Approche de Bernoulli.

Nota : Le texte qui suit est un extrait du document "Probabilités" de Eduscol.
Voir le document complet :
http://media.eduscol.education.fr/file/Mathematiques/24/3/Probablites_17_03_08_maj2011_197243.pdf

Du point de vue historique, le résultat fondamental permettant de comprendre la façon dont se stabilisent les fréquences des résultats au fur et à mesure que le nombre des essais augmente est le théorème de Bernoulli. Ce théorème, qui confirme l'intuition sur la stabilité des fréquences, est plus difficile à démontrer que la formule de Bayes. La démonstration "moderne" dans la théorie axiomatique de Kolmogorov repose sur une majoration assez grossière donnée par la formule de Bienaymé - Tchebychev, majoration peu utile en pratique. La démonstration donnée par Bernoulli est beaucoup plus compliquée, mais présente l'intérêt de présenter très clairement sa problématique : il remarque que, dans les jeux de dés ou de tirages dans une urne, la détermination des probabilités a priori ne pose pas de problème : il suffit de prendre le ratio entre le nombre de tirages "fertiles" et le nombre total de tirages ou le ratio entre le nombre de tirages "fertiles" et le nombre de tirages "stériles". Mais, il constate que cette méthode est inutilisable dans des problèmes concernant les maladies, la météorologie, où les causes sont cachées, et où l'énumération des cas équiprobables est impossible. Au lieu de cela, il propose de déterminer la probabilité d'un cas favorable a posteriori :
« On peut supposer qu'une chose particulière se produira ou non autant de fois qu'elle s'est produite ou non dans le passé, dans des circonstances semblables ».

Il cherche donc à déterminer empiriquement la proportion de cas favorables dans le cas où elle est inconnue. L'originalité de la tentative de Bernoulli consiste à donner un traitement formel de la vague notion qu'il décrit ainsi :
« Même le plus stupide des hommes, par quelque instinct de la nature, par lui-même et sans aucune instruction (et c'est une chose remarquable), est convaincu que plus on fait d'observations, moins on risque de s'écarter de notre but ».

Bernoulli veut démontrer ce principe, et montrer que la "certitude morale" à propos de la proportion inconnue peut être approchée d'aussi près que l'on veut. Il considère une urne contenant 3000 galets blancs et 2000 galets noirs. On tire un galet, avec remise. On regarde combien de fois on tire un galet blanc, combien de fois on tire un galet noir. Se pose alors la question : Peut-on tirer un nombre suffisant de fois de manière à ce qu'il devienne 10 fois, 100 fois, 1000 fois plus probable que les nombres de galets blancs et noirs tirés soient dans le ratio 3:2 plutôt que dans tout autre ratio ? Bernoulli précise :
« Pour éviter les malentendus, on doit noter que le ratio que nous essayons de déterminer expérimentalement ne doit pas être considéré comme précis et indivisible (sinon, c'est le contraire qui se produirait, et il deviendrait moins probable que le vrai ratio soit trouvé en augmentant les observations). Ce que l'on veut, en revanche, c'est un ratio pris avec quelque latitude, c'est-à-dire situé entre deux limites qui peuvent être aussi proches l'une de l'autre que l'on veut. Par exemple, on prend deux ratios 301:200 et 299:200 ou 3001:2000 et 2999:2000 … l'un qui est immédiatement supérieur et l'autre immédiatement inférieur au ratio 3:2. On prouvera que l'on peut rendre plus probable que le ratio trouvé après des expériences répétées tombe entre ces limites plutôt qu'il tombe à l'extérieur. »

Après avoir fait la démonstration[21], il traite l'exemple où $r = 30$, $s = 20$, $p = 3/5$ et $\varepsilon = 1/50$. Il trouve $n = 25\,550$ pour $c = 1000$,
25 550 est à l'époque un nombre astronomique, inutilisable dans la pratique (Il est pourtant bien meilleur que celui obtenu en employant l'inégalité de Bienaymé Tchebychev, qui est égal à 600 600) : ceci conduit Bernoulli à ne pas publier ses travaux. Remarquons que Bernoulli ne répond pas à la question qu'il s'est posée au départ, car la proportion est ici connue au départ. Il détermine le nombre de tirages suffisant pour que, avec une probabilité très forte (supérieure à 0,999, alors qu'aujourd'hui on se contente selon les secteurs d'activité des niveaux 0,95 ou 0,99), la proportion de cas favorables ne s'écarte pas de 3/5 de plus de 1/50. En d'autres termes, il détermine n pour un intervalle de probabilité au niveau 1000/1001 d'amplitude donnée, alors qu'il cherchait n pour en faire un intervalle de confiance au même niveau.

Bayes s'est posé et a résolu un problème voisin de celui de Bernoulli, qu'il formule ainsi :
« Etant donné le nombre de fois qu'un événement inconnu s'est réalisé ou non, on cherche la chance que la probabilité de sa réalisation lors d'une seule épreuve soit comprise entre deux degrés quelconques de probabilité que l'on puisse assigner. »
problème que l'on peut traduire avec un langage plus moderne de la manière suivante :
Un événement se produit à chaque tirage avec la probabilité θ. Soit X le nombre de fois qu'il se produit au cours de n essais. On demande $P(a < \theta < b \mid X)$, probabilité que θ soit comprise entre a et b, connaissant le nombre de fois où l'événement s'est produit au cours des n essais.
Il a illustré sa résolution avec un dispositif original et intéressant, connu sous le nom de "billard de Bayes". Mais ses travaux n'ont connu aucune diffusion, en raison des difficultés de calcul des intégrales mises en jeu dans la solution.

Les travaux de De Moivre sur le développement du binôme, de De Moivre et de Laplace sur le théorème qui porte leurs noms (et qui est un cas particulier du théorème "central limite" établi par Laplace) et ceux de Gauss sur la loi de Laplace - Gauss vont permettre d'obtenir les résultats essentiels et de prolonger le travail de Bernoulli (et de Bayes).
On peut modéliser n tirages au hasard et avec remise dans une urne contenant des boules noires et des boules blanches, la proportion de boules noires dans l'urne étant égale à p, en introduisant n variables aléatoires indépendantes X_i, prenant la valeur 1 si la $i$ème boule tirée est noire, et 0 si elle est blanche. La somme S_n des X_i, i variant de 1 à n, donne le nombre de boules noires obtenu à l'issue des n tirages. S_n suit la loi binomiale de paramètres n et p, son espérance mathématique est égale à np et son écart type est égal à $np(1-p)$. On s'intéresse à la fréquence F_n du nombre de boules noires à l'issue des n tirages, égale à S_n/n.
Le théorème de Moivre - Laplace dit que la variable aléatoire centrée réduite R_n associée à S_n, égale à $(S_n - np)/sqrt(np(1-p))$, converge en loi vers la loi normale centrée réduite, c'est-à-dire que lorsque n tend vers l'infini, la probabilité pour que R_n prenne des valeurs comprises entre a et b tend vers $\frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$.
La variable centrée réduite associée à F_n est $R_n = \frac{F_n - p}{\sqrt{\frac{p(1-p)}{n}}}$.
Plus précisément, sous les hypothèses suivantes : $n > 30$, $np > 5$ et $n(1 - p) > 5$, on approxime avec une très bonne précision la probabilité pour que R_n soit dans l'intervalle $[a, b]$ par sa limite donnée par le théorème de Moivre - Laplace.

À l'aide d'une table de la loi normale ou d'un tableur, on peut trouver une valeur approchée du nombre réel u en prenant a=-u et b-u tel que $\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx$, intégrale notée $\varphi(u)$, prenne une valeur donnée.
Il est utile de retenir que $\varphi(1{,}96) \approx 0{,}95$; $\varphi(1{,}65) \approx 0{,}90$; $\varphi(3) \approx 0{,}99$. Par exemple, sous les conditions rappelées plus haut : $P(-1{,}96 \leq R_n \leq 1{,}96) \approx 0{,}95$.

Donc avec une probabilité voisine de 95%, l'inégalité précédente est vraie. L'intervalle ainsi défini est appelé intervalle de probabilité (ou de pari) de niveau 95%, ou encore intervalle de fluctuation de niveau 95%. Son interprétation est la suivante : dans 95% des séries de n tirages que l'on peut faire, la fréquence empirique f_n obtenue expérimentalement (modélisée par F_n) appartient à cet intervalle.