

Une démonstration élémentaire du théorème de Jacques Bernoulli par Charles de La Vallée Poussin

Hervé Le Ferrand
Université de Bourgogne, Institut de Mathématiques de Bourgogne



Figure 1: *Timbre suisse à l'effigie de Jacques Bernoulli. Édité en 1994 à l'occasion du Congrès International des Mathématiciens, Zürich.*

CONTEXTE

Le 16 Février 1907, le célèbre mathématicien belge Charles de La Vallée Poussin adresse, au mathématicien français Robert de Montessus de Ballore¹, une lettre, dans laquelle il donne une nouvelle démonstration du théorème de Jacques Bernoulli, c'est à dire de la loi faible des grands nombres dans le cas d'une urne dite de Bernoulli. Si Robert de Montessus, dont les recherches en ce début de siècle portent plutôt sur les fractions continues algébriques, s'intéresse au théorème de Bernoulli, c'est qu'il est en train d'écrire un livre de leçons de probabilités [Montessus, 1908]. Cet ouvrage intéresse certains historiens des sciences car il reprend les travaux de Louis Bachelier sur la théorie de la spéculation qui datent de 1900.

1. Robert de Montessus est lauréat de l'Académie des Sciences en 1906 pour ses travaux sur les fractions continues algébriques.



Figure 2 : (à g.) Charles de La Vallée Poussin (1866-1962) ; (à dr.) Robert de Montessus de Ballore (1870-1937).

En 1907, Charles de La Vallée Poussin² est professeur à l'Université Catholique de Louvain. Robert de Montessus est quant à lui maître de conférences³ à l'Université Catholique de Lille. Ils sont tous deux catholiques et proches géographiquement. De plus, les deux mathématiciens font partie de la Société Scientifique de Bruxelles créée en 1875⁴. Robert de Montessus n'est d'ailleurs pas le seul professeur lillois à faire partie de cette société : c'est le cas de son ami et collègue de l'Université Catholique de Lille, Robert d'Adhémar.

Cette même année 1907, le mathématicien belge Paul Mansion⁵, dont le nom apparaît dans la lettre de Charles de La Vallée Poussin, est le secrétaire de la Société Scientifique de Bruxelles. La Société est organisée en sections. La première section est celle des Mathématiques, Astronomie, Géodésie et Mécanique, présidée par Georges Humbert. Charles de La Vallée Poussin et Robert d'Adhémar en

2. On pourra lire la description de sa carrière à l'Université Catholique de Louvain faite par Jean Mawhin [Mawhin J., 2011].

3. Il est nommé professeur en 1908.

4. Un des créateurs de cette société savante est le père jésuite, et scientifique, belge Ignace Carbonnelle (1829-1889) [Voelke J.-D., 2005]. En 1875, Ignace Carbonnelle enseigne au Collège Saint Michel de Bruxelles.

5. Paul Mansion est alors professeur émérite de l'Université de Gand.

assurent la vice-présidence et Hector Dutordoir⁶ le secrétariat. Edouard Goedseels, cité aussi dans la lettre, est membre fondateur de la Société et administrateur-inspecteur de l'Observatoire royal de Belgique⁷. C'est dans la revue de la Société, les *Annales de la Société Scientifique de Bruxelles*, que les travaux des trois mathématiciens belges sur le théorème de Bernoulli sont publiés. On peut donc considérer que la lettre s'inscrit dans un processus d'échanges entre deux membres d'une société savante, échanges dans lesquels figurent d'autres noms de sociétaires.



Figure 3 : Paul Mansion (1844-1919)

Charles de La Vallée Poussin a écrit, d'après l'analyse d'Eugene Seneta dans [LVP, 2001], en théorie des Probabilités uniquement deux articles. Cette observation rend encore plus intéressantes les lettres adressées à Robert de Montessus. Les deux articles du mathématicien belge sur le théorème de Bernoulli, *Etude sur le théorème de Bernoulli* et *Démonstration nouvelle du théorème de Bernoulli* paraissent tous les deux au cours de l'année 1907 dans les *Annales de la Société Scientifique de Bruxelles* [LVP 1907 & 1907-2]. Ainsi, dans une lettre datée du 1^{er} février 1907, Charles de La Vallée Poussin écrit à Robert de Montessus⁸ :

Mon article sur le théorème de Bernoulli va être imprimé par la Société Scientifique de Bruxelles. Je ne pourrai vous le communiquer qu'en épreuves d'ici à une quinzaine de jours, ce qui sera fait.

6. Hector Dutordoir est un ingénieur en chef, directeur du service technique provincial, demeurant à Gand.

7. Édouard Goedseels (1857-1928) était aussi professeur à l'Université Catholique de Louvain [Mawhin, 2011].

8. Fonds Robert de Montessus, Université Pierre-et-Marie-Curie Paris-VI.

Il s'agit du premier article parmi les deux articles cités plus haut. Quant au second article, Charles de La Vallée Poussin indique, dans sa lettre du 16 février :

Enfin M. Mansion est en possession d'un travail manuscrit que je lui ai remis dans la dernière séance de la société et où j'établis [...]

Comment Charles de La Vallée Poussin en est-il venu à écrire sur le théorème de Bernoulli ? Dans cette même lettre, il fait référence à un article d'Edouard Goedseels paru quelques années plus tôt dans les Annales de la Société Scientifique de Bruxelles [Goedseels, 1895] :

M. Goedseels a donné dans les Annales de la société scientifique t. XVII p. 8 une démonstration très simple du théorème de Bernoulli.

L'article de Goedseels est publié en 1895. Ensuite deux articles de Paul Mansion sur le théorème de Bernoulli, paraissent en 1902 et 1904 [Mansion, 1902 & 1904]. Les deux articles de La Vallée Poussin s'inscrivent donc dans la continuité des travaux de Goedseels et de Mansion. Les trois mathématiciens cherchent à

préciser la vitesse de convergence de la probabilité $P\left(\left|\frac{S_n}{n} - p\right| < \varepsilon\right)$ vers 1, $n \rightarrow \infty$,

S_n étant une somme de variables de Bernoulli indépendantes de probabilité p de succès. Dans la lettre du 16 février, on trouve l'énoncé du théorème VII de l'article. Nous reviendrons sur la formulation de la loi faible des grands nombres⁹ dans le paragraphe suivant. Paul Mansion [1904] souligne la difficulté de l'utilisation du théorème de Bernoulli, dans sa forme originale, dans les applications pratiques. Les travaux de Mansion et de La Vallée Poussin, pour reprendre l'expression de Seneta [2013], présagent d'un retour aux méthodes exactes.

Signalons qu'en 1907, date de la lettre et des publications de La Vallée Poussin, les probabilités ne sont pas axiomatisées. Il faut attendre l'année 1933, année de publication de l'ouvrage de Kolmogorov, pour avoir l'axiomatisation classique que nous connaissons¹⁰. Deux années plus tard, en 1909, Emile Borel publie *Les probabilités dénombrables et leurs applications arithmétiques*, où se trouve la première version d'un résultat que l'on nomme loi forte des grands nombres. La loi forte des grands nombres traduit une *convergence presque-sûre*

9. Notons que ce théorème a conservé le nom de loi, initialement donné par Siméon-Denis Poisson dans [Poisson S.D., 1837] (voir [Fischer H., 2011]).

10. Une version anglaise paraît dans les années 1950 : voir [Kolmogorov A.N., 1956]. Il existe des tentatives d'axiomatisation antérieures à 1930, comme celle en 1919 de Richard von Mises [Mises (von) R., 1919].

tandis que la loi faible des grands nombres indique une *convergence en probabilité* selon le vocabulaire à présent consacré.

Loi forte des grands nombres

Soit X_1, X_2, \dots une suite infinie de variables aléatoires réelles indépendantes et de même loi de probabilités, définie par sa fonction de répartition $F(x) = P(X_1 \leq x)$. Alors, la limite de la moyenne d'ordre n

définie par $l = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i$ existe avec probabilité 1 si et seulement si

l'espérance de la valeur absolue de $X := X_1$, soit $E(|X|) = \int_{-\infty}^{+\infty} |x| dF(x)$, est finie. Dans ce cas, on a, nécessairement $l = E(X)$.

THÉORÈME DE BERNOULLI, LOIS DES GRANDS NOMBRES

En 1713, soit huit années après la mort de Jacques Bernoulli¹¹, paraît l'ouvrage, écrit en latin, *Ars Conjectandi*.

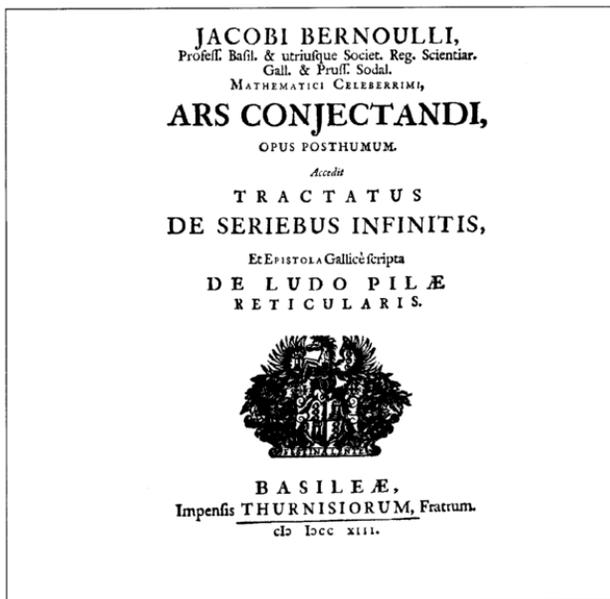


Figure 4 : couverture de l'Ars Conjectandi (1713).

Le résultat qui nous intéresse, nommé à présent *théorème de Bernoulli* ou encore *loi faible des grands nombres de Bernoulli*, se trouve dans la partie 4,

11. C'est au neveu de Jacques, Nicolas (I) Bernoulli (Bâle 1687-Bâle 1759) que l'on doit cette publication. Nicolas

chapitre 5 de l'ouvrage. Jacques Bernoulli a débuté ses travaux sur la théorie des probabilités vers 1685. Son ouvrage couronne sa réflexion sur, selon l'expression de Lanier-Trotoux [Lanier & Trotoux, 1995], le *couple fréquence-probabilité*. Ainsi Siméon-Denis Poisson en 1837 écrit-il au sujet du théorème de Bernoulli :

Le théorème sur lequel est fondée la règle précédente est dû à Jacques Bernouilli [sic], 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 [...]

Dans le préambule de son ouvrage, Poisson souligne l'aspect fondamental de la loi des grands nombres :

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, c'est-à-dire sans que leur variation soit progressive dans aucun sens déterminé, on trouvera, entre ces nombres, des rapports à très peu près constants. Pour chaque nature de choses, ces rapports auront une valeur spéciale dont ils s'écarteront de moins en moins, à mesure que la série des événements observés augmentera davantage, et qu'ils atteindraient rigoureusement s'il était possible de prolonger cette série à l'infini.

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 Bernouilli.

Pour Kolmogorov, la loi des grands nombres marque réellement le début de la théorie des probabilités (page 12 de [AMS-LMS, 2000]) :

The cognitive value of probability theory is due to the fact that large-scale random phenomena in their cumulative effect create strict regularities. The very concept of mathematical probability would be barren if it did not find its realization as the frequency of occurrence of some result when uniform

condition are repeated many times. Therefore, the work of Pascal and Fermat can be regarded as only the prehistory of probability theory, while its genuine history begins with Bernoulli's law of large numbers.

Loi faible des grands nombres

Un énoncé moderne de la loi faible des grands nombres est le suivant : soit $(X_n)_n$ une suite de variables aléatoires indépendantes et de même loi, on considère la suite des moyennes arithmétiques des X_i , i.e. $(\bar{X}_n)_n$ où $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$. Alors si la loi commune des X_i admet un moment d'ordre 2 ($m = E(X_i)$ et $\sigma^2 = Var(X_i)$), alors la suite $(\bar{X}_n)_n$ converge en probabilité vers la constante m , c'est à dire que pour $\varepsilon > 0$ quelconque donné, $\lim_{n \rightarrow +\infty} P(|\bar{X}_n - m| > \varepsilon) \rightarrow 0$.

La Vallée Poussin, comme Goedseels et Mansion, se place dans le cas d'un tirage dans une urne de Bernoulli. On effectue n tirages avec remise dans une urne contenant des boules blanches (succès) et des boules noires (échec) en proportions respectives p et q ($p + q = 1$) et on s'intéresse au nombre de succès. On peut voir ce nombre comme une variable aléatoire X . La variable X est la somme des variables indépendantes X_i mesurant le succès au $i^{\text{ème}}$ tirage ($P(X_i = 1) = p$) pour $i = 1$ à n . Les variables X_i suivent la loi de Bernoulli de paramètre p et X suit alors une loi binomiale de paramètres n et p :

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n.$$

Les quantités à droite sont les différents termes du développement de $(p+q)^n$:

$$1 = (p+q)^n = \sum_{x=0}^n \binom{n}{x} p^x (1-p)^{n-x}.$$

Le théorème de Bernoulli nous indique que la limite, quand $n \rightarrow +\infty$, de

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - p\right| > \varepsilon\right) = P\left(\left|\frac{X}{n} - p\right| > \varepsilon\right) = P(|X - np| > n\varepsilon)$$

est 0 ou, ce qui revient au même, que la limite de

$$P(|X - np| \leq n\varepsilon) = P(n(p - \varepsilon) \leq X \leq n(p + \varepsilon))$$

vaut 1. On peut formuler¹² cela par :

$$\lim_{n \rightarrow +\infty} \sum_{x=n(p-\varepsilon)}^{n(p+\varepsilon)} \binom{n}{x} p^x (1-p)^{n-x} = 1 \quad (1)$$

@@@@@@

L'idée principale dans la démonstration du théorème par Bernoulli est de considérer le plus grand terme apparaissant dans le développement de $(p+q)^n$. Michel Henry [2004] analyse la démonstration de Bernoulli telle qu'elle est donnée dans l'*Ars Conjectandi* paru en 1713. Bernoulli se livre à une étude très fine des coefficients binomiaux, utilisant même des logarithmes décimaux (voir aussi la traduction de Oscar Sheynin, [Bernoulli,2005]). En 1837, Poisson redonne dans son livre *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, une démonstration du théorème de Bernoulli, et revient sur l'idée qui est de considérer le plus grand terme apparaissant dans le développement du binôme. Il donne ensuite différentes étapes (voir [Lanier & Trotoux]), sans cependant manipuler d'inégalités. Comme l'indique La Vallée Poussin, on trouve dans le *Traité d'Analyse*, tome II, de Camille Jordan, en 1894, une démonstration *rigoureuse* du théorème de Bernoulli. Jordan utilise la fonction Γ et des approximations de celle-ci. Notons que la formule de Stirling, qui donne le comportement de $n!$ quand n tend vers l'infini, a été utilisée par Abraham de Moivre et Pierre-Simon de Laplace dans leurs démonstrations respectives du théorème que l'on nomme à présent *théorème central limite*. C'est donc une tout autre démarche que propose La Vallée Poussin, plus proche du travail original de Bernoulli, mais plus simple – d'ailleurs il n'hésite pas à dire de ses prédécesseurs :

À mon avis, ces démonstrations ont le tort de s'appuyer sur la formule de Stirling et de faire appel à des moyens analytiques inutiles pour le résultat qu'on obtient.

12. Les nombres $n(p-\varepsilon)$ et $n(p+\varepsilon)$ ne sont pas nécessairement des nombres entiers, mais pour ne pas alourdir l'écriture, nous faisons comme si c'en étaient.

LA DÉMONSTRATION DE CHARLES DE LA VALLÉE POUSSIN DANS SA LETTRE

Nous transcrivons ci-dessous une partie de la démonstration de Charles de La Vallée Poussin (en mode citation¹³), en expliquant les différentes étapes et notations (en mode texte).

$$(p+q)^N = \sum_{k=0}^N \frac{N!}{k!(N-k)!} p^{N-k} q^k, \quad (2)$$

Appelons A_k le terme générique de sommation : $A_0 = p^N$ et $A_N = q^N$. L'auteur suppose, et démontre qu'entre ces deux termes, existe un terme A_k plus grand que tous les autres. On le sait pour $p = q = \frac{1}{2}$, où les A_k sont proportionnels à la suite des coefficients binomiaux, pour lesquels on a effectivement une croissance puis une décroissance : 1 2 1 / 1 3 3 1 / 1 4 6 4 1, etc. (triangle de Pascal). Démontrons-le pour p et q quelconques, en examinant le rapport entre deux A_k successifs :

$$A_k > A_{k-1} \Leftrightarrow (N-k+1)q > kp \Leftrightarrow (N+1)q > k \Leftrightarrow k < (Nq-p) + 1 \quad (3)$$

$$A_k < A_{k-1} \Leftrightarrow k > (Nq-p) + 1 \quad (4)$$

Appelons n la partie entière de $(Nq-p) + 1$: on a $A_n > A_{n-1}$ (suivant (3) ci-dessus) et $A_{n+1} < A_n$ (suivant (4) ci-dessus). La suite est donc décroissante à partir de A_n , qui est bien le plus grand terme. Reprenons à présent une des notations de La Vallée Poussin en considérant la suite T qui démarre à A_n ($T_0 = A_n$: il ne s'intéresse pas aux premiers termes, pendant laquelle la suite croît). C'est bien le même n (partie entière de $(Nq-p) + 1$, avec ε positif) dont il s'agit :

On pourra poser (ε étant une fraction) $n = Nq - p + \varepsilon$,

d'où $N - n = (N+1)p - \varepsilon$ et $n + 1 = (N+1)q + \varepsilon$.

Il vient alors :

$$T_1 = T_0 \frac{N-n}{n+1} \frac{q}{p} = T_0 \frac{(N+1)p - \varepsilon}{(N+1)q + \varepsilon} \frac{q}{p} = T_0 \frac{1 - \frac{\varepsilon}{(N+1)p}}{1 + \frac{\varepsilon}{(N+1)q}} < \frac{T_0}{1 + \frac{\varepsilon}{(N+1)pq}} \quad (5)$$

13. Nous avons changé, y compris dans les citations, deux notations utilisées par la Vallée Poussin : l'entier à la puissance duquel est élevé $(p+q)$ est non plus μ mais N ; l'indice générique de la somme est k , et non plus n .

La dernière étape en (5) nécessite démonstration. On peut se demander comment La Vallée Poussin a l'idée de cette inégalité, symétrique en p et q dans son membre de droite. Il s'agit de prouver que :

$$\left(1 + \frac{\varepsilon}{(N+1)pq}\right) \left(1 - \frac{\varepsilon}{(N+1)p}\right) < 1 + \frac{\varepsilon}{(N+1)q}$$

Faisons, avec le changement de notation (pour alléger) $a = \varepsilon / (N + 1)$, la différence :

$$\left(1 + \frac{a}{q}\right) - \left(1 + \frac{a}{pq}\right) \left(1 - \frac{a}{p}\right) = \frac{a}{q} - \frac{a}{pq} + \frac{a}{p} + \frac{a^2}{p^2q} = \frac{1}{pq} (ap - a + aq) + \frac{a^2}{p^2q} = \frac{a^2}{p^2q} > 0$$

puisque'on a $p + q = 1$. L'inégalité en (5) est donc démontrée. Reprenons le texte de La Vallée Poussin :

$$T_2 = T_1 \frac{N-n-1}{n+1+1} \frac{q}{p} = T_1 \frac{1 - \frac{\varepsilon+1}{(N+1)p}}{1 + \frac{\varepsilon+1}{(N+1)q}} < \frac{T_1}{1 + \frac{\varepsilon+1}{(N+1)pq}} \quad (6)$$

$$T_{\lambda+1} = T_\lambda \frac{N-n-\lambda}{n+1+\lambda} \frac{q}{p} = T_\lambda \frac{1 - \frac{\varepsilon+\lambda}{(N+1)p}}{1 + \frac{\varepsilon+\lambda}{(N+1)q}} < \frac{T_\lambda}{1 + \frac{\varepsilon+\lambda}{(N+1)pq}} \quad (7)$$

En multipliant ces inégalités membre à membre ; il vient, a fortiori (en négligeant ε),

$$\frac{T_{\lambda+1}}{T_0} < \prod_{k=1}^{\lambda} \frac{1}{1 + \frac{k}{(N+1)pq}} \quad (8)$$

et, ce qui est la même chose

$$\frac{T_{\lambda+1}}{T_0} < \prod_{k=1}^{\lambda} \frac{1}{1 + \frac{\lambda+1-k}{(N+1)pq}} \quad (9)$$

Multiplions ces deux inégalités membre à membre facteur par facteur et extrayons la racine carrée. Comme on a¹⁴

$$\left(1 + \frac{k}{(N+1)pq}\right) \left(1 + \frac{\lambda+1-k}{(N+1)pq}\right) > 1 + \frac{\lambda+1}{(N+1)pq}, \quad (10)$$

14. Le produit en (10) donne 4 facteurs : 3 d'entre eux deviennent le membre de droite, le quatrième est positif, d'où l'inégalité (10). Par ailleurs, à partir de (8) et (9), on a remplacé par k l'entier générique du produit utilisé par l'auteur, x .

il vient a fortiori (T_0 étant < 1)

$$T_{\lambda+1} < T_0 \left(1 + \frac{\lambda+1}{(N+1)pq} \right)^{\frac{\lambda}{2}} < \left(1 + \frac{\lambda+1}{(N+1)pq} \right)^{\frac{\lambda}{2}} \quad (11)$$

Donnons une explication à ce stade : (8) et (9) sont des produits constitués chacun de λ termes. En les multipliant, on arrive à $(T_{\lambda+1}/T_0)^2$ d'un côté, et λ fois l'inégalité (10) de l'autre, et en prenant la racine carrée, on obtient (11). Et La Vallée Poussin poursuit :

Cette inégalité est symétrique en p et q. On en conclut que tout terme à une distance supérieure à λ du terme principal [NB : T_0] est moindre que

$$\left(1 + \frac{\lambda+1}{(N+1)pq} \right)^{\frac{\lambda}{2}} \quad (12)$$

Donc la somme de tous ces termes (dont le nombre est $< N$) est moindre que

$$\frac{N}{\left(1 + \frac{\lambda}{(N+1)pq} \right)^{\frac{\lambda}{2}}} \quad (13)$$

Soit t un nombre fixe aussi petit qu'on veut. Prenons λ supérieur à $t(N+1)$, la somme précédente sera plus petite que

$$\frac{N}{\left(1 + \frac{t}{pq} \right)^{\frac{t}{2}N+1}} \quad (14)$$

Quand N croît indéfiniment le dénominateur croît comme une exponentielle et est infiniment grand par rapport à N, le quotient tend vers 0. La somme de tous les termes de $(p + q)^N$ étant un, on en conclut que la somme des termes à une distance moindre que $t(N+1)$ du terme principal a pour limite l'unité, quand N tend vers l'infini, quel que petit soit le nombre fixe t : c'est le théorème de Jacques Bernoulli.

À partir de (12), la démonstration de La Vallée Poussin ne pose pas de difficulté. Il s'agit d'isoler, à partir d'un certain rang, les termes de la somme initiale, et en majorant leur somme partielle, de montrer que celle-ci tend vers 0 quand N croît (loi des grands nombres).

CONCLUSION

Charles de La Vallée Poussin a donné une démonstration simple car purement algébrique du théorème de Bernoulli, sans faire appel ni à la formule de Stirling, ni à l'inégalité de Bienaymé-Tchebychev. En dehors de l'élégance de la preuve, typique des travaux de La Vallée Poussin, il faut noter l'esprit de collaboration du mathématicien. Il communique cette preuve à Robert de Montessus et l'autorise à l'utiliser à sa guise. Elle trouvera toute sa place dans le livre de Robert de Montessus de 1908 (*Leçons élémentaires sur le calcul des probabilités*).



(février 2014)

RÉFÉRENCES

- [AMS-LMS, 2000] “Kolmogorov in perspective”, *History of Mathematics*, volume 20, American Mathematical Society, London Mathematical Society, 2000.
- [Bernoulli J., 1713] *Ars conjectandi: Opus posthumum: accedit tractatus de seriebus infinitis, et Epistola Gallicè Scripta de ludo pilae reticularis*. Published 1713 by Impensis Thurnisiorum, fratrum in Basileae. (Latin.)
- [Bernoulli J., 1899] *Wahrscheinlichkeitsrechnung, Ars conjectandi, 1713*, Übersetzt und hrsg. von R. Haussner. Published 1899 by W. Engelmann in Leipzig . Written in German.
- [Bernoulli J., 2005] *On the law of large numbers. Translation of Pars Quarta tradens Usus & Applicationem Praecedentis Doctrinae in Civilibus, Moralibus & Oeconomicis*, Translated to English by Oscar Sheynin, Berlin 2005. <http://www.sheynin.de/download/bernoulli.pdf>.
- [Borel E., 1909] *Les probabilités dénombrables et leurs applications arithmétiques*. Rendiconti del Circolo Matematico di Palermo, Volume 27, Issue 1, pp 247-271, 1909.
- [Borel E., 1958] *Probabilités, Erreurs*, 10^e édition, Librairie Armand Colin, 1958.
- [Deheuvels P., 1982] *La probabilité, le hasard et la certitude*, Que sais-je ? PUF, 1982.
- [Denker M., 2013] “Tercentennial anniversary of Bernoulli’s of large numbers”, *Bulletin (New series) of the American Mathematical Society*, volume 50, number 3, July 2013, pp 373-390 (article electronically published on March 28, 2013).
- [Fischer H., 2011] *A History of the Central Limit Theorem, Sources and Studies in the History of Mathematics and Physical Sciences*, Springer Science -Business Media, LLC 2011
- [Goedseels E., 1895] « Démonstration du théorème de Jacques Bernoulli, en calcul des probabilités », *Annales de la Société Scientifique de Bruxelles*, XIX, 1895.
- [Henry M., 2004] « La démonstration par Jacques Bernoulli de son théorème », *Histoires de Probabilités et de Statistiques*, IREM - Histoire des Mathématiques coordonné par Evelyne Barbin et Jean-Pierre Lamarche, Ellipses, 2004.
- [Issac R., 2005] *Une initiation aux probabilités*, traduction de Roger Mansuy, Vuibert-Springer, 2005.
- [Jacod J., Protter P., 2003] *L’essentiel en théorie des probabilités*, Cassini, Paris, 2003 (pour l’édition française)
- [Jordan Camille, 1894] *Cours d’Analyse de l’Ecole Polytechnique, tome deuxième : calcul intégral*, Gauthier-Villars, Paris, 1894 (ouvrage numérisé.)
- [Kolmogorov A.N., 1933] *Grundbegriffe der Wahrscheinlichkeitrechnung*, Springer 1933.
- [Kolmogorov A.N., 1956] *Foundations of the theory of probability*, second English edition, Chelsea publishing company, New-York, 1958.
- [Lanier D., Trotoux D., 1995] « La loi des grands nombres : le théorème de De Moivre-Laplace », in *Contribution à une approche historique de l’enseignement des mathématiques*, IREM de Besançon, Presses Universitaires de Franche-Comté, 1995.
- [La Vallée Poussin (de) C.J., 1907] « Etude sur le théorème de Bernoulli », *Annales de la Société Scientifique de Bruxelles* (2e série) 31 (1907), 73-77.
- [La Vallée Poussin (de) C.J., 1907-2] « Démonstration nouvelle du théorème de Bernoulli », *Annales de la Société Scientifique de Bruxelles* (2e série) 31 (1907), 219-236.

- [La Vallée Poussin (de) C.J., 2001] *Collected works, oeuvres scientifiques, Volume II : Integration and measure, probability, ordinary differential equations, mechanics, geometry (Intégration, mesure, probabilités, équations différentielles ordinaires géométrie, mécanique,* Edited by (Editée par) Paul Butzer, Jean Mawhin, Pasquale Vetro. Académie Royale de Belgique, Circolo matematico di Palermo, 2001.
- [Lesigne E., 2005] *Heads or Tails: An Introduction to Limit Theorems in Probability,* American Mathematical Society, 2005.
- [Mansion P., 1902] « Démonstration du théorème de Jacques Bernouilli », *Brux. S. sc.* 27, 1902.
- [Mansion P., 1904] « Sur la loi des grands nombres de Poisson », *Brux. S. sc.* 28 A, 72-77, 1904.
- [Mawhin J., 2011] *Une brève histoire des mathématiques à l'Université catholique de Louvain,* en ligne sur <http://www.sciences.be>, 2011.
- [Mises (von) R., 1919] *Grundlagen der Wahrscheinlichkeitsrechnung,* *Math. Zeitsch.*, Vol. 5, 1919, pp. 55-99.
- [Montessus (de) Robert, 1908] *Leçons élémentaires sur le calcul des probabilités,* Gauthier-Villars, Paris, 1908.
- [Poisson S.D., 1837] *Recherches sur la probabilité des jugements en matière criminelle et en matière civile,* Bachelier, Paris, 1837.
- [Seneta E., 2013] "A tricentenary history of the law of large numbers", *Bernoulli* 19 (4), 2013, 1088-1121.
- [Voelke J.-D., 2005] *Renaissance de la géométrie non euclidienne entre 1860 et 1900,* Peter Lang, 2005.

