

RÉFLEXIONS SUR LE CALCUL DES PROBABILITÉS

Le nom seul de calcul des probabilités est un paradoxe : la probabilité, opposée à la certitude, c'est ce qu'on ne sait pas, et comment peut-on calculer ce que l'on ne connaît pas? Cependant, beaucoup de savants éminents se sont occupés de ce calcul, et l'on ne saurait nier que la science n'en ait tiré quelque profit. Comment expliquer cette apparente contradiction?

La probabilité a-t-elle été définie? Peut-elle même être définie? Et, si elle ne peut l'être, comment ose-t-on en raisonner? La définition, dira-t-on, est bien simple : la probabilité d'un événement est le rapport du nombre des cas favorables à cet événement au nombre total des cas possibles.

Un exemple simple va faire comprendre combien cette définition est incomplète. Je jette deux dés; quelle est la probabilité pour que l'un des deux dés au moins amène un six? Chaque dé peut amener six points différents : le nombre des cas possibles est $6 \times 6 = 36$; le nombre des cas favorables est 11; la probabilité est $\frac{11}{36}$.

C'est là la solution correcte. Mais ne pourrais-je pas dire tout aussi bien : Les points amenés par les deux dés peuvent former $\frac{6 \times 7}{2} = 21$ combinaisons différentes? Parmi ces combinaisons, 6 sont favorables; la probabilité est $\frac{6}{21}$.

Pourquoi la première manière d'énumérer les cas possibles est-elle plus légitime que la seconde? En tous cas, ce n'est pas notre définition qui nous l'apprend.

On est donc réduit à compléter cette définition en disant : «... au nombre total des cas possibles, pourvu que ces cas soient également probables. » Nous voilà donc réduits à définir le probable par le probable.

Comment saurons-nous que deux cas possibles sont également probables? Sera-ce par une convention? Si nous plaçons au début de chaque problème une convention explicite, tout ira bien; nous n'aurons plus qu'à appliquer les règles de l'arithmétique et de l'algèbre et nous irons jusqu'au bout du calcul sans que notre résultat puisse laisser place au doute. Mais, si nous voulons en faire la moindre application, il faudra démontrer que notre convention était légitime, et nous nous retrouverons en face de la difficulté que nous avons cru éluder.

Dira-t-on que le bon sens suffit pour nous apprendre quelle convention il faut faire? Hélas! M. Bertrand s'est amusé à traiter un problème

simple : « Quelle est la probabilité pour que, dans une circonférence, une corde soit plus grande que le côté du triangle équilatéral inscrit? » L'illustre géomètre a adopté successivement deux conventions que le bon sens semblait également imposer, et il a trouvé avec l'une $\frac{1}{2}$, avec l'autre $\frac{1}{3}$.

La conclusion qui semble résulter de tout cela, c'est que le calcul des probabilités est une science vaine, qu'il faut se défier de cet instinct obscur que nous nommons bon sens et auquel nous demandons de légitimer nos conventions.

Mais, cette conclusion, nous ne pouvons non plus y souscrire; cet instinct obscur, nous ne pouvons nous en passer; sans lui la science serait impossible, sans lui nous ne pourrions ni découvrir une loi, ni l'appliquer. Avons-nous le droit, par exemple, d'énoncer la loi de Newton? Sans doute, de nombreuses observations sont en concordance avec elle; mais n'est-ce pas là un simple effet du hasard? comment savons-nous d'ailleurs si cette loi, vraie depuis tant de siècles, le sera encore l'an prochain? A cette objection, vous ne trouverez rien à répondre, sinon : « Cela est bien peu probable ».

Mais admettons la loi; grâce à elle, je crois pouvoir calculer la position de Jupiter dans un an. En ai-je le droit? qui me dit qu'une masse gigantesque, animée d'une vitesse énorme, ne va pas d'ici-là passer près du système solaire et produire des perturbations imprévues? Ici encore il n'y a rien à répondre, sinon : « Cela est bien peu probable ».

A ce compte, toutes les sciences ne seraient que des applications inconscientes du calcul des probabilités; condamner ce calcul, ce serait condamner la science tout entière.

J'insisterai moins sur les problèmes scientifiques où l'intervention du calcul des probabilités est plus évidente. Tel est en première ligne celui de l'interpolation, où, connaissant un certain nombre de valeurs d'une fonction, on cherche à deviner les valeurs intermédiaires.

Je citerai également : la célèbre théorie des erreurs d'observation, sur laquelle je reviendrai plus loin; la théorie cinétique des gaz, hypothèse bien connue, où chaque molécule gazeuse est supposée décrire une trajectoire extrêmement compliquée, mais où, par l'effet des grands nombres, les phénomènes moyens, seuls observables, obéissent à des lois simples qui sont celles de Mariotte et de Gay-Lussac.

Toutes ces théories reposent sur les lois des grands nombres, et le calcul des probabilités les

entraînerait évidemment dans sa ruine. Il est vrai qu'elles n'ont qu'un intérêt particulier et que, sauf en ce qui concerne l'interpolation, ce sont là des sacrifices auxquels on pourrait se résigner.

Mais, je l'ai dit plus haut, ce ne serait pas seulement de ces sacrifices partiels qu'il s'agirait, ce serait la science tout entière dont la légitimité serait révoquée en doute.

Je vois bien ce qu'on pourrait dire : « Nous sommes ignorants et pourtant nous devons agir. Pour agir, nous n'avons pas le temps de nous livrer à une enquête suffisante pour dissiper notre ignorance ; d'ailleurs, une pareille enquête exigerait un temps infini. Nous devons donc nous décider sans savoir ; il faut bien le faire au petit bonheur et suivre des règles sans trop y croire. Ce que je sais, ce n'est pas que telle chose est vraie, mais que le mieux pour moi est encore d'agir comme si elle était vraie ». Le calcul des probabilités, et par conséquent la science, n'aurait plus qu'une valeur pratique.

Malheureusement la difficulté ne disparaît pas ainsi : Un joueur veut tenter un coup ; il me demande conseil. Si je le lui donne, je m'inspirerai du calcul des probabilités, mais je ne lui garantirai pas le succès. C'est là ce que j'appellerai la *probabilité subjective*. Dans ce cas, on pourrait se contenter de l'explication que je viens d'esquisser. Mais je suppose qu'un observateur assiste au jeu, qu'il en note tous les coups et que le jeu se prolonge longtemps ; quand il fera le relevé de son carnet, il constatera que les événements se sont répartis conformément aux lois du calcul des probabilités. C'est là ce que j'appellerai la *probabilité objective*, et c'est ce phénomène qu'il faudrait expliquer.

Il existe de nombreuses sociétés d'assurances qui appliquent les règles du calcul des probabilités, et elles distribuent à leurs actionnaires des dividendes dont la réalité objective ne saurait être contestée. Il ne suffit pas, pour les expliquer, d'invoquer notre ignorance et la nécessité d'agir.

Ainsi, le scepticisme absolu n'est pas de mise ; nous devons nous méfier, mais nous ne pouvons condamner en bloc ; il est nécessaire de discuter.

I. — CLASSIFICATION DES PROBLÈMES DE PROBABILITÉ.

Pour classer les problèmes qui se présentent à propos des probabilités, on peut se placer à plusieurs points de vue différents, et d'abord au *point de vue de la généralité*. J'ai dit plus haut que la probabilité est le rapport du nombre des cas favorables au nombre des cas possibles. Ce que, faute d'un meilleur terme, j'appelle la généralité, croîtra avec le nombre des cas possibles. Ce nombre peut être fini ; comme, par exemple, si l'on envisage un coup de dés où le nombre des cas possibles

est 36. C'est là le premier degré de généralité.

Mais, si nous demandons, par exemple, quelle est la probabilité pour qu'un point intérieur à un cercle soit intérieur au carré inscrit, il y a autant de cas possibles que de points dans le cercle, c'est-à-dire une infinité. C'est le second degré de généralité. La généralité peut être poussée plus loin encore : on peut se demander la probabilité pour qu'une fonction satisfasse à une condition donnée ; il y a alors autant de cas possibles qu'on peut imaginer de fonctions différentes. C'est le troisième degré de généralité, auquel on s'élève, par exemple, quand on cherche à deviner la loi la plus probable d'après un nombre fini d'observations.

On peut se placer à un point de vue tout différent. Si nous n'étions pas ignorants, il n'y aurait pas de probabilité, il n'y aurait de place que pour la certitude ; mais notre ignorance ne peut être absolue, sans quoi il n'y aurait pas non plus de probabilité, puisqu'il faut encore un peu de lumière pour parvenir même à cette science incertaine. Les problèmes de probabilité peuvent ainsi se classer d'après la profondeur plus ou moins grande de cette ignorance.

En Mathématiques, on peut déjà se proposer des problèmes de probabilité. Quelle est la probabilité pour que la 5^e décimale d'un logarithme pris au hasard dans une table soit un 9 ? On n'hésitera pas à répondre que cette probabilité est $\frac{1}{10}$. Ici nous

possédons toutes les données du problème ; nous saurions calculer notre logarithme sans recourir à la table ; mais nous ne voulons pas nous en donner la peine. C'est le premier degré de l'ignorance.

Dans les sciences physiques, notre ignorance est déjà plus grande. L'état d'un système, à un instant donné, dépend de deux choses : son état initial et la loi d'après laquelle cet état varie. Si nous connaissions à la fois cette loi et cet état initial, nous n'aurions plus qu'un problème mathématique à résoudre et nous retomberions sur le premier degré d'ignorance.

Mais il arrive souvent qu'on connaît la loi et qu'on ne connaît pas l'état initial. On demande, par exemple, quelle est la distribution actuelle des petites planètes ; nous savons que, de tout temps, elles ont obéi aux lois de Képler, mais nous ignorons quelle était leur distribution initiale.

Dans la théorie cinétique des gaz, on suppose que les molécules gazeuses suivent des trajectoires rectilignes et obéissent aux lois du choc des corps élastiques ; mais, comme on ne sait rien de leurs vitesses initiales, on ne sait rien de leurs vitesses actuelles.

Seul, le calcul des probabilités permet de prévoir les phénomènes moyens qui résulteront de

la combinaison de ces vitesses. C'est là le second degré d'ignorance.

Il est possible, enfin, que non seulement les conditions initiales, mais les lois elles-mêmes, soient inconnues; on atteint alors le troisième degré de l'ignorance et, généralement, on ne peut plus rien affirmer du tout au sujet de la probabilité d'un phénomène.

Il arrive souvent qu'au lieu de chercher à deviner un événement d'après une connaissance plus ou moins imparfaite de la loi, on connaisse les événements et qu'on cherche à deviner la loi; qu'au lieu de déduire les effets des causes, on veuille déduire les causes des effets. Ce sont là les problèmes dits de *probabilité des causes*, les plus intéressants au point de vue de leurs applications scientifiques.

Je joue à l'écarté avec un monsieur que je sais parfaitement honnête; il va donner; quelle est la probabilité pour qu'il tourne le roi? c'est $\frac{1}{8}$; c'est là un problème de probabilité des effets. Je joue avec un monsieur que je ne connais pas; il a donné 10 fois et il a tourné 6 fois le roi; quelle est la probabilité pour que ce soit un grec? c'est là un problème de probabilité des causes.

On peut dire que c'est le problème essentiel de la méthode expérimentale. J'ai observé n valeurs de x et les valeurs correspondantes de y ; j'ai constaté que le rapport des secondes aux premières est sensiblement constant. Voilà l'événement; quelle est la cause?

Est-il probable qu'il y ait une loi générale d'après laquelle y serait proportionnel à x et que les petites divergences soient dues à des erreurs d'observations? Voilà un genre de question qu'on est sans cesse amené à se poser et qu'on résout inconsciemment toutes les fois que l'on fait de la science.

Je vais maintenant passer en revue ces différentes catégories de problèmes en envisageant successivement ce que j'ai appelé plus haut la probabilité subjective et ce que j'ai appelé la probabilité objective.

II. — LA PROBABILITÉ DANS LES SCIENCES MATHÉMATIQUES.

L'impossibilité de la quadrature du cercle est démontrée depuis 1885; mais, bien avant cette date récente, tous les géomètres considéraient cette impossibilité comme tellement « probable », que l'Académie des Sciences rejetait sans examen les mémoires, hélas! trop nombreux, que quelques malheureux fous lui envoyaient tous les ans sur ce sujet.

L'Académie avait-elle tort? Evidemment non, et elle savait bien qu'en agissant ainsi, elle ne ris-

quait nullement d'étouffer une découverte sérieuse. Elle n'aurait pu démontrer qu'elle avait raison; mais elle savait bien que son instinct ne la trompait pas. Si vous aviez interrogé les académiciens, ils vous auraient répondu : « Nous avons comparé la probabilité pour qu'un savant inconnu ait trouvé ce qu'on cherche vainement depuis si longtemps, et celle pour qu'il y ait un fou de plus sur la terre; la seconde nous a paru plus grande. » Ce sont là de très bonnes raisons, mais elles n'ont rien de mathématique, elles sont purement psychologiques.

Et si vous les aviez pressés davantage, ils auraient ajouté : « Pourquoi voulez-vous qu'une valeur particulière d'une fonction transcendante soit un nombre algébrique; et si π était racine d'une équation algébrique, pourquoi voulez-vous que cette racine soit une période de la fonction $\sin 2x$ et qu'il n'en soit pas de même des autres racines de cette même équation? » En somme, ils auraient invoqué le principe de raison suffisante sous sa forme la plus vague.

Mais que pouvaient-ils en tirer? Tout au plus une règle de conduite pour l'emploi de leur temps, plus utilement dépensé à leurs travaux ordinaires qu'à la lecture d'une élucubration qui leur inspirait une légitime défiance. Mais ce que j'appelais plus haut la probabilité objective n'a rien à voir avec ce premier problème.

Il en est autrement du second problème.

Envisageons les nombres :

$$\log \left(1 + \frac{x}{100.000} \right),$$

où je donne successivement à x les valeurs 1, 2, ... jusqu'à 10.000. Parmi ces 10.000 nombres, j'en prends un au hasard; quelle est la probabilité pour que sa troisième décimale soit un nombre pair?

Vous n'hésitez pas à répondre $\frac{1}{2}$, et, en effet, si vous relevez dans une table les troisièmes décimales de ces 10.000 nombres, vous trouverez à peu près autant de chiffres pairs que de chiffres impairs.

Il sera plus simple d'examiner un problème un peu différent, mais analogue. Ce que je dirai de l'un s'appliquerait également à l'autre. Considérons les 10.000 expressions :

$$\sin \left[1.000 \pi \log \left(1 + \frac{x}{100.000} \right) \right].$$

Je n'hésiterai pas à dire que la moyenne de ces 10.000 expressions est probablement nulle, et, si je la calculais effectivement, je vérifierais qu'elle est très petite¹.

¹ L'analogie des deux problèmes est évidente; celui que

Au lieu de faire la somme et de diviser par 10.000, je calcule l'intégrale de notre expression depuis $\frac{1}{2}$ jusqu'à $10.000 + \frac{1}{2}$ et je divise par 10.000. Je démontrerais aisément que l'erreur ainsi commise est plus petite que :

$$\frac{\pi^2}{10.000} + \frac{\pi}{10.000.000} < \frac{1}{1.000},$$

et ensuite, en intégrant par parties, que l'intégrale divisée par 10.000 est elle-même plus petite que :

$$\frac{22}{1.000 \pi \log e} + \frac{1,1}{1.000 \pi \log^2 e} < \frac{1}{500}.$$

Pour établir ce double résultat, je n'ai besoin que de m'appuyer sur deux faits, à savoir que les dérivées première et seconde du logarithme restent, dans l'intervalle considéré, comprises entre certaines limites.

D'où cette première conséquence que la propriété est vraie non seulement du logarithme, mais d'une fonction continue quelconque, puisque les dérivées de toute fonction continue sont limitées.

Si j'étais certain d'avance du résultat, c'est d'abord que j'avais souvent observé des faits analogues pour d'autres fonctions continues; c'est ensuite parce que je faisais dans mon for intérieur, d'une façon plus ou moins inconsciente et imparfaite, le raisonnement qui m'a conduit aux inégalités précédentes, comme un calculateur qui, avant d'avoir achevé une multiplication, se rend compte que « cela va faire à peu près tant ».

Et d'ailleurs, comme ce que j'appelais mon intuition n'était qu'un aperçu incomplet d'un véritable raisonnement, on s'explique que l'observation ait confirmé mes prévisions, que la probabilité objective ait été d'accord avec la probabilité subjective.

Comme troisième exemple, je choisirai le problème suivant : Un nombre u est pris au hasard, n est un entier donné très grand; quelle est la valeur probable de $\sin nu$? Ce problème n'a aucun sens par lui-même. Pour lui en donner un, il faut une convention; nous *conviendrons* que la probabilité pour que le nombre u soit compris entre les limites a et b est :

$$\int_a^b \varphi(u) du,$$

$\varphi(u)$ étant une fonction que je choisis arbitraire-

j'avais d'abord posé se ramène à la recherche de la valeur moyenne de l'expression :

$$F \left[\log \left(1 + \frac{x}{100.000} \right) \right],$$

$F(y)$ étant une fonction qui est égale à $+1$, si la troisième décimale de y est paire, et à -1 si elle est impaire. Or, cette fonction, comme $\sin 1.000 \pi y$, est une fonction périodique dont la période est $\frac{1}{500}$.

ment, mais que je suppose continue. La valeur de $\sin nu$ restant la même quand u augmente de 2π , je puis, sans restreindre la généralité, supposer que u est compris entre 0 et 2π et je serai ainsi conduit à supposer que $\varphi(u)$ est une fonction périodique dont la période est 2π . Comme je suis *certain* que u est compris entre 0 et 2π , comme, en d'autres termes, la probabilité de cet événement est égale à 1, j'aurai :

$$\int_0^{2\pi} \varphi(u) du = 1.$$

La valeur probable cherchée est :

$$\int_0^{2\pi} \varphi(u) \sin nu du,$$

et il est aisé de montrer que cette intégrale est plus petite que :

$$\frac{2\pi M_k}{n^k},$$

M_k étant la plus grande valeur de la dérivée k^e de $\varphi(u)$. On voit donc que, si la dérivée k^e est finie, notre valeur probable tendra vers zéro quand n croîtra indéfiniment et cela plus vite que $\frac{1}{n^{k-1}}$.

La valeur probable de $\sin nu$ pour n très grand est donc nulle; pour définir cette valeur, j'ai eu besoin d'une convention; mais le résultat reste le même *quelle que soit cette convention*. Je ne me suis imposé que de faibles restrictions en supposant que la fonction $\varphi(u)$ est continue et périodique, et ces hypothèses sont tellement naturelles qu'on se demande comment on pourrait y échapper.

L'examen des trois exemples précédents, si différents à tous égards, nous a fait déjà entrevoir d'une part le rôle de ce que les philosophes appellent le principe de raison suffisante, et d'autre part l'importance de ce fait que certaines propriétés sont communes à toutes les fonctions continues. L'étude de la probabilité dans les sciences physiques nous conduira au même résultat.

III. — LA PROBABILITÉ DANS LES SCIENCES PHYSIQUES.

Arrivons maintenant aux problèmes qui se rapportent à ce que j'ai appelé plus haut le second degré d'ignorance; ce sont ceux où l'on connaît la loi, mais où on ignore l'état initial du système. Je pourrais multiplier les exemples, je n'en prendrai qu'un : Quelle est la distribution actuelle probable des petites planètes sur le zodiaque ?

Nous savons qu'elles obéissent aux lois de Képler; nous pouvons même, sans rien changer à la nature du problème, supposer que leurs orbites sont toutes circulaires et situées dans un même plan et que nous le sachions. En revanche, nous ignorons absolument quelle était leur distribution initiale. Cepen-

dant nous n'hésitons pas à affirmer qu'aujourd'hui cette distribution est à peu près uniforme. Pourquoi ?

Soit b la longitude d'une petite planète à l'époque initiale, c'est-à-dire à l'époque 0 ; soit a son moyen mouvement ; sa longitude à l'époque actuelle, c'est-à-dire à l'époque t , sera $at + b$. Dire que la distribution actuelle est uniforme, c'est dire que la valeur moyenne des sinus et des cosinus des multiples de $at + b$ est nulle. Pourquoi l'affirmons-nous ?

Soit $\varphi(a, b) da db$ la probabilité pour que le moyen mouvement d'une petite planète soit compris entre a et $a + da$ et que sa longitude initiale soit comprise entre b et $b + db$. La valeur moyenne de $\sin(at + b)$ sera donnée par l'intégrale double :

$$\iint \varphi(a, b) \sin(at + b) da db.$$

Ne sachant rien de la distribution initiale, je ne sais rien de la fonction $\varphi(a, b)$; je n'ai pas plus de raison pour choisir telle fonction plutôt que telle autre ; je serai cependant conduit à choisir une fonction continue. Si je suppose, par conséquent, que la fonction φ a des dérivées, notre intégrale double deviendra, en intégrant successivement par parties :

$$\frac{1}{t} \iint \frac{d\varphi}{da} \cos(at + b) da db, - \frac{1}{t^2} \iint \frac{d^2\varphi}{da^2} \sin(at + b) da db, \dots$$

La présence des facteurs $\frac{1}{t}$, $\frac{1}{t^2}$ nous montre qu'elle tend vers zéro quand t augmente.

Ainsi, je ne savais trop quelle hypothèse faire au sujet de la probabilité de telle ou telle distribution initiale ; mais, quelle que soit l'hypothèse faite, le résultat sera le même et c'est ce qui me tire d'embarras.

Quelle que soit la fonction φ , la valeur moyenne tend vers zéro quand t augmente, et comme les petites planètes ont certainement accompli un très grand nombre de révolutions, je puis affirmer que cette valeur moyenne est très petite.

Je puis choisir φ comme je le veux, sauf une restriction toutefois : cette fonction doit être continue ; et, en effet, au point de vue de la probabilité subjective, le choix d'une fonction discontinue aurait été déraisonnable ; quelle raison pourrai-je avoir, par exemple, de supposer que la longitude initiale peut être égale à 0° juste, mais qu'elle ne peut être comprise entre 0° et 1° ?

Mais la difficulté reparaît si l'on se place au point de vue de la probabilité objective.

La valeur moyenne de $\sin(at + b)$ sera représentée tout simplement par :

$$\frac{1}{n} \sum \sin(at + b),$$

n étant le nombre des petites planètes. Au lieu d'une intégrale double portant sur une fonction continue, nous avons une somme de termes discrets. Et pourtant personne ne doutera sérieusement que cette valeur moyenne ne soit effectivement très petite.

C'est que notre somme discrète différera en général très peu d'une intégrale.

Une intégrale est la limite vers laquelle tend une somme de termes quand le nombre de ces termes croît indéfiniment. Si les termes sont très nombreux, la somme différera très peu de sa limite, c'est-à-dire de l'intégrale, et ce que j'ai dit de cette dernière sera encore vrai de la somme elle-même.

Il y a des cas d'exception néanmoins. Si, par exemple, l'on avait pour toutes les petites planètes :

$$b = \frac{\pi}{2} - at,$$

la valeur moyenne serait évidemment égale à 1. Pour cela, il faudrait qu'à l'époque 0, les petites planètes eussent été toutes placées sur une sorte de spirale d'une forme particulière à spires extrêmement serrées. Tout le monde jugera qu'une pareille distribution initiale est extrêmement improbable (et, même en la supposant réalisée, la distribution ne serait pas uniforme à l'époque actuelle, par exemple le 1^{er} janvier 1900, mais elle le redeviendrait quelques années plus tard).

Toutefois, pourquoi jugeons-nous cette distribution initiale improbable ? Il est nécessaire de l'expliquer, car, si nous n'avons pas de raison de rejeter comme invraisemblable cette hypothèse saugrenue, tout s'écroulerait et nous ne pourrions plus rien affirmer au sujet de la probabilité de telle ou telle distribution actuelle.

Ce que nous invoquerons, c'est encore le principe de raison suffisante, auquel il faut toujours revenir. Nous pourrions admettre qu'à l'origine les planètes étaient distribuées à peu près en ligne droite ; nous pourrions admettre qu'elles étaient irrégulièrement distribuées ; mais il nous semble qu'il n'y a pas de raison suffisante pour que la cause inconnue qui leur a donné naissance, ait agi suivant une courbe si régulière et pourtant si compliquée, et qui paraîtrait précisément avoir été choisie exprès pour que la distribution actuelle ne fût pas uniforme.

IV. — ROUGE ET NOIR.

Les questions soulevées par les jeux de hasard, comme celui de la roulette, sont, au fond, tout à fait analogues à celles que nous venons de traiter.

Par exemple, un cadran est partagé en un grand nombre de subdivisions égales, alternativement

rouges et noires; une aiguille est lancée avec force, et, après avoir fait un grand nombre de tours, elle s'arrête devant une de ces subdivisions. La probabilité, pour que cette division soit rouge, est évidemment $\frac{1}{2}$.

L'aiguille va tourner d'un angle θ , comprenant plusieurs circonférences; j'ignore quelle est la probabilité pour que l'aiguille soit lancée avec une force telle que cet angle soit compris entre θ et $\theta + d\theta$; mais, je puis faire une convention; je puis supposer que cette probabilité est $\varphi(\theta)d\theta$; quant à la fonction $\varphi(\theta)$, je puis la choisir d'une façon entièrement arbitraire; il n'y a rien qui puisse me guider dans mon choix; cependant, je suis naturellement conduit à supposer cette fonction continue.

Soit ε la longueur (comptée sur la circonférence de rayon 1) de chaque subdivision rouge ou noire. Il faut calculer l'intégrale $\int \varphi(\theta)d\theta$ en l'étendant, d'une part, à toutes les divisions rouges, d'autre part, à toutes les divisions noires, et comparer les résultats.

Considérons un intervalle 2ε , comprenant une division rouge et la division noire qui la suit. Soit M et m , la plus grande et la plus petite valeur de la fonction $\varphi(\theta)$ dans cet intervalle. L'intégrale étendue aux divisions rouges sera plus petite que $\Sigma M\varepsilon$; l'intégrale étendue aux divisions noires sera plus grande que $\Sigma m\varepsilon$; la différence sera donc plus petite que $\Sigma(M - m)\varepsilon$. Mais, si la fonction φ est supposée continue; si, d'autre part, l'intervalle ε est très petit par rapport à l'angle total parcouru par l'aiguille, la différence $M - m$ sera très petite. La différence des deux intégrales sera donc très petite, et la probabilité sera très voisine de $\frac{1}{2}$.

On comprend que, sans rien savoir de la fonction φ , je dois agir comme si la probabilité était $\frac{1}{2}$. On s'explique, d'autre part, pourquoi, si, me plaçant au point de vue objectif, j'observe un certain nombre de coups, l'observation me donnera à peu près autant de coups noirs que de coups rouges.

Tous les joueurs connaissent cette loi objective; mais elle les entraîne dans une singulière erreur, qui a été souvent relevée, et dans laquelle ils retombent toujours. Quand la rouge est sortie, par exemple, six fois de suite, ils mettent sur la noire, croyant jouer à coup sûr; parce que, disent-ils, il est bien rare que la rouge sorte sept fois de suite.

En réalité, leur probabilité de gain reste $\frac{1}{2}$. L'observation montre, il est vrai, que les séries de sept rouges consécutives sont très rares; mais, les séries de six rouges suivies d'une noire sont tout

aussi rares. Ils ont remarqué la rareté des séries de sept rouges; s'ils n'ont pas remarqué la rareté des séries de six rouges et une noire, c'est uniquement parce que de pareilles séries frappent moins l'attention.

V. — LA PROBABILITÉ DES CAUSES.

J'arrive aux problèmes de probabilité des causes, les plus importants au point de vue des applications scientifiques. Deux étoiles, par exemple, sont très rapprochées sur la sphère céleste; ce rapprochement apparent est-il un pur effet de hasard, et ces étoiles, quoique à peu près sur un même rayon visuel, sont-elles placées à des distances très différentes de la Terre et, par conséquent, très éloignées l'une de l'autre? Ou bien, correspond-il à un rapprochement réel? C'est là un problème de probabilité des causes.

Je rappelle d'abord qu'au début de tous les problèmes de probabilité des effets qui nous ont occupés jusqu'ici, nous avons toujours dû placer une convention plus ou moins justifiée. Et, si le plus souvent le résultat était, dans une certaine mesure, indépendant de cette convention, ce n'était qu'à la condition de certaines hypothèses qui nous permettraient de rejeter *a priori* les fonctions discontinues, par exemple, ou certaines conventions saugrenues.

Nous retrouverons quelque chose d'analogue, en nous occupant de la probabilité des causes. Un effet peut être produit par la cause A ou par la cause B. L'effet vient d'être observé; on demande la probabilité pour qu'il soit dû à la cause A; c'est la probabilité de la cause *a posteriori*. Mais, je ne pourrais la calculer, si une convention plus ou moins justifiée ne me faisait connaître *d'avance* quelle est la probabilité *a priori*, pour que la cause A entre en action; je veux dire la probabilité de cet événement, pour quelqu'un qui n'aurait pas encore observé l'effet.

Pour mieux m'expliquer, je reviens à l'exemple du jeu d'écarté, cité plus haut; mon adversaire donne pour la première fois et il tourne le roi; quelle est la probabilité pour que ce soit un grec?

Les formules ordinairement enseignées donnent $\frac{8}{9}$, résultat évidemment bien surprenant. Si on les examine de plus près, on voit qu'on fait le calcul comme si, *avant de nous asseoir à la table de jeu*, j'avais considéré qu'il y avait une chance sur deux pour que mon adversaire ne fût pas honnête. Hypothèse absurde, puisque, dans ce cas, je n'aurais certainement pas joué avec lui; et c'est ce qui explique l'absurdité de la conclusion.

La convention sur la probabilité *a priori* était

injustifiée; c'est pour cela que le calcul de la probabilité *a posteriori* m'avait conduit à un résultat inadmissible. On voit l'importance de cette convention préalable; j'ajouterai même que, si l'on n'en faisait aucune, le problème de la probabilité *a posteriori* n'aurait aucun sens; il faut toujours le faire, soit explicitement, soit tacitement.

Passons à un exemple d'un caractère plus scientifique. Je veux déterminer une loi expérimentale; cette loi, quand je la connaîtrai, pourra être représentée par une courbe; je fais un certain nombre d'observations isolées; chacune d'elles sera représentée par un point. Quand j'ai obtenu ces différents points, je fais passer une courbe entre ces points en m'efforçant de m'en écarter le moins possible et, cependant, de conserver à ma courbe une forme régulière, sans points anguleux, sans inflexions trop accentuées, sans variation brusque du rayon de courbure. Cette courbe me représentera la loi probable, et j'admets, non seulement qu'elle me fait connaître les valeurs de la fonction intermédiaires entre celles qui ont été observées, mais encore qu'elle me fait connaître les valeurs observées elles-mêmes plus exactement que l'observation directe (c'est pour cela que je la fais passer près de mes points et non pas par ces points eux-mêmes).

C'est là un problème de probabilité des causes. Les effets, ce sont les mesures que j'ai enregistrées; ils dépendent de la combinaison de deux causes: la loi véritable du phénomène et les erreurs d'observation. Il s'agit, connaissant les effets, de chercher la probabilité pour que le phénomène obéisse à telle loi, et pour que les observations aient été affectées de telle erreur. La loi la plus probable correspond alors à la courbe tracée, et l'erreur la plus probable d'une observation est représentée par la distance du point correspondant à cette courbe.

Mais, le problème n'aurait aucun sens si, avant toute observation, je ne me faisais une idée *a priori* de la probabilité de telle ou telle loi, et des chances d'erreur auxquelles je suis exposé.

Si mes instruments sont bons (et cela, je le savais avant d'avoir observé), je ne permettrai pas à ma courbe de s'écarter beaucoup des points qui représentent les mesures brutes. S'ils sont mauvais, je pourrai m'en éloigner un peu plus, afin d'obtenir une courbe moins sinueuse; je sacrifierai davantage à la régularité.

Pourquoi donc est-ce que je cherche à tracer une courbe sans sinuosités? C'est parce que je considère *a priori* une loi représentée par une fonction continue (ou par une fonction dont les dérivées d'ordre élevé sont petites), comme plus probable qu'une loi ne satisfaisant pas à ces conditions. Sans

cette croyance, le problème dont nous parlons n'aurait aucun sens; l'interpolation serait impossible; on ne pourrait déduire une loi d'un nombre fini d'observations; la science n'existerait pas.

Il y a cinquante ans, les physiciens considéraient une loi simple comme plus probable qu'une loi compliquée, toutes choses égales d'ailleurs. Ils invoquaient même ce principe en faveur de la loi de Mariotte contre les expériences de Regnaut. Aujourd'hui, ils ont répudié cette croyance; que de fois pourtant ne sont-ils pas obligés d'agir comme s'ils l'avaient conservée! Quoi qu'il en soit, ce qui reste de cette tendance, c'est la croyance à la continuité, et nous venons de voir que, si cette croyance disparaissait à son tour, la science expérimentale deviendrait impossible.

VI. — LA THÉORIE DES ERREURS.

Nous sommes ainsi amenés à parler de la théorie des erreurs, qui se rattache directement au problème de la probabilité des causes. Ici encore nous constatons des *effets*, à savoir un certain nombre d'observations discordantes, et nous cherchons à deviner les *causes*, qui sont d'une part la véritable valeur de la quantité à mesurer, d'autre part l'erreur commise dans chaque observation isolée. Il faudrait calculer quelle est *a posteriori* la grandeur probable de chaque erreur, et, par conséquent, la valeur probable de la quantité à mesurer.

Mais, ainsi que je viens de l'expliquer, on ne saurait entreprendre ce calcul, si l'on n'admettait *a priori*, c'est-à-dire avant toute observation, une loi de probabilité des erreurs. Y a-t-il une loi des erreurs?

La loi des erreurs admise par tous les calculateurs est la loi de Gauss, qui est représentée par une certaine courbe transcendante connue sous le nom de « courbe en cloche ».

Mais d'abord il convient de rappeler la distinction classique entre les erreurs systématiques et accidentelles. Si nous mesurons une longueur avec un mètre trop long, nous trouverons toujours un nombre trop faible et il ne servira à rien de recommencer la mesure plusieurs fois; c'est là une erreur systématique. Si nous la mesurons avec un mètre exact, nous pourrions nous tromper cependant, mais nous nous tromperons tantôt en plus, tantôt en moins, et, quand nous ferons la moyenne d'un grand nombre de mesures, l'erreur tendra à s'atténuer. Ce sont là des erreurs accidentelles.

Il est évident d'abord que les erreurs systématiques ne peuvent satisfaire à la loi de Gauss; mais les erreurs accidentelles y satisfont-elles? On a tenté un grand nombre de démonstrations; presque toutes sont de grossiers paralogismes.

On peut néanmoins démontrer la loi de Gauss en partant des hypothèses suivantes : l'erreur commise est la résultante d'un très grand nombre d'erreurs partielles et indépendantes ; chacune des erreurs partielles est très petite et obéit d'ailleurs à une loi de probabilité quelconque, sauf que la probabilité d'une erreur positive est la même que celle d'une erreur égale et de signe contraire. Il est évident que ces conditions seront remplies souvent, mais pas toujours, et nous pourrions réserver le nom d'accidentelles aux erreurs qui y satisfont.

On voit que la méthode des moindres carrés n'est pas légitime dans tous les cas ; en général, les physiciens s'en défient plus que les astronomes. Cela tient sans doute à ce que ces derniers, outre les erreurs systématiques qu'ils rencontrent comme les physiciens, ont à lutter avec une cause d'erreur extrêmement importante et qui est tout à fait accidentelle ; je veux parler des ondulations atmosphériques. Aussi, il est très curieux d'entendre un physicien discuter avec un astronome au sujet d'une méthode d'observation : le physicien, persuadé qu'une bonne mesure vaut mieux que beaucoup de mauvaises, se préoccupe avant tout d'éliminer à force de précautions les dernières erreurs systématiques, et l'astronome lui répond : « Mais vous ne pourrez observer ainsi qu'un petit nombre d'étoiles ; les erreurs accidentelles ne disparaîtront pas ».

Que devons-nous conclure ? Faut-il continuer à appliquer la méthode des moindres carrés ? Nous devons distinguer : nous avons éliminé toutes les erreurs systématiques que nous avons pu soupçonner ; nous savons bien qu'il y en a encore, mais nous ne pouvons les découvrir ; cependant il faut prendre un parti et adopter une valeur définitive, qui sera regardée comme la valeur probable ; pour cela, il est évident que ce que nous avons de mieux à faire, c'est d'appliquer la méthode de Gauss. Nous n'avons fait qu'appliquer une règle pratique se rapportant à la « probabilité subjective ». Il n'y a rien à dire.

Mais l'on veut aller plus loin, et affirmer que non seulement la valeur probable est de tant, mais que l'erreur probable commise sur le résultat est de tant. *Cela est absolument illégitime* ; cela ne serait vrai que si nous étions sûrs que toutes les erreurs systématiques sont éliminées, et nous n'en savons

absolument rien. Nous avons deux séries d'observations ; en appliquant la règle des moindres carrés, nous trouvons que l'erreur probable sur la première série est deux fois plus faible que sur la seconde. La seconde série peut cependant être meilleure que la première, parce que la première est peut-être affectée d'une grosse erreur systématique. Tout ce que nous pouvons dire, c'est que la première série est *probablement* meilleure que la seconde, puisque son erreur accidentelle est plus faible, et que nous n'avons aucune raison d'affirmer que l'erreur systématique est plus grande pour une des séries que pour l'autre, notre ignorance à ce sujet étant absolue.

VII. — CONCLUSIONS.

Dans les lignes qui précèdent, j'ai posé bien des problèmes sans en résoudre aucun. Je ne regrette pas cependant de les avoir écrites, car elles inviteront peut-être le lecteur à réfléchir sur ces délicates questions.

Quoi qu'il en soit, il y a certains points qui semblent bien établis. Pour entreprendre un calcul quelconque de probabilité, et même pour que ce calcul ait un sens, il faut admettre, comme point de départ, une hypothèse ou une convention qui comporte toujours un certain degré d'arbitraire. Dans le choix de cette convention, nous ne pouvons être guidés que par le principe de raison suffisante. Malheureusement ce principe est bien vague et bien élastique et, dans l'examen rapide que nous venons de faire, nous l'avons vu prendre bien des formes différentes. La forme sous laquelle nous l'avons rencontré le plus souvent, c'est la croyance à la continuité, croyance qu'il serait difficile de justifier par un raisonnement apodictique, mais sans laquelle toute science serait impossible. Enfin, les problèmes où le calcul des probabilités peut être appliqué avec profit sont ceux où le résultat est indépendant de l'hypothèse faite au début, pourvu seulement que cette hypothèse satisfasse à la condition de continuité.

H. Poincaré,

de l'Académie des Sciences,
Président du Bureau des Longitudes,
Professeur de Mécanique Céleste
à la Sorbonne.