, on obtient très vite la Probabilité 28/41 ou le Rapport de Probabilités de 28 contre 13 ; d'où nous pouvons être naturellement amené à chercher, quelles sont les Bornes entre lesquelles on trouvera la Proportion d'Egalité ? Je réponds, que ces Bornes doivent être placées à une distance du Terme central, exprimée à peu près par $\frac{1}{4} \sqrt{2n}$; ainsi dans le Cas mentionné plus haut, où était posé $n = 3600$, $\frac{1}{4} \sqrt{2n}$ sera à peu près 21,2, qui n'est pas plus de la 1/169 partie de 3600 : ainsi il y a une Chance à peu près égale, ou plutôt un petit peu plus, que pour 3600 Expériences, en chacune desquelles un Evénement peut aussi bien se produire que ne pas se produire, l'Excès des succès ou des échecs au delà de 1800 fois ne dépasse pas environ 21.

COROLLAIRE 6.

Si l est remplacé par \sqrt{n} , la Série ne convergera pas aussi rapidement qu'elle ne le faisait dans le Cas précédent où l était remplacé par $1/2 \sqrt{n}$, car ici il ne faudra pas moins de 12 ou 13 Termes de la Série pour avoir une approximation passable, et il faudrait encore plus de Termes, dans la mesure où l est dans une proportion plus grande avec \sqrt{n} : pour cette raison je fais usage dans ce Cas de l'Artifice des Quadratures Mécaniques, inventé pour la première fois par Sir *Isaac Newton*, et poursuivi depuis par Mr. *Cotes*, Mr. *James Stirling*, moi-même, et peut-être d'autres ; il consiste à déterminer l'Aire approchée d'une Courbe, connaissant un certain nombre de ses Ordonnées A, B, C, D, E, F, &c. placées à des Intervalles égaux, plus il y a d'Ordonnées, plus la Quadrature sera exacte ; mais ici je me borne à quatre, ceci étant suffisant pour mon propos : supposons ainsi que les quatre Ordonnées sont A, B, C, D, et que la Distance entre la première et la dernière soit notée l , alors l'Aire

contenue entre la première et la dernière sera $\frac{1(A+D)+3(B+C)}{8} \times 1$; maintenant prenons les Distances $0 \sqrt{n}$, $1/6 \sqrt{n}$, $2 \sqrt{n} /6$, $3 \sqrt{n} /6$, $4 \sqrt{n} /6$, $5 \sqrt{n} /6$, $6 \sqrt{n} /6$, chacune dépassant la précédente de $1/6 \sqrt{n}$, et dont la dernière est \sqrt{n} ; alors en prenant les Carrés des quatre dernières, en doublant chacune d'elles, en les divisant toutes par n , et en les faisant toutes précéder du Signe -, nous aurons $-1/2$, $-8/9$, $-25/18$, $-2/1$, qui doivent être considérés comme des Logarithmes Hyperboliques, dont en conséquence les nombres correspondants, à savoir 0,60653, 0,41111, 0,24935, 0,13534 seront les valeurs des quatre Ordonnées A, B, C, D. Or en ayant remplacé l par $1/2 \sqrt{n}$, on trouvera l'Aire = 0,170203. \sqrt{n} dont le double étant multiplié par $\frac{2}{\sqrt{nc}}$, le produit sera 0,27160 ; ajoutons ceci à l'Aire trouvée avant, c'est-à-dire, à 0,682688, et la Somme 0,95428 montrera ce qu'est, après un nombre d'Essais noté n , la Probabilité que l'Événement n'ait lieu ni plus que $1/2 n + \sqrt{n}$ fois, ni moins que $1/2 n - \sqrt{n}$, et donc que la Probabilité du contraire sera 0,04572 : ce qui montre que le Rapport des Probabilités pour qu'un Événement n'ait lieu ni plus ni moins que dans les Limites assignées est environ 21 contre 1.

Et en raisonnant de la même manière, on trouvera que la Probabilité que l'Événement ne se produise ni plus que $1/2 n + 3/2 \sqrt{n}$ fois, ni moins que $1/2 n + 3/2 \sqrt{n}$ sera 0,99874, ce qui donnera dans ce cas le Rapport des Probabilités d'environ 369 contre 1.

Pour appliquer ceci à des Exemples particuliers, il sera nécessaire d'estimer la fréquence de la réalisation ou de l'échec d'un Événement par la Racine Carrée du nombre qui désigne combien d'Expériences ont été faites, ou sont à faire ; et cette Racine Carrée, suivant ce qui a été indiqué au quatrième Corollaire, sera un *Modulus* [Norme] par lequel nous pouvons apprécier notre Estimation ; et ainsi supposons que le nombre d'Expériences à faire est 3600, et qu'on veut connaître la Probabilité que l'Événement ne se produise ni plus que 1850 fois, ni moins que 1750, ces deux nombres pouvant être changés à volonté, à condition qu'ils soient également distant du centre de la Somme 1800, alors faisons la demi-différence des deux nombres 1850 et 1750, c'est-à-dire, dans ce Cas, $50 = s \sqrt{n}$; or comme nous avons supposé $3600 = n$, alors \sqrt{n} sera = 60, qui fera que 50 sera = 60s, et en conséquence $s = 50/60 = 5/6$; et ainsi si nous prenons la proportion, dans le cas d'une puissance infinie, entre la double Somme des Termes correspondant à l'Intervalle $5/6 \sqrt{n}$, et la Somme de tous les Termes, nous aurons une approximation par excès de la Probabilité cherchée.

LEMME 2.

Dans le développement de $(a+b)^n$ à une Puissance quelconque, le plus grand Terme est celui dans lequel les Exposants des Puissances de a et b sont dans la même proportion que les Quantités a et b elles-mêmes ; ainsi si on prend la 10ème Puissance de $a+b$, qui est

$$a^{10} + 10a^9b + 45a^8b^2 + 120a^7b^3 + 210a^6b^4 + 252a^5b^5 + 210a^4b^6 + 120a^3b^7 + 45a^2b^8 + 10ab^9 + b^{10}$$

et en supposant que la proportion de a à b soit comme de 3 à 2, alors le Terme $210a^6b^4$ sera le plus grand, parce que les Exposants des Puissances de a et b , qui sont dans ce Terme, sont dans la proportion de 3 à 2 ; mais en supposant que la proportion de a à b ait été de 4 à 1, alors le Terme $45a^8b^2$ aurait été le plus grand.

LEMME 3.

Si un Evénement dépend du Hasard, de telle manière que les Probabilités de sa réalisation ou de son échec sont dans une proportion donnée, supposée être celle de a à b , et si un certain nombre d'Expériences est envisagé, afin d'observer combien de fois l'Evénement se produira ou non ; alors la Probabilité qu'il ne se produise ni plus qu'un nombre de fois noté $\frac{an}{a+b}+1$, ni moins qu'un nombre de fois noté $\frac{an}{a+b}-1$, sera trouvée comme suit :

Soient L et R également distants d'un Intervalle l du plus grand Terme ; soit aussi S la Somme des Termes compris entre L et R, y compris ces Extrémités, alors la Probabilité cherchée sera exprimée de manière exacte par $\frac{S}{(a+b)^n}$.

COROLLAIRE 8.

Le Rapport, pour une Puissance infinie notée n , du plus grand Terme à la Somme de tous les autres, sera correctement exprimée par la Fraction $\frac{a+b}{\sqrt{abnc}}$, où c désigne, comme avant, la Circonférence d'un Cercle de Rayon égal à l'Unité.

COROLLAIRE 9.

Si, pour une Puissance infinie, un Terme est distant du plus Grand d'un Intervalle l , alors le Logarithme Hyperbolique du Rapport de ce Terme au plus Grand sera exprimé par la Fraction $-\frac{(a+b)^2}{2abn} \times l$; pourvu que le Rapport de l à n ne soit pas un Rapport fini, mais qu'on puisse le considérer comme celui d'un nombre donné p à \sqrt{n} ; de telle manière qu'on puisse écrire l comme $p\sqrt{n}$, auquel Cas les deux Termes L et R seront égaux.

COROLLAIRE 10.

Si les Probabilités de succès et d'échec sont dans un Rapport donné d'inégalité, les Problèmes relatifs à la Somme des Termes du Binôme $(a+b)^n$ seront résolus avec la même facilité que ceux dans lesquels les Probabilités de succès et d'échec sont dans un Rapport d'Egalité.

REMARQUE I.

De ce qui a été dit, il suit, que le Hasard perturbe très peu les Evénements qui dans leur Institution naturelle sont amenés à se produire ou non, suivant une certaine Loi déterminée ; car si afin d'aider notre compréhension, nous imaginons une pièce ronde de Métal, avec deux faces opposées polies, ne différant en rien sinon leur couleur, qu'on peut supposer être l'une blanche, et

l'autre noire ; il est clair que nous pouvons dire, que cette pièce peut avec une égale facilité montrer une face blanche ou noire, et nous pouvons même supposer qu'elle a été construite en vue de montrer quelquefois une face, quelquefois l'autre, et que en conséquence si elle est lancée le Hasard décidera de l'apparition ; Mais nous avons vu dans notre LXXIIème Problème, que bien que le Hasard puisse produire une inégalité d'apparition, et même une inégalité plus grande suivant la durée où il peut s'exercer, pourtant les apparitions, d'une manière ou d'une autre, tendront perpétuellement vers une proportion d'Egalité : Mais de plus, nous avons vu dans le présent Problème, que pour un grand nombre d'Expériences, comme 3600, on aurait un Rapport de Probabilités de plus de 2 contre 1, qu'une des faces, supposons la blanche, n'apparaisse pas plus fréquemment que 1830 fois, ni plus rarement que 1770, ou en d'autres Termes, que l'écart avec l'Egalité parfaite ne soit d'un côté ou de l'autre supérieur à la $1/120$ partie du nombre total d'apparitions ; et par la même Règle, que si le nombre d'Essais était 14400 au lieu de 3600, alors il y aurait encore un Rapport de Probabilités de plus de 2 contre 1, que la déviation avec l'Egalité parfaite n'excède pas d'un côté ou de l'autre $1/260$ du total : et en 1000000 Essais il y aurait un Rapport de Probabilités de plus de 2 contre 1, que la déviation de l'Egalité parfaite ne soit pas plus que la $1/2000$ partie du tout. Mais les Rapports de Probabilités augmenteraient à une vitesse prodigieuse, si au lieu de prendre des limites aussi étroites des deux côtés du Terme d'Egalité, que celles représentées par $1/2\sqrt{n}$, nous doublions ces Limites ou les triplions : car dans le premier Cas le Rapport de Probabilités deviendrait 21 contre 1, et dans le second 369 contre 1, et encore plus immensément grand si nous devions les quadrupler, et à la fin serait infiniment grand ; et pourtant que nous les doublions, triplions ou quadruplions, &c. l'Etendue des ces Limites sera dans une proportion insignifiante du tout, et nulle à la fin, si le tout est infini ; la raison de ceci sera facilement perçue par les Mathématiciens, qui savent, que la Racine carrée d'une Puissance est avec cette Puissance dans une proportion d'autant plus petite, que son Exposant est grand.

On peut appliquer aussi ce que nous avons dit à un Rapport d'Inégalité, comme il apparaît de notre 9ème Corollaire. Et ainsi dans tous les Cas, on trouvera, que *bien que le Hasard produise des Irrégularités, cependant les Rapports de Probabilités seront infiniment grands, et que avec l'avancement du Temps, ces Irrégularités n'auront aucune proportion avec le retour de l'Ordre qui résulte naturellement du DESSEIN ORIGINEL.*

REMARQUE II.

De même que, avec la Supposition d'une certaine Loi déterminée suivant laquelle un Evénement doit avoir lieu, nous démontrons que le Rapport des Réalisations [Succès : Echecs] approchera continuellement cette Loi, d'autant que les Expériences ou Observations sont multipliées : de même, *inversement*, si à partir d'Observations innombrables, nous trouvons que le Rapport des Evénements converge vers une quantité déterminée, comme le Rapport de P à Q ; alors nous concluons que ce Rapport exprime la Loi déterminée suivant laquelle l'Evénement se produit.

Car supposons que cette Loi soit exprimée non par le rapport P : Q, mais par un autre R : S ; alors le Rapport des Evénements convergerait vers ce dernier, non vers le premier : ce qui contredit notre *Hypothèse*. Et une Absurdité semblable, ou plus grande, découle, si nous supposons que

l'Événement ne se produise pas suivant une Loi, mais d'une manière à la fois décousue et incertaine ; car alors les Événements ne convergeraient vers aucun Rapport fixé.

De plus, de même qu'il est ainsi démontrable qu'il y a, dans la constitution des choses, certaines Lois suivant lesquelles les Événements se produisent, il n'est pas moins évident à partir de l'Observation, que ces Lois servent à des buts sages, utiles et bénéfiques ; préserver l'Ordre bien établi de l'Univers, propager les différentes Espèces d'Êtres, et fournir au Genre pensant des degrés de bonheur tels qu'ils conviennent à leur Etat.

Mais de telles lois, autant que le Dessein et le But originels de leur Etablissement, doivent être toutes *issues de l'extérieur* : l'*Inertie* de la matière, et la nature de tous les Êtres créés, rendant impossible qu'une chose puisse modifier sa propre essence, ou donner à elle-même, ou à une autre chose, une détermination ou un penchant original. Et ainsi, si nous ne nous aveuglons pas nous-mêmes avec de la poussière métaphysique, nous serons conduits, d'une manière rapide et évidente, à la reconnaissance du grand CREATEUR et MAITRE de toutes choses ; *Lui-même toute sagesse, toute puissance et bonté.*

Mr. *Nicolas Bernoulli*, Homme très savant et très bon, en ne reliant pas la dernière partie de notre raisonnement avec la première, fut conduit à écarter et même vilipender cet Argument des *Causes finales*, sur lequel ont tant insisté nos meilleurs Ecrivains ; en particulier, pour l'Exemple de la quasi-égalité des nombres de Naissances *mâles* et *femelles*, introduit par une excellente Personne le regretté Dr. *Arbuthnot*, dans les *Phil. Trans.* N° 328.

Mr. *Bernoulli* tire des Tables d'Observations faites 82 ans durant, c'est-à-dire de 1629 à 1711, que le nombre de Naissances à *Londres* était, en *moyenne*, d'environ 14000 par an : et de même, que le nombre de *Mâles* était à celui des *Femelles*, ou plutôt la facilité de leur naissance, comme environ 18 à 17. Mais il pense que c'est le comble de la légèreté d'en tirer quelque Argument contre l'Influence du *Hasard* dans l'apparition des deux Sexes. En effet, dit-il, "Si on lance 14000 dés, ayant chacun 35 faces, 18 blanches et 17 noires, il y a une grande Probabilité que les nombres de faces blanches et noires s'approcheront l'un de l'autre autant, ou plus, que les nombres de Garçons et de Filles le font dans les Tables."

Ceci amène la courte réponse suivante : Dr. *Arbuthnot* n'a jamais dit "qu'en supposant *fixé* auparavant que la facilité de naissance d'un Mâle était à celle de naissance d'une femelle comme environ le Rapport d'égalité, ou comme 18 à 17 ; il était *étonné* que le Rapport de ces nombres de naissances de Mâles et de Femelles puisse, pendant beaucoup d'années, rester entre des bornes aussi étroites ." c'est la seule Proposition contre laquelle le raisonnement de Mr. *Bernoulli* a une certaine force.

Mais il aurait pu dire, et nous insistons encore dessus, que "de même que, des Observations, nous pouvons, avec Mr. *Bernoulli*, déduire que les facilités de naissance des deux Sexes sont environ dans un Rapport d'égalité ; ainsi de la découverte de ce Rapport, et *servant manifestement un but avisé*, nous concluons que le Rapport lui-même, ou si vous voulez la *Forme du Dé*, est un Effet de l'*Intelligence* et du *Dessein*." De même que si on nous montrait un certain nombre de Dés, chacun avec

18 faces blanches et 17 noires, ce qui est la supposition de Mr. *Bernoulli*, nous ne douterions aucunement que ces Dés aient été fabriqués par quelque Artiste ; et que leur forme ne devait rien au *Hasard*, mais était adaptée au but particulier qu'il avait en Vue.

Ainsi il s'en fallait de beaucoup qu'on efface l'impression que l'autorité d'un si grand nom puisse porter préjudice à notre argument. Qui, après tout, étant au niveau de la compréhension la plus basse, et tombant sous le sens commun de l'humanité, ne demandait pas de Démonstration formelle, si ce n'est à cause des subtilités scholastiques dont il peut être embarrassé ; et à cause de l'abus de certains mots et expressions ; qu'on imagine quelquefois avoir un sens simplement parce qu'ils sont souvent prononcés.

Le *Hasard*, tel que nous le comprenons, suppose l'*Existence* des choses, et leurs *Propriétés* généralement connues : par exemple, que si on lance un certain nombre de Dés, chacun d'eux s'établira sur l'une ou l'autre de ses Bases. Après quoi, la *Probabilité* d'une Chance donnée, c'est-à-dire d'une disposition particulière des Dés, devient un sujet d'Investigation aussi approprié que peut l'être une autre quantité ou Rapport.

Mais le *Hasard*, dans des écrits ou discours athées, est un son complètement dénué de sens : il n'implique de détermination à aucun *mode d'Existence* ; ni en fait à l'*Existence* elle-même, pas plus qu'à la *non-existence* ; il ne peut être ni défini ni compris ; aucune Proposition le concernant ne peut être ni affirmée ni réfutée, excepté celle-ci, "C'est simplement un mot".

On peut dire la même chose d'autres mots d'usage fréquent : comme *destin*, *nécessité*, *nature*, une *direction de la nature* en opposition à l'*Energie divine* : tous, ainsi qu'on les utilise en certaines occasions, sont simplement des sons : et cependant, par un maniement plein d'art, ils servent à fonder des conclusions spécieuses : qui toutefois, aussitôt que la fausseté latente du *Terme* est détectée, n'apparaissent pas moins absurdes en elles-mêmes, qu'elles ne sont ordinairement nuisibles à la société.

J'ajouterai seulement, Qu'on peut appliquer utilement cette méthode de Raisonnement à d'autres Recherches très intéressantes ; si ce n'est pour forcer l'Assentiment des autres par une Démonstration rigoureuse, au moins pour la Satisfaction du Chercheur lui-même : et je conclurai cette --Remarque avec un passage tiré de l'*Ars Conjectandi* de Mr. *James Bernoulli*, Part IV, Chap. 4, où cet Auteur judicieux et précis introduit ainsi sa Solution du Problème pour *Trouver les Limites entre lesquelles, par la répétition des Expériences, la Probabilité d'un Evénement peut approcher indéfiniment une Probabilité donnée*, "Hoc igitur est illud Problema &c." Voici, dit-il, le Problème dont j'ai proposé ici la Publication, après l'avoir mis à jour il y a vingt ans déjà, et dont la nouveauté d'une part, la grande utilité d'autre part, jointes à une égale difficulté, peut ajouter du poids et du prix à tous les autres Points de cette Doctrine. Cependant il y a des Auteurs, en fait d'une Classe très différente de celle de *James Bernoulli*, qui laissent entendre que la *Doctrine des Probabilités* ne pourrait trouver de place dans une Recherche sérieuse ; et que des Etudes de ce type, aussi simples et faciles soient-elles, disqualifient un homme pour raisonner sur tout autre sujet. Le Lecteur choisira.

ANNEXE II

Pierre-Simon LAPLACE
(1749-1827)

THÉORIE ANALYTIQUE DES PROBABILITÉS
(1812)
in Oeuvres, Paris, 1847

CHAPITRE III.

DES LOIS DE LA PROBABILITÉ QUI RÉSULTENT DE LA MULTIPLICATION
INDÉFINIE DES ÉVÉNEMENTS.

16. A mesure que les événements se multiplient, leurs probabilités respectives se développent de plus en plus : leurs résultats moyens et les bénéfices ou les pertes qui en dépendent convergent vers des limites dont ils approchent avec des probabilités toujours croissantes. La détermination de ces accroissements et de ces limites est une des parties les plus intéressantes et les plus délicates de l'analyse des hasards.

Considérons d'abord la manière dont les possibilités de deux événements simples, dont un seul doit arriver à chaque coup, se développent lorsqu'on multiplie le nombre de coups. Il est visible que l'événement dont la facilité est la plus grande doit probablement arriver plus souvent dans un nombre donné de coups ; et l'on est porté naturellement à penser qu'en répétant les coups un très grand nombre de fois, chacun de ces événements arrivera proportionnellement à sa facilité, que l'on pourra ainsi découvrir par l'expérience. Nous allons démontrer analytiquement cet important théorème.

On a vu dans le n° 6 que si p et $1 - p$ sont les probabilités respectives de deux événements a et b , la probabilité que dans $x + x'$ coups, l'événement a arrivera x fois, et l'événement b , x' fois est égale à

$$\frac{1.2.3..\dots(x+x')}{1.2.3..\dots x.1.2.3..\dots x'} \cdot p^x \cdot (1-p)^{x'}$$

c'est le $(x+1)^{\text{ième}}$ terme du binôme $[p + (1-p)]^{x+x'}$. Considérons le plus grand de ces termes que nous désignons par k : le terme antérieur sera $\frac{kp}{1-p} \cdot \frac{x'}{x+1}$, et le terme suivant sera $k \cdot \frac{1-p}{p} \cdot \frac{x}{x'+1}$.

Pour que k soit le plus grand terme, il faut que l'on ait à la fois

$$\frac{p}{1-p} < \frac{x+1}{x'} > \frac{x}{x'+1}$$

il est facile d'en conclure que si l'on fait $x + x' = n$, on aura

$$x < (n+1) \cdot p > (n+1) \cdot p - 1 :$$

ainsi x est le plus grand nombre entier compris dans $(n+1) \cdot p$; en faisant donc

$$x = (n+1) \cdot p - s$$

ce qui donne

$$p = \frac{x+s}{n+1}, \quad 1-p = \frac{x'+1-s}{n+1}, \quad \frac{p}{1-p} = \frac{x+s}{x'+1-s}$$

s sera moindre que l'unité. Si x et x' sont de très grands nombres, on aura à très-peu près,

$$\frac{p}{1-p} = \frac{x}{x'}$$

c'est-à-dire que les exposants de p et $1-p$, dans le plus grand terme du binôme, sont à fort peu près dans le rapport de ces quantités ; en sorte que de toutes les combinaisons qui peuvent avoir lieu dans un très grand nombre n de coups, la plus probable est celle dans laquelle chaque événement est répété proportionnellement à sa probabilité.

Le terme $l^{\text{ième}}$, après le plus grand, est

$$\frac{1.2.3\dots n}{1.2.3\dots(x-l).1.2.3\dots(x'+1)} \cdot p^{x-1} \cdot (1-p)^{x'+1}.$$

On a, par le n° 32 du premier livre,

$$1.2.3\dots n = n^{\frac{n+1}{2}} \cdot c^{-n} \cdot \sqrt{2\pi} \cdot \left\{ 1 + \frac{1}{12n} + \text{etc.} \right\};$$

ce qui donne

$$\frac{1}{1.2.3\dots(x-l)} = (x-l)^{l-x-\frac{1}{2}} \cdot \frac{c^{x-l}}{\sqrt{2\pi}} \cdot \left\{ 1 - \frac{1}{12(x-l)} - \text{etc.} \right\},$$

$$\frac{1}{1.2.3\dots(x'+1)} = (x'+1)^{-x'-l-\frac{1}{2}} \cdot \frac{c^{x'+1}}{\sqrt{2\pi}} \cdot \left\{ 1 - \frac{1}{12(x'+1)} - \text{etc.} \right\}.$$

Développons le terme $(x-l)^{l-x-\frac{1}{2}}$: son logarithme hyperbolique est

$$\left(l - x - \frac{1}{2} \right) \cdot \left[\log x + \log \left(1 - \frac{l}{x} \right) \right];$$

or on a

$$\log \left(1 - \frac{l}{x} \right) = -\frac{l}{x} - \frac{l^2}{2x^2} - \frac{l^3}{3x^3} - \frac{l^4}{4x^4} - \text{etc.}$$

nous négligerons les quantités de l'ordre de $\frac{1}{n}$, et nous supposerons que l^2 ne surpasse point l'ordre

de n ; alors on pourra négliger les termes de l'ordre $\frac{l^4}{x^3}$, parce que x et x' sont de l'ordre n . On aura

ainsi

$$\left(l - x - \frac{1}{2} \right) \cdot \left[\log x + \log \left(1 - \frac{l}{x} \right) \right] = \left(l - x - \frac{1}{2} \right) \cdot \log x + l + \frac{l}{2x} - \frac{l^2}{2x} - \frac{l^3}{6x^2};$$

ce qui donne, en repassant des logarithmes aux nombres,

$$(x-l)^{l-x-\frac{1}{2}} = c^{l-\frac{l^2}{2x}} \cdot x^{l-x-\frac{1}{2}} \cdot \left(1 - \frac{l}{2x} - \frac{l^3}{6x^2} \right);$$

on aura pareillement

$$(x'+1)^{-x'-l-\frac{1}{2}} = c^{-l-\frac{l^2}{2x'}} \cdot x'^{-x'-l-\frac{1}{2}} \cdot \left(1 - \frac{l}{2x'} + \frac{l^3}{6x'^2} \right).$$

On a ensuite par ce qui précède, $p = \frac{x+s}{n+1}$, s étant moindre que l'unité ; en faisant donc $p = \frac{x-z}{n}$, z sera compris dans les limites $\frac{x}{n+1}$ et $-\frac{n-x}{n+1}$, et par conséquent il sera, abstraction faite du signe,

au-dessous de l'unité. La valeur de p donne $1-p = \frac{x'+z}{n}$; on aura donc par l'analyse précédente,

$$p^{x-l} \cdot (1-p)^{x+l} = \frac{x^{x-l} \cdot x'^{x+l}}{n^n} \cdot \left(1 + \frac{nz \cdot l}{xx'}\right);$$

par là on tire

$$\frac{1.2.3\dots n}{1.2.3\dots(x-l).1.2.3\dots(x'+l)} \cdot p^{x-l} \cdot (1-p)^{x+l} = \frac{\sqrt{n} \cdot c^{\frac{nl^2}{2xx'}}}{\sqrt{\pi} \cdot \sqrt{2xx'}} \cdot \left(1 + \frac{nzl}{xx'} + \frac{l(x'-x)}{2xx'} - \frac{l^3}{6x^2} + \frac{l^3}{6x'^2}\right).$$

On aura le terme antérieur au plus grand terme, et qui en est éloigné à la distance l , en faisant l négatif dans cette équation ; en réunissant ensuite ces deux termes, leur somme sera

$$\frac{2 \cdot \sqrt{n}}{\sqrt{\pi} \cdot \sqrt{2xx'}} \cdot c^{\frac{nl^2}{2xx'}}.$$

L'intégrale finie

$$\sum \cdot \frac{2 \cdot \sqrt{n}}{\sqrt{\pi} \cdot \sqrt{2xx'}} \cdot c^{\frac{nl^2}{2xx'}},$$

prise depuis $l = 0$ inclusivement, exprimera donc la somme de tous les termes du binôme $[p + (1-p)]^n$, comprise entre les deux termes, dont l'un a p^{x+l} pour facteur, et l'autre a p^{x-l} pour facteur, et qui sont ainsi équidistants du plus grand terme ; mais il faut retrancher de cette somme le plus grand terme, qui y est évidemment compris deux fois.

Maintenant, pour avoir cette intégrale finie, nous observerons que l'on a, par le n° 10 du premier livre, y étant fonction de l ,

$$\sum \cdot y = \frac{1}{c^{\frac{dy}{dl}} - 1} = \left(\frac{dy}{dl}\right)^{-1} - \frac{1}{2} \cdot \left(\frac{dy}{dl}\right)^0 + \frac{1}{12} \frac{dy}{dl} + \text{etc.}$$

d'où l'on tire, par le même numéro,

$$\sum \cdot y = \int y dl - \frac{1}{2} \cdot y + \frac{1}{12} \frac{dy}{dl} + \text{etc.} + \text{constante.}$$

y étant ici égal à $\frac{2 \cdot \sqrt{n}}{\sqrt{\pi} \cdot \sqrt{2xx'}} \cdot c^{\frac{nl^2}{2xx'}}$, les différentielles successives de y acquièrent pour facteur $\frac{nl}{2xx'}$ et ses puissances ; ainsi l étant supposé ne pouvoir être au plus que de l'ordre \sqrt{n} , ce facteur est de l'ordre $\frac{1}{\sqrt{n}}$, et par conséquent ses différentielles, divisées par les puissances respectives de dl , décroissent de plus en plus ; en négligeant donc, comme on l'a fait précédemment, les termes de l'ordre $\frac{1}{n}$, on aura, en faisant commencer avec l les deux intégrales finies et infiniment petites, et désignant par Y le plus grand terme du binôme,

$$\sum \cdot y = \int y dl - \frac{1}{2} \cdot y + \frac{1}{2} \cdot Y.$$

La somme de tous les termes du binôme $[p + (1-p)]^n$ compris entre les deux termes équidistants du plus grand terme, du nombre l , étant égale à $\sum y - \frac{1}{2} \cdot Y$, elle sera

$$\int y dl - \frac{1}{2} \cdot y;$$

et si l'on y ajoute la somme de ces termes extrêmes, on aura pour la somme de tous ces termes,

$$\int y dl + \frac{1}{2} \cdot y.$$

Si l'on fait

$$t = \frac{l\sqrt{n}}{\sqrt{2xx'}},$$

cette somme devient

$$\frac{2}{\sqrt{\pi}} \int dt \cdot c^{-t^2} + \frac{\sqrt{n}}{\sqrt{\pi} \cdot \sqrt{2xx'}} \cdot c^{-t^2}. \quad (o)$$

Les termes que l'on a négligés étant de l'ordre $\frac{1}{n}$, cette expression est d'autant plus exacte que n est grand : elle est rigoureuse lorsque n est infini. Il serait facile, par l'analyse précédente, d'avoir égard aux termes de l'ordre $\frac{1}{n}$ et des ordres supérieurs.

On a, par ce qui précède, $x = np + z$, z étant un nombre plus petit que l'unité ; on a donc

$$\frac{x+l}{n} - p = \frac{l+z}{n} = \frac{t \cdot \sqrt{2xx'}}{n \cdot \sqrt{n}} + \frac{z}{n};$$

ainsi la formule (o) exprime la probabilité que la différence entre le rapport du nombre de fois que l'événement a doit arriver, au nombre total de coups, et la facilité p de cet événement, est comprise dans les limites

$$\pm \frac{t \cdot \sqrt{2xx'}}{n \cdot \sqrt{n}} + \frac{z}{n}. \quad (l)$$

$\sqrt{2xx'}$ étant égal à

$$n \cdot \sqrt{2p(1-p) + \frac{2z}{n} \cdot (1-2p) - \frac{2z^2}{n^2}};$$

on voit que l'intervalle compris entre les limites précédentes est de l'ordre $\frac{1}{\sqrt{n}}$.

Si la limite de t , que nous désignerons par T , est supposée invariable, la probabilité déterminée par la fonction (o) reste la même à très-peu près ; mais l'intervalle compris entre les limites (l) diminue sans cesse à mesure que les coups se répètent, et il devient nul, lorsque leur nombre est infini.

Cet intervalle étant supposé invariable, lorsque les événements se multiplient, T croît sans cesse, et à fort peu près comme la racine carrée du nombre des coups. Mais lorsque T est considérable, la formule (o) devient par le n° 27 du premier livre,

$$1 - \frac{c^{-T^2}}{2T\sqrt{\pi}} \cdot \frac{1}{1 + \frac{q}{1 + \frac{2q}{1 + \frac{3q}{1 + \text{etc.}}}}} + \frac{c^{-T^2}}{\sqrt{2n\pi} \cdot \left[p(1-p) + \frac{z}{n(1-2p)} - \frac{z^2}{n^2} \right]}$$

q étant égal à $\frac{1}{2T^2}$. Lorsqu'on fait croître T, c^{-T^2} diminue avec une extrême rapidité, et la probabilité précédente s'approche rapidement de l'unité, à laquelle elle devient égale lorsque le nombre de coups est infini.

Il y a ici deux sortes d'approximations : l'une d'elles est relative aux limites prises de part et d'autre de la facilité de l'événement a ; l'autre approximation se rapporte à la probabilité que le rapport des arrivées de cet événement, au nombre total des coups, sera dans ces limites. La répétition indéfinie des coups accroît de plus en plus cette probabilité, les limites restant les mêmes ; elle resserre de plus en plus l'intervalle de ces limites, la probabilité restant la même. Dans l'infini, cet intervalle devient nul, et la probabilité se change en certitude.

L'analyse précédente réunit à l'avantage de démontrer ce théorème, celui d'assigner la probabilité que dans un grand nombre n de coups, le rapport des arrivées de chaque événement sera compris dans des limites données. Supposons, par exemple, que les facilités des naissances des garçons et des filles soient dans le rapport de 18 à 17, et qu'il naisse dans une année 14000 enfants ; on demande la probabilité que le nombre des garçons ne surpassera pas 7363, et ne sera pas moindre que 7037.

Dans ce cas, on a

$$p = \frac{18}{35}, x = 7200, x' = 6800, n = 14000, l = 163 ;$$

la formule (o) donne à fort peu près 0,994303 pour la probabilité cherchée.

Si l'on connaît le nombre de fois que sur n coups, l'événement a est arrivé, la formule (o) donnera la probabilité que sa facilité p supposée inconnue sera comprise dans des limites données. En effet, si l'on nomme i ce nombre de fois, on aura, par ce qui précède, la probabilité que la différence $\frac{i}{n} - p$ sera comprise dans les limites $\pm \frac{T \cdot \sqrt{2xx'}}{n \cdot \sqrt{n}} + \frac{z}{n}$; par conséquent, on aura la probabilité que p sera compris dans les limites

$$\frac{i}{n} \mp \frac{T \cdot \sqrt{2xx'}}{n \cdot \sqrt{n}} - \frac{z}{n}.$$

La fonction $\frac{T \cdot \sqrt{2xx'}}{n \cdot \sqrt{n}}$ étant de l'ordre $\frac{1}{\sqrt{n}}$, on peut, en négligeant les quantités de l'ordre $\frac{1}{n}$, y substituer i au lieu de x, et n - i au lieu de x' ; les limites précédentes deviennent ainsi, en négligeant les termes de l'ordre $\frac{1}{n}$,

$$\frac{i}{n} \mp \frac{T \cdot \sqrt{2i(n-i)}}{n \cdot \sqrt{n}} ;$$

et la probabilité que la facilité de l'événement a est contenue dans ces limites est égale à

$$\frac{2}{\sqrt{\pi}} \cdot \int dt \cdot c^{-t^2} + \frac{\sqrt{n} \cdot c^{-T^2}}{\sqrt{\pi} \cdot \sqrt{2i(n-i)}}. \quad (o')$$

On voit ainsi qu'à mesure que les événements se multiplient, l'intervalle des limites se resserre de plus en plus, et la probabilité que la valeur de p tombe dans ces limites approche de plus en plus de l'unité ou de la certitude. C'est ainsi que les événements, en se développant, font connaître leurs probabilités respectives.