**La Maxime de Hume et les probabilités**

**Croyance et probabilité**

*La probabilité au sens de Hume*

La « Maxime de Hume » par laquelle se termine la première partie de l’enquête sur les miracles est présentée comme un résultat incontestable, dérivé du principe plus général que l’homme sage doit proportionner sa croyance à l’*evidence*. Comme il s’agit de questions de fait, *a posteriori*, et non de questions relevant des relations d’idées, *a priori*, comme la logique et les mathématiques, la fondation d’une croyance sur l’*evidence* ne peut en ce cas avoir le statut d’une démonstration. La fausseté d’une croyance ainsi fondée n’est ni une contradiction en soi, ni la contradiction de principes *a priori*. Ce fondement peut ressortir à la *preuve* s’il s’avère qu’il est impossible (psychologiquement) de douter, parce que l’*evidence* donne toute l’assurance possible, ou seulement à la *probabilité*, si l’assurance reste limitée, et si des raisons en sens contraire viennent jeter un doute sur la croyance. Cela dit, Hume semble admettre qu’il est sage d’accorder sa croyance à ce qui paraît le plus probable, et qu’il ne serait pas sage de la refuser, encore moins de croire contre la probabilité.

Evidemment, on doit se demande ici à quelle notion de probabilité Hume fait référence. Dans le *Traité de la Nature Humaine*, la troisième partie du Livre I est intitulée « Connaissance et probabilité », et Hume explique qu’il s’agit d’une division qui n’est pas de lui (mais qu’on trouve chez Locke[[1]](#footnote-1), et qui est très commune à l’époque de Hume, par exemple chez Butler) et qui oppose la certitude démonstrative des propositions portant sur des relations d’idées à la certitude, qu’on pourrait dire inductive, reposant sur des vérités de fait, notamment des relations causales. L’important est que la démonstration (déduction) donne une certitude absolue, sans degré, alors que le raisonnement inductif donne une certitude relative, avec des degrés : la probabilité[[2]](#footnote-2). Comme Hume estime que certaines croyances de fait, en particulier certaines croyances causales (comme celle que le soleil se lèvera demain, ou que les hommes sont mortels), sont au-delà de tout doute possible (attention il ne s’agit pas de la justification de notre croyance en un fondement objectif du lien causal), il introduit la notion de *preuve*, comme troisième terme entre « connaissance » proprement dite, et « probabilité ». Il écrit à la section XI (du *TNH* I, iii) : « Pour cette raison, il serait peut-être plus juste, pour conserver le sens courant des mots et marquer en même temps les divers degrés d’évidence, de distinguer trois genres du raisonnement humain, *celui qui résulte de la connaissance, celui qui résulte de preuves, celui qui résulte des probabilités*. Par connaissance, j’entends la certitude qui naît d’une comparaison d’idées. Par preuves, les arguments tirés de la relation de la cause à l’effet, complètement délivrés du doute et de l’incertitude. Par probabilité, l’évidence qui s’accompagne encore d’incertitude ». En gros la preuve correspondrait au degré maximal de probabilité, et encore. Mais Hume parle parfois (comme dans l’*Enquête* X) de preuves plus ou moins fortes. On pourrait dire alors que les preuves constituent *les* degrés les plus élevés (sans atteindre la certitude absolue de la démonstration) des probabilités. La probabilité se distingue de la preuve en ce qu’il y a des expériences qui ne vont pas dans le même sens (par exemple des témoignages contradictoires), même s’ils sont globalement favorables à une proposition :

Un homme sage, donc, proportionne sa croyance à l'évidence. Dans des conclusions du type de celles qui se fondent sur une expérience infaillible, il s'attend à l'événement avec le dernier degré d'assurance, et considère son expérience passée comme une *preuve* complète de l'existence future de cet événement. Pour d'autres cas, il procède avec plus de prudence : il soupèse les expériences opposée; il considère quel côté est soutenu par le plus grand nombre d'expériences. Il penche de ce côté sans doute et sans hésitation, et quand enfin il arrête son jugement, l'évidence n'excède pas ce que nous appelons proprement la *probabilité*. (…) Dans tous les cas, il faut mettre en balance les expériences opposées, quand elles sont opposées, et déduire le petit nombre du plus grand, afin de connaître la force exacte de l'évidence la plus forte. (*Enquête* X, §4)

Millican (2010) note deux choses à propos de l’usage de ‘probabilité’ par Hume : d’une part le terme semble qualifier surtout les raisonnements, ou les inférences (comme les démonstrations et les preuves sont des arguments), plutôt que les propositions elles-mêmes. D’autre part, Millican propose de réserver la notion de « raisonnement moral » (et de certitude morale) à l’ensemble des raisonnements non démonstratifs, incluant donc les raisonnements fondés sur la probabilité et ceux fondés sur des preuves. Le raisonnement probable serait celui qui infère une conclusion à partir d’une expérience inconstante, par opposition à la preuve qui est une inférence à partir d’une expérience constante.

Ce que je voudrais souligner, c’est surtout que le sens de « probabilité » est ici assez général, non technique, et renvoie plutôt à la tradition du scepticisme antique, quand Ciceron, dans le *De natura deorum* (5, 12) fait répondre le sceptique, au défi du stoïcien qui lui objecte que sans connaissance (ferme) on ne peut agir, que l’action n’a pas besoin de plus que des probables. On retrouve l’idée chez Locke (*Essay* IV, 15), et chez le contemporain de Hume, l’évêque Joseph Butler, qui énonce que *la probabilité est le guide de la vie[[3]](#footnote-3)*. Le *probable* c’est en gros le *vraisemblable*, ce qui a de meilleures raisons en sa faveur qu’en sa défaveur. Un point notable est que, contrairement à ce qui a pu être l’objet d’une démonstration, et qui reste valable quelles que soient les considérations qu’on y ajoute (raisonnement monotone), ce qui fait l’objet d’un raisonnement par probabilités peut être *défait* par la prise en considération d’éléments nouveaux : le verdict sur une enquête policière, médicale ou géologique peut être abandonné au vu d’éléments nouveaux (des *evidences* au sens anglais) qui font *pencher la balance* dans l’autre sens. Deux idées donc : 1) le probable admet des degrés, 2) le probable peut être modifié dans ses degrés (voire rendu assuré ou improbable) par la prise en compte de nouvelles considérations.

*Degrés de croyance*

Une question se pose alors (ou doit être posée pour la suite) qui ressortit à la philosophie de la croyance. L’idée de degré s’applique-t-il à la croyance comme à la probabilité, et si oui, est-il directement fonction du degré de probabilité justifiant cette croyance ? Faut-il dire que la *croyance* fondée sur une probabilité (un argument de probabilité) a un degré moins élevé que la croyance fondée sur une démonstration, et que son degré varie en fonction de la probabilité ? La question est doublement intéressante. 1) On peut se demander si la croyance (que Napoléon est mort à Sainte Hélène, que 2 et 2 font 4, ou que les miracles n’ont pas lieu) est une affaire de tout ou rien, sans degré – auquel cas le degré de probabilité serait seulement un degré de la justification de la croyance (il est justifié avec une probabilité de x% qu’il pleuvra demain), voire de son contenu (qu’il est probable à x% qu’il pleuvra demain) mais pas de l’état psychologique du *belief -* ou au contraire si la croyance a des degrés (de fermeté) : on croit plus ou moins, plus ou moins fort, entre le point maximal de la certitude, et le point zéro de l’incertitude totale ou du doute (si l’on va dans le négatif il s’agit alors de degrés de la croyance opposée). 2) On peut aussi se demander si cette manière de poser la question ne nous fait pas passer subrepticement à une notion plus technique et plus moderne de probabilité, quelque chose que l’on peut attribuer à un événement (et non à la nature des *evidences* dont on dispose), comme lorsqu’on dit que la probabilité qu’il pleuve demain à Nantes est de 30%, ou que la probabilité d’être contaminé par la grippe dans une salle de classe où il y a un cas de grippe déclaré est de 10%, tandis que sinon elle est de 0,001% (rassurez-vous, je n’en sais rien). En ce cas, on pourrait être tenté de penser que le degré de croyance dans la proposition p est indexé sur la probabilité que l’on attribue à p. Si l’on pense que p a 50% de chances (probabilité) d’être vrai, cela reviendrait à dire que l’on croit au degré 0.5 que p – cas où l’on serait dans le doute, 100% serait la certitude, 0% la certitude que non-p. Entre 50% et 100% (ou entre 0.5 et 1) on aurait des degrés de fermeté de la croyance. Si l’on dit que la croyance est une question de tout ou rien, la question se pose de savoir si on croit que p dès que l’on attribue à p la probabilité de 51%, et sinon à quel degré de probabilité (qui pourrait bien paraître arbitraire) commence la croyance.

Toutes ces questions ont été posées et traitées par les philosophes depuis l’époque de Hume (et même un peu avant). Mais il ne semble pas que Hume lui-même en ait eu une idée claire. Il ne semble pas, en particulier, qu’il ait pensé à rapprocher sa conception de la croyance et de la croyance probable des théories des probabilités qui étaient en train de se constituer, même s’il est difficile de penser qu’il n’en avait aucune connaissance. On peut néanmoins penser que son penchant philosophique était ouvert à l’introduction de considérations scientifiques (mathématiques en l’occurrence) dans le traitement de questions philosophiques, dès lors que ces considérations étaient pertinentes.

Pour justifier les pages qui suivent, je propose donc de tenir les considérations suivantes pour de plausibles expressions ou extrapolations de la pensée de Hume :

1) la croyance est un sentiment (*feeling*), en tant que tel plus ou moins fort, et auquel il n’y a pas d’objection à appliquer la notion de degré. Il y a un saut peut-être entre le doute et la croyance, puis entre un degré très fort (assurance) et la certitude produite par une démonstration, mais entre ces extrêmes il y a un *continuum* de degrés de croyance.

2) La probabilité est rapportée par Hume à l’*evidence* ou aux *raisons de croire* qui n’emportent pas la certitude, au moins au sens où il est intelligible que ces raisons soient vaincues par des raisons en sens contraire (en cela il s’exprime comme nombre de ses contemporains).

3) Hume fait certainement une place à une probabilité objective (d’un fait) fondée sur la statistique, la fréquence, ou le relevé des occurrences observées jusqu’ici, il parlera plus volontiers alors de *chances*: il y a une chance sur 36 que les deux dés, non truqués, roulent et s’arrêtent sur un double 6 ; il y a toutes les chances (et c’en est même presque une certitude de raison) que les deux dés composent un nombre supérieur ou égal à 2.

4) Sans parler de probabilité en ce cas, Hume fait une place très importante à la mesure de la crédibilité : une proposition (ou un fait) est plus ou moins croyable, voire plus ou moins miraculeuse. Il ne s’agit pas alors de mesurer le degré de croyance d’un individu particulier, mais plutôt de mesurer le degré de croyance qu’un homme raisonnable ou sage, devrait avoir en ladite proposition au vu des éléments (*evidences*) dont il dispose. Aujourd’hui certains philosophes parlent de *probabilité subjective* comme mesure de la crédibilité en référence à des connaissances d’arrière-plan, mais la notion est discutable, car on pourrait dire qu’il s’agit d’une mesure objective (pas relative à quelqu’un en particulier). Voir Appendice sur la probabilité.

5) Le principe général de ‘proportionner sa croyance à l’évidence’ semble bien dire que l’on devrait ajuster son degré de croyance au degré de probabilité de l’évidence.

Ce dernier point soulève un problème que je mentionne sans vouloir m’y étendre (mais il est très intéressant) : demander de régler, proportionner, adapter sa croyance (ou son degré de croyance) en fonction d’une donnée particulière, c’est supposer que l’on a le *pouvoir* de le faire. Or la croyance, du moins telle que Hume l’entend, est un sentiment, un *feeling*, passif, qui se produit en nous mais qui ne dépend pas de nous. On ne peut pas décider de croire ou de ne pas croire une proposition, ou de la croire avec telle force, en tout cas pas directement (je peux m’induire à avoir des croyances sur l’histoire de Nantes en lisant un livre sur l’histoire de Nantes, mais quand je le lis, le fait que se produise en moi la croyance que l’édit de Nantes a été signé au Château, cela ne dépend pas de moi). Cela dit, on peut comprendre de deux manières la maxime de proportionner sa croyance à l’évidence : a) ou bien il ne s’agit pas d’une *prescription* mais d’une *description*, disant que les hommes sages ou raisonnables, ont ainsi, de fait, des croyances réglées sur leurs évidences, b) ou bien il faut dire que la maxime ne porte pas sur le sentiment de croyance, mais plutôt sur l’*acceptation*, le fait de *tenir pour vrai* (volontairement) une proposition, qu’on la croie ou pas – à la manière dont le doute méthodique conduit Descartes à rejeter comme fausses (c’est-à-dire à ne pas *accepter*) des propositions qu’il croit naturellement (comme celle qu’il a un corps)[[4]](#footnote-4). Aucune de ces deux interprétations ne me semble pleinement satisfaisante : je ne vois pas de place pour la distinction entre croyance et acceptation chez Hume, et il me semble bien que la maxime a une valeur prescriptive et pas seulement descriptive.

*Le calcul des probabilités et l’évaluation des croyances religieuses*

L’émergence des probabilités à la fin du XVIIe siècle a eu un impact sur les discussions religieuses, dans les deux sens : certains les ont utilisées pour fortifier une apologétique des miracles, ou de la révélation chrétienne (comme Locke, Butler ou Price, voire Pascal avec l’argument du pari), d’autres au contraire pour affaiblir la crédibilité des témoignages historiques en faveur des évangiles (comme Hume et avant lui Wolson). Il y a en tout cas une idée partagée, c’est que le sort de la rationalité de la foi chrétienne se joue avec la probabilité de son contenu, et l’apport des *evidences*. L’idée est notamment exprimée par Butler (sans user du terme de probabilité, la notion est bien présente, notamment avec l’idée que les évidences alléguées en faveur du christianisme ne sont pas un accident) :

[T]he truth of our religion, like the truth of common matters, is to be judged of by all the evidence taken together. And unless the whole series of things which may be alleged in this argument, and every particular thing in it, can reasonably be supposed to have been by accident (for here the stress of the argument for Christianity lies); then is the truth of it proved: in like manner, as if in any common case, numerous events acknowledged, were to be alleged in proof of any other event disputed; the truth of the disputed event would be proved, not only if any one of the acknowledged ones did of itself clearly imply it but, though no one of them singly did so, if the whole of the acknowledged events taken together could not in reason be supposed to have happened, unless the disputed one were true. Joseph Butler *The Analogy of Religion* (1736/1819: 194)

Certains historiens estiment que l’irréligion moderne a une source importante dans la mise en question de la démontrabilité de la doctrine chrétienne, de l’autorité des évangiles, ou simplement du caractère suffisant des *évidences* en faveur de l’existence de Dieu (un autre source serait la conviction croissante qu’un sceptique ou un athée peut être moralement bon). Pourtant, paradoxalement, c’est l’usage des probabilités fait par certains théoriciens en faveur de la doctrine chrétienne qui a manifesté la plus grande maîtrise dans le calcul. C’est notamment le cas de Richard Price et de Thomas Bayes.

Price a été un auteur presque aussi renommé que Hume à son époque en Angleterre (et plus encore dans les colonies du Nouveau Monde)[[5]](#footnote-5). Ses *Quatre dissertations sur les miracles* constituent un ouvrage important de la littérature sur la question au XVIIIe siècle, avec une réponse explicite aux arguments de Hume dans la IVe dissertation. Les deux auteurs se sont mutuellement loués, tout en se critiquant, et Earman pense que Hume a sans doute modifié son texte dans les éditions ultérieures en tenant compte de critiques de Price. Mais à l’époque de Hume, le plus remarquable développement dans la théorisation de la probabilité est venu du mathématicien et pasteur britannique Thomas Bayes (1702-1761), dont les résultats ont été consignés dans l’ouvrage posthume *Essais sur la manière de résoudre un problème dans la doctrine des risques* (*Essay Towards Solving a Problem in the Doctrine of Chances* - 1763), publié justement par Richard Price[[6]](#footnote-6).

Contrairement aux objections de Price, Hume n’a sans doute pas eu connaissance des textes de Bayes, et il est resté ignorant des développements techniques du calcul des probabilités. Il est clair qu’il n’entend pas par ‘probabilité’ un raisonnement qui procède par application des théorèmes du calcul des probabilités. Mais Price et d’autres contemporains de Hume ont compris que ses analyses le conduisaient à parler de degrés de croyances ou de crédibilité, qui sont des quantités auxquelles on peut appliquer le calcul des probabilités. Aujourd’hui certains commentateurs (dont John Earman) considèrent qu’il est justifié de traiter les idées de Hume avec l’appareil technique en question, puisque dans l’*Abstract* du *Traité*, Hume se joint à Leibniz pour se plaindre des auteurs qui, comme Locke, traitent de probabilités de manière trop rapide, alors qu’il s’agit de notre guide dans l’action et dans la vie (cf. la formule de Butler). Hume prétend avoir essayé de parer à ce défaut, et il aurait donc dû utiliser les ressources les plus élaborées s’il en avait eu connaissance (or, il aurait pu en avoir connaissance, puisque les premiers travaux sur les probabilités, dans le sillage de Pascal, remontent à la fin du XVIIe siècle, notamment, par exemple, la question des témoins multiples).

En conséquence, il paraît légitime d’appliquer à Hume, et au texte sur les miracles, des outils dont il n’a pas fait usage, mais qui permettent de comprendre et aussi d’évaluer son approche. Dans un premier temps, je vais proposer une formulation probabiliste de la maxime qui ne fait pas référence à des théories particulière, en suivant l’exposé de Millican (2010). Puis j’exposerai, succinctement, en quoi consiste l’approche bayesienne de la maxime de Hume.

*Quelques notations et principes simples*

Pour les besoins de l’exposé, il convient de se donner quelques notions et notations de base. On notera la probabilité affectée à une proposition p : Pr(p). Si tant est que l’on puisse donner des valeurs numériques (c’est parfois justifié, comme la probabilité de tirer un double 6 aux dés, non pipés, est de 1/36), la probabilité d’une proposition est comprise entre 0 et 1. Si la probabilité est de 1 (ou 100%), cela revient à dire que la proposition est *nécessaire* (si l’on envisage une probabilité objective) ou qu’elle est *certaine* (si l’on envisage que la croyance tient la proposition pour nécessaire). A l’inverse une probabilité de 0 revient à dire que la proposition est impossible, ou certainement fausse. Une probabilité inférieure à 0.5 signifie que la proposition est moins probable que sa négation (et donc sans doute qu’il n’est pas rationnel de la croire), si elle est supérieure à 0.5, elle est plus probable que sa négation (et donc il est rationnel de la croire). La probabilité de la disjonction (p ou ~p) est forcément de 1, comme pour toute tautologie. Or la probabilité d’une disjonction est équivalente à la somme des probabilités des propositions disjointes : Pr(p v q) = Pr(p) + Pr(q). On peut donc dire que la somme des probabilités d’une proposition et de sa disjonction est de 1 : Pr(p) + Pr(~p) = 1. Ce qui permet de dire que la probabilité de la négation Pr(~p) = 1 – Pr(p).

Pour la suite nous aurons également et surtout besoin de la notion de *probabilité conditionnelle*. Il faut entendre par là, la probabilité qu’une proposition soit vraie *étant admis qu’ (si) une autre proposition est vraie*. On la note Pr(X|Y). Et cette probabilité est équivalente à la probabilité de la conjonction des deux propositions, divisée par la probabilité de la condition : Pr(X|Y) = Pr(X&Y)/Pr(Y), quand Pr(Y)≠0. A noter que la probabilité d’une conjonction est le produit des probabilités des conjoints seulement si ceux-ci sont indépendants (la probabilité de tirer un 6 avec le premier dé, puis un 6 avec le second dé est ainsi de 1/6 x 1/6 = 1/36 qui est bien la probabilité globale de tirer deux 6). La notion de probabilité conditionnelle sera surtout importante quand nous aborderons l’approche bayesienne, car il s’agira d’envisager l’*actualisation* des probabilités que nous donnons aux propositions *en fonction* de l’acquisition de nouvelles connaissances. On remarquera alors qu’en fait la probabilité affectée à une proposition est toujours fonction de connaissances d’arrière-plan, donc d’autres propositions tenues pour vraies : qu’il s’agisse des vérités logiques, ou de vérités empiriques qui vont de soi et ne sont pas en cause. On notera cet ensemble de connaissances d’arrière-plan ‘k’ et la probabilité intrinsèque d’une proposition sera en fait compte tenu de ces connaissances d’arrière-plan, soit Pr(p|k). Pour aller vite, j’omettrai souvent cet ajout (mais le mentionnerai quand il sera important).

Rappelons qu’il y a un argument que Hume ne défend certainement pas et qui dirait que puisque tout miracle est une violation des lois de la nature, et que *par définition* ou *par nécessité* une loi ne peut pas être violée, un miracle est impossible. C’est peut-être l’argument de Spinoza, mais cela revient à produire une *connaissance* par une *démonstration* *a priori* de l’impossibilité du miracle, et Hume n’aurait pas rédigé toute l’enquête sur les miracles si la solution était aussi simple. Il produit plutôt un argument, complexe, pour montrer qu’on n’est jamais fondé à croire un récit de miracle. Bien qu’il distingue *preuve* et *probabilité*, et estime que les lois de la nature, et l’absence d’exception (violation), font l’objet d’une preuve (en sons sens), il ne peut pas non plus vouloir dire que la probabilité d’une loi sans exception est de 1. Cela rendrait effectivement le miracle impossible (probabilité de 0), mais ce serait à nouveau un argument trop rapide, qui reviendrait, au moins psychologiquement, au même qu’une démonstration de l’impossibilité du miracle. Qui plus est, le processus de l’induction (dont Hume a montré qu’il correspondait sans doute à une tendance de la nature, mais n’avait pas de justification), repose sur l’observation de conjonctions répétées, et ne permet que de tenir le futur pour probable, jamais pour absolument certain (inévitable, nécessaire). C’est ce qu’avait noté Price, et qui a sans doute conduit Hume à *affaiblir* son texte, au moins dans la première partie de l’enquête sur les miracles:

“It must, however, be remembered, that the greatest uniformity and frequency of experience will not offer a proper *proof,* that an event will happen in a future trial, or even render it so much as probable, that it will *always* happen in all future trials” (*Four Dissertations* 392–393: 160).

On voit que l’usage de ‘preuve’ par Price n’est pas celui de Hume. Mais c’est que Hume, on l’a vu, reprend la notion de preuve pour faire une distinction au sein de ce que tous ces auteurs appellent probabilités (et que nous avons proposé d’appeler ‘raisonnement moral), pour marquer celles qui donnent une grande assurance. En ce sens Hume envisage qu’un miracle puisse faire lui aussi l’objet de preuve, et qu’il faille alors distinguer la preuve la plus forte (la meilleure), soit celle qui produit la plus grande probabilité. On pense à la situation d’un procès où le procureur établit avec des ‘preuves’ accablantes la culpabilité de M.X. Ce serait une preuve au sens de Hume : les jurés, après la plaidoirie, sont *certains* que M.X est coupable. Vient l’avocat qui produit des ‘preuves’ en sens contraire : alibi solide, réfutations de certains arguments du procureur, preuves en faveur de la culpabilité de M.Y. Les jurés pourraient alors cesser d’être certains de la culpabilité de M.X, voire être totalement hésitants, ou même tenir l’innocence de M.X pour probable, voire très probable, et passer d’une certitude à la certitude opposée. L’important est qu’aucune des deux *preuves* ne donnent une probabilité de 1, même si elle peut être très proche. Bref, une *preuve* au sens de Hume ne peut donner qu’une probabilité élevée (et une *probabilité* au sens de Hume donnerait une probabilité plus faible). Par conséquent, il est toujours possible, moyennant des informations supplémentaires, qu’une proposition très probable devienne improbable, ou qu’une proposition très improbable devienne très probable.

**Lecture probabiliste de la Maxime de Hume**

Revenons maintenant à la Maxime de Hume (première partie de la phrase qui énonce la maxime générale) :

Aucun témoignage n'est suffisant pour établir un miracle à moins que le témoignage soit d'un genre tel que sa fausseté serait plus *miraculeuse* que le fait qu'il veut établir

Millican note qu’en disant ‘suffisant pour établir’, Hume ne doit pas vouloir dire ‘suffisant pour établir un miracle au-delà de tout doute raisonnable’ ou ‘avec une certitude entière’, mais plutôt ‘suffisant pour établir que l*’occurrence* d’un miracle *est plus probable que sa non-occurrence*’. Cela s’accorde mieux avec la suite de la phrase, et surtout avec l’illustration suivante :

Quand quelqu'un me dit qu'il a vu un mort revenu à la vie, je considère immédiatement en moi-même s'il est *plus probable* que cette personne me trompe ou soit trompée, ou que le fait qu'elle relate ait réellement eu lieu. *Je soupèse les deux miracles*, et selon la supériorité que je découvre, je rends ma décision *et rejette toujours le plus grand miracle*. Si la fausseté de son témoignage était plus miraculeuse que l'événement qu'elle relate, alors, et alors seulement, cette personne pourrait prétendre commander ma croyance et mon opinion

On voit que ‘plus probable’ correspond ici à ‘moins miraculeux’ si l’on peut dire, en tout cas que le caractère miraculeux est retraduit en termes de probabilité. Bref, il s’agit bien de comparer les deux probabilités, correspondant de manière inverse à deux degrés de miraculosité (plus la chose est miraculeuse, moins elle est probable) : celle concernant la fausseté du témoignage (évaluée indépendamment de ce qui est arrivé) et celle concernant l’occurrence du miracle (évaluée indépendamment du témoignage). Ce n’est que si les deux probabilités sont égales que la suspension du jugement sera justifiée. Mais il suffirait que la probabilité que le témoignage soit faux soit inférieure (même de peu) à celle de l’occurrence du miracle pour que la croyance (opinion) puisse être *commandée* par le témoignage.

Une interprétation de la maxime en termes probabilistes aura la forme suivante, ou la probabilité est présentée comme mesure de la crédibilité d’une proposition ou d’un fait :

Aucun témoignage n'est suffisant pour rendre l’occurrence d’un miracle plus probable (crédible) que sa non-occurrence à moins que le témoignage soit d'un genre tel que sa fausseté soit encore moins probable (crédible) que le fait qu'il veut établir[[7]](#footnote-7)

Comme le dit Millican, il y a comme un procès où l’on doit évaluer la plus grande crédibilité : celle du ou des témoins humains (en faveur du miracle), et celle du témoignage de l’expérience (contre le miracle)[[8]](#footnote-8).

La maxime de Hume peut donc être formulée en termes de probabilité conditionnelle, en notant la proposition que le miracle a eu lieu, M, et le témoignage rapportant le miracle t(M). Plusieurs tentatives différentes ont été faites, car l’interprétation de la maxime est sujette à discussion. J’en indique deux. Celle de Price, sans symbolisation, hésite entre

(P) Pr(M|t(M)) > 0.5 si et seulement si Pr(M) > Pr(~M|t(M))[[9]](#footnote-9)

et

(P’) Pr(M|t(M)) > 0.5 si et seulement si Pr(M) > Pr(t(M) |~M)[[10]](#footnote-10)

John Earman et Peter Millican (que je suis) s’accordent plutôt pour noter la maxime de la manière suivante

(E/M) Pr(M|t(M)) > 0.5 si et seulement si Pr(M|t(M)) > Pr(~M|t(M))

Autrement dit : pour que la probabilité d’un miracle faisant l’objet d’un témoignage le rende crédible (plus probable que sa négation ou plus probable que 0.5) il faut et suffit que la probabilité que le miracle ait lieu si le témoignage en faveur du miracle est rendu soit supérieure à la probabilité qu’il n’ait pas lieu, alors que le témoignage est rendu. C’est une formulation beaucoup moins intéressante, et presque triviale[[11]](#footnote-11). A moins que l’on ne puisse en tirer quelque chose. C’est ce que nous allons voir maintenant.

Dans ce qui suit, je vais reprendre l’analyse de Millican (2010) aux §8 et 9 et la discussion du §19.

Il est clair que dans la situation envisagée, où l’on donne les meilleures chances possibles aux témoignages en faveur du miracle, *la probabilité que le témoignage soit faux,* que l’on peut noter *f*, doit être petite. Mais s’il doit bien s’agir d’un miracle, d’une violation des lois de la nature, allant contre tout le témoignage de l’expérience, *la probabilité que cet événement se produise*, que l’on notera *m*, doit aussi être (très) petite. Il y a ainsi ‘preuve contre preuve’, et il faut parvenir à trancher, car ou bien le miracle *M* a eu lieu, et alors le témoignage humain est vrai et le témoignage inductif de l’expérience n’est pas probant (il peut justifier qu’il y a une régularité, mais pas que cette régularité n’admet pas d’exception), ou alors *M* ne s’est pas produit et c’est l’inverse. Pour trancher, il semble que la seule possibilité soit de comparer les forces des deux témoignages (ou des deux preuves), soit *f* et *m*. En reprenant les principes de notation donnés plus haut, la maxime de Hume revient simplement à ceci :

Pr(*M*|*t(M)*) > 0.5 —> f < m

Qu’on peut lire : Il est probable (plus probable que non, ou d’une probabilité supérieure à 0.5) qu’un événement de type M (miracle) a eu lieu étant donné un certain témoignage en sa faveur (t(M)) *seulement si* la probabilité que le témoignage soit faux (*f*) est inférieure à la probabilité (antérieure) que l’événement ait eu lieu (*m*).

Ainsi reformulée, la maxime de Hume peut se voir exposée à plusieurs critiques radicales.

* l’une dit que la maxime est simplement *triviale* et ne fait que reformuler avec une apparente technicité une vérité de La Palisse[[12]](#footnote-12)
* Une autre objection dit au contraire que la maxime est *fausse* puisqu’elle aboutit à dire que pour tout événement historique, par essence improbable et ne pouvant faire fond sur aucune induction, elle devrait aboutir à le tenir pour aussi peu crédible qu’un miracle saugrenu. Mais nous nous appuyons sans hésiter sur les témoignages historiques[[13]](#footnote-13)
* Enfin, une troisième objection fait valoir que la maxime rendrait impossible le progrès scientifique qui se fait parfois sur la base de découvertes très improbables et qui n’auraient pas passé le test de la maxime[[14]](#footnote-14)

*La maxime n’est pas triviale*

Pour que la Maxime de Hume ne soit pas triviale, il faut que la probabilité du témoignage ne soit pas équivalente par définition à celle de l’événement rapporté. Ce serait le cas si la probabilité du témoignage était comprise comme la probabilité que soit vrai tel témoignage particulier, c’est-à-dire non seulement de telle personne particulière mais aussi portant sur tel événement particulier. En ce cas la probabilité que l’évènement rapporté ait eu lieu serait par définition la même que celle que le témoignage est vrai. Et la maxime reviendrait à dire qu’un récit de miracle n’est croyable que si l’occurrence du miracle, étant donné ce témoignage, est plus probable que sa non-occurrence, ce qui est trivial.

Mais la maxime n’est pas triviale si la probabilité porte sur un *type* de témoignage (de cette personne éventuellement, mais considérée comme témoin en général, quel que soit le contenu de son témoignage particulier). Or, Hume semble bien parler de la probabilité d’une *sorte* ou d’un ou d’un *type* de témoignage, abstraction faite du contenu du témoignage : c’est le type de témoignage donné par une certaine personne ou par un certain type de personne, voire par un certain nombre de témoins. Il adopte pour ce faire une *hypothèse d’indépendance*. La probabilité du témoignage est une évaluation de la crédibilité du ou des témoins *indépendamment* du contenu du témoignage. La crédibilité est évaluée par des critères portant sur la personne du témoin (ou des témoins) : fiabilité dans le passé, qualités d’honnêteté, éventuellement absence d’intérêt relativement à ce qui est rapporté, etc., nombre des témoins et cohérence de leurs témoignages.

Hypothèse d’Indépendance :

Différentes *sortes* de témoignages (spécifiées en termes du caractère et du nombre des témoins, de leur consistance, de la façon dont ils délivrent leur témoignage) ont une probabilité de vérité et de fausseté typique, *indépendamment de l’événement qu’ils rapportent*.

Dérivation mathématique de la maxime :

Une fois adoptée l’Hypothèse d’Indépendance, on peut dériver la maxime, en suivant l’analyse de Millican (2010, §8). Considérons un certain type de témoignage (témoin), dont la probabilité de fausseté est *f* (la probabilité que le témoignage soit vrai est complémentaire : 1–*f*)[[15]](#footnote-15). Dans un cas particulier ce témoignage porte sur un événement de type *M*, dont la probabilité d’une occurrence est *m* (et la probabilité qu’il ne se produise pas est donc 1–*m*). Avec l’Hypothèse d’Indépendance, qui nous assure que la probabilité de la fausseté du témoignage ne dépend pas de ce qui est rapporté, on peut calculer les probabilités du *Vrai Positif* (le miracle a lieu et le témoignage rapporte qu’il a lieu) et du *Faux Positif* (le miracle n’a pas lieu, mais il est rapporté par le témoin)[[16]](#footnote-16):

*Vrai Positif*: Pr(*M* & *t(M)*) = Pr(*M*) x Pr(le témoignage de ce type est vrai) = *m*.(1-*f*)

*Faux Positif*: Pr(~*M* & *t*(*M*)) = Pr(~*M*) x Pr(le témoignage de ce type est faux) = (1-*m*).f

Si un témoignage rapporte qu’un événement de type *M* a eu lieu, ce témoignage sera probablement vrai seulement si un *Faux positif* est moins probable qu’un *Vrai Positif*, soit :

Pr(*M*|*t(M)*) > 0.5 —> (1-*m*).f < *m*.(1-*f*), ce qui revient à Pr(*M*|*t(M)*) > 0.5 —> f < m

La maxime peut bien être dérivée par ce calcul simple (que Hume ne fait pas).

On voit que la maxime n’est pas triviale à ce qu’elle permet de corriger nos intuitions, comme souvent égarantes en matière de probabilités[[17]](#footnote-17). Je reprends l’illustration de Millican (§9). Soit une condition génétique G qui affecte une personne sur un million. Seul un certain test bien établi permet de savoir si vous avez cette condition (autrement inapparente), et ce test a une probabilité de 99,9% de donner un résultat correct, que ce résultat soit positif (l’individu a la condition G) ou négatif (il n’a pas la condition G) – de sorte que l’Hypothèse d’Indépendance s’applique (la probabilité que le test donne un résultat correct est indépendante du contenu de ce résultat). Si le test *témoigne* que vous avez la condition G, quelle est la probabilité pour que vous l’ayez ? Les enquêtes psychologiques montrent que la plupart des gens pensent que la probabilité d’avoir la condition G serait alors très élevée (voire serait simplement de 99,9%). Mais la maxime donne un résultat bien différent, et correct : le test est faux une fois sur mille, c’est notre quantité *f* (*f* = 0.001), mais la présence de G chez quelqu’un est beaucoup moins probable, une sur un million (0.000001), grandeur qui correspond au *m* de la maxime. Puisque *f* (la probabilité que le test/témoignage soit faux) est bien supérieure à *m* (la probabilité que la condition G soit présente), la probabilité d’avoir la condition G sur la base d’un test positif fiable à 99,9% reste très faible (de l’ordre de 0.001)[[18]](#footnote-18). Ce n’est que si la probabilité initiale de la condition G est inférieure à un sur mille (0.001) que le test sera utile et commencera à être crédible.

*La maxime est fausse car elle n’est pas universelle : le cas des témoignages historiques*

La maxime n’est pas triviale donc, mais elle est fausse. Ou plutôt, elle n’est correcte que dans l’hypothèse d’indépendance, et cette hypothèse n’est que très rarement satisfaite. Pour reprendre l’exemple du test de la condition génétique G : il est rare qu’un test ait la même probabilité d’être correct quand il donne un résultat positif et un résultat négatif. Surtout, l’hypothèse d’indépendance conduirait effectivement à rendre très difficile à croire tout récit d’un événement en soi improbable. Or tout événement peut être décrit avec suffisamment de précision pour qu’il soit improbable. Un accident de voiture est déjà improbable, mais l’on ne doute guère d’un témoignage rapportant un accident, même si le témoin n’est pas particulièrement fiable. Qui plus est, un accident mettant en cause une voiture jaune avec une camionette verte, juste devant la porte de ma maison à 12h01 est *très* improbable. Pourtant, même si le témoin qui rapporte cet événement est peu fiable (sans être systématiquement ou le plus souvent trompeur), il paraît irrationnel de tenir le récit de cet accident pour incroyable. Ou encore, votre victoire à la loterie est très improbable (a priori), mais si un grand journal très fiable (*Ouest France*?) rapportait que vous avez gagné la loterie, même si vous pourriez avoir du mal à le croire, les autres lecteurs ne remettraient pas en cause cette information sur la base de l’improbabilité du résultat.

Pourtant une application de la maxime de Hume, avec l’Hypothèse d’Indépendance, ne rend pas compte de ces intuitions. Reprenons l’exemple de la loterie. Soit votre probabilité de gagner avec un ticket acheté parmi un million de tickets de une sur un million (10-6), car la loterie est équitable. Soit la probabilité que *Ouest France* se trompe dans les annonces de ce genre de une sur dix mille (10-4) – sa fiabilité est donc de 1 – 10—4 ou 99,9%. Une application de la maxime dirait que vous ne devez pas croire *Ouest France* puisque la probabilité que le journal se trompe (la quantité *f*, qui est de 10-4) est nettement supérieure à la probabilité que l’événement se produise (10-6). Mais c’est aberrant, on voudrait au contraire dire que, dès lors que le journal l’a annoncé, la probabilité que vous ayez gagné est élevée.Nous ne sommes pas enclins à dire que nous ne croirons les témoignages portant sur des faits anodins mais improbables s’ils sont décrits avec suffisamment de précision que si la fausseté du témoignage était plus improbable (était un plus grand miracle) que l’occurrence de l’événement rapporté. La probabilité que mon fils se trompe, ou que *Ouest France* se trompe est plus grande que celle de l’événement rapporté (ou : la fausseté du témoignage est *moins* improbable que l’événement), et pourtant il est rationnel de croire sur la base de ces témoignages que l’événement a eu lieu plutôt que de refuser de le croire (sinon, à ce compte, on ne devrait jamais croire les récits des historiens, ni les témoignages d’événements anodins).

Que se passe-t-il ? Deux choses

1) la probabilité qu’un événement a eu lieu doit être *révisée* sur la base des informations dont nous disposons, notamment sur la base des témoignages rapportant cet événement. Il est très improbable, *a priori* (avant que l’événement ait lieu, ou après, mais sans information particulière), que se produise un accident mettant en cause une voiture jaune avec une camionette verte, juste devant la porte de ma maison à 12h01. Mais il n’est pas du tout improbable *qu’un tel événement ait eu lieu* *si un témoin relativement fiable le rapporte*. Il est très improbable *que vous gagniez à la loterie*, mais il n’est pas du tout improbable *que vous ayez gagné si le journal le rapporte*. On peut même aller plus loin. Dès lors que le journal rapporte un événement (même très improbable a priori), la probabilité que cet événement a eu lieu doit seulement être estimée par la fiabilité du journal : donc il est probable à 1-10-4 ou à 99,9% que vous ayez gagné à la loterie si *Ouest France* le dit. Sans l’information, cette probabilité était très faible (10-6), avec cette information, elle est très forte[[19]](#footnote-19).

2) Ce qui ne va pas, c’est donc l’hypothèse d’indépendance, qui paraît solidaire de la maxime de Hume. Bien qu’elle puisse être maintenue dans certains cas (comme celui envisagé de la condition G et du test dont la fiabilité est *indépendante* du contenu), cette condition n’est pas universellement valide, et l’exemple des témoignages historiques le révèle crûment. On ne peut pas simplement considérer la fiabilité du témoin, indépendamment de la nature de l’événement rapporté : dans de nombreux cas, le contenu du témoignage affecte ce que nous pensons de la fiabilité du témoignage[[20]](#footnote-20). Qui plus est, dans la 2e partie de l’Enquête, Hume lui-même dit que le témoignage humain est sujet à des distorsions en fonction de la matière sur laquelle il porte. Il faut donc évaluer le témoignage en tenant compte de son contenu.

Précisons. L’Hypothèse d’Indépendance conduit à considérer qu’un témoignage est simplement vrai ou faux (ce qui peut sembler la bonne alternative dans le cas du test de la condition génétique). Mais en fait, alors qu’un témoignage ne peut être vrai que d’une façon (à savoir, si ce qu’il rapporte est le cas), il y a le plus souvent bien des manières pour un témoignage d’être faux (certes à chaque fois, il sera faux parce que ce qu’il rapporte n’est pas le cas, mais il peut rapporter quantité de choses différentes qui ne sont pas le cas). Il n’y a qu’un manière de gagner à la loterie : en ayant le bon numéro, mais il y en a bien plus de perdre : en ayant un numéro ou un autre qui n’est pas le bon numéro. La probabilité de perdre est sans doute de 1-10-6, dans une loterie à un million de tickets, mais la probabilité de *perdre avec tel numéro* est bien, bien moindre[[21]](#footnote-21). Ce qu’il faut comparer à la probabilité de l’événement (*m*), ce n’est donc pas la probabilité que le journal ou mon fils dise le faux (*f*), mais la probabilité que l’un ou l’autre donne un témoignage faux avec tel contenu précis. La probabilité que le témoignage soit faux peut être supérieure à la probabilité que l’événement ait eu lieu alors qu’il serait bien moins probable qu’un témoignage faux avec ce contenu précis soit *donné*. Si l’on se contente de la probabilité qu’un témoin dise le faux (par exemple que mon fils dise quelque chose de faux en rapportant un événement survenu dans la rue) on néglige le fait que cette probabilité peut être bien différente de la probabilité qu’un témoin dise le faux en donnant tel contenu. Mon fils peut se tromper (ou mentir) deux fois sur cinq, mais cela ne veut pas dire qu’il y a 2/5 chances qu’il dise le faux *en disant qu*’un accident a mis en cause une voiture jaune avec une camionette verte, juste devant la porte de ma maison à 12h01. La probabilité qu’il dise *cette fausseté* est beaucoup, beaucoup plus faible[[22]](#footnote-22).

Millican propose alors de réviser la Maxime de Hume de la manière suivante :

Maxime de Hume révisée :

Aucun témoignage n’est suffisant pour rendre l’occurrence d’un miracle de type M plus probable que sa non occurrence, à moins que le témoignage ne soit d’une sorte telle que l’occurrence d’un témoignage de cette sorte qu’un M a eu lieu (*étant donné qu’aucun M ne s’est produit*) serait encore moins probable que l’occurrence d’un M

On voit qu’il y a là un changement de perspective: on ne compare plus la probabilité d’un événement à la probabilité qu’un témoin ait dit le faux, quoi qu’il ait dit (simplement sur la base des propriétés du témoin), mais à la probabilité que le témoin dise le faux en délivrant tel message (ou à la probabilité qu’il délivre tel message, faux). Ce changement de perspective paraît bien s’accorder avec le contenu de la 2e partie de l’Enquête, puisqu’il va justement être question des raisons qui rendent probables la production de *faux* témoignages de miracles.

*Une discussion évitable ? Comment distinguer le miraculeux du simplement improbable ?*

Je me demande si toute cette discussion ne pourrait pas être évitée (au sens de rendue inutile pour la question qui nous occupe), si l’on répondait à l’objection selon laquelle la maxime rend impossible de croire les récits historiques d’événements improbables qu’il y a simplement lieu de distinguer entre ‘improbable’ et ‘miraculeux’. Non pas au sens où ‘miraculeux’ serait‘très improbable’ (il va de soi que le miraculeux est très improbable), mais au sens où l’improbabilité (même très élevée) n’épuise pas le sens de ‘miraculeux’. En effet, pour qu’un événement soit miraculeux, il faut surtout qu’il *viole les lois de la nature*. Or un événement peut être très improbable (comme les événements historiques), sans violer les lois de la nature. Hume ne voudrait donc pas dire qu’il est très difficile de croire qu’un événement très improbable a eu lieu, mais qu’un événement miraculeux, violant les lois de la nature, a eu lieu. On pourrait faire ressortir la différence en disant qu’un événement très improbable mais non miraculeux est un événement d’un *type* qui n’est pas improbable : il est improbable que vous gagniez à la loterie (si vous avez acheté un ticket sur un million et qu’il n’y a pas de biais), mais il n’est pas improbable que quelqu’un gagne à la loterie. En revanche, l’improbabilité que telle personne ressuscite est une improbabilité qui vaut pour le type d’événement : il est improbable que quelqu’un, qui que ce soit, ressuscite.

Cela dit, cette remarque, qui me semble juste en soi, n’est peut-être pas accessible à Hume. En effet, l’idée même de loi de la nature est pour lui celle d’une régularité, qui fait que les événements naturels sont probables au regard de leurs causes (antécédents réguliers). Ce que Hume devrait dire c’est qu’un événement comme le gain à la loterie ou à la roulette, ou l’accident de la voiture jaune et de la camionnette verte à 12h01 devant chez moi, est improbable *pour nous*, mais pas *en soi*, ou pour un individu qui connaîtrait toutes les causes de l’événement, alors qu’un événement miraculeux est improbable *en soi*. Cela marquerait la différence du miraculeux et du simplement improbable. Mais une chose est de savoir que cette différence est justifiée, une autre est de pouvoir l’appliquer. Nous ne pouvons connaître une probabilité en soi qui ne soit pas pour nous.

Il faudrait que Hume puisse dire qu’un événement se présente comme un miracle non seulement si aucune régularité connue ne permet de s’y attendre (ne le rend probable pour nous), mais si en plus les régularités connues *excluent* que cet événement puisse arriver (on n’a jamais vu de saumon de 3m, mais est-ce exclu par les lois de la nature ?). Apparemment, aucune régularité connue n’exclut que vous gagniez à la loterie, à la roulette, ou que l’accident mentionné se produise. Mais on pourrait répondre que si : c’est une régularité que vous n’ayiez *jamais* gagné avant, c’est une régularité qu’un tel accident ne s’est *jamais* produit avant. Il faudrait alors montrer que ces régularités-là ne sont pas des lois de la nature. Mais pourquoi, dans la perspective de Hume, où les lois sont connues comme des régularités observées[[23]](#footnote-23) ? Sans doute pourrait-on ajouter que ces régularités doivent *faire système* avec d’autres dont elles se déduisent ou sont des cas particuliers, ou des combinaisons (comme dans les théories scientifiques). Je crois que cette direction est prometteuse, mais c’est un défi depuis longtemps pour les conceptions positivistes des théories scientifiques et des lois de la nature (distinguer les régularités qui sont des lois de celles qui n’en sont pas).

De toute façon, si cette analyse réussisait, elle permettrait de justifier ce que nous avons dit sur les croyances : le témoignage en faveur d’un argument improbable mais non miraculeux est tout à fait crédible parce qu’il est probable qu’un tel témoignage soit produit si l’événement a lieu et très peu probable sinon, de sorte que l’événement est lui-même très probable si le témoignage a été donné. Alors que le témoignage en faveur d’un événement miraculeux, sans doute probable si le miracle a eu lieu, ne contribue pas (ou peu) à rendre probable l’événement en question. Cette considération va également rendre pertinente la considération de l’approche bayesienne de la maxime (révisée) et plus généralement de la problématique de Hume sur la crédibilité des miracles.

**L’approche bayesienne**

De nombreux commentateurs ont adopté l’approche bayesienne, et c’est à la lumière de cette analyse qu’ils critiquent en général l’argument de Hume dans la première partie de l’enquête sur les miracles. Mais il faut noter que d’autres commentateurs pensent qu’il n’y a pas lieu de faire référence aux idées bayesienne pour rendre compte de la pensée de Hume (voir le traitement de Millican), certains considérant même que c’est une mauvaise idée, puisque Hume ne fait aucune référence au calcul des probabilités. Mon avis n’est pas assuré, mais je trouve suffisamment intéressant de regarder le principe de l’application du raisonnement bayesien à la question de Hume pour l’exposer ci-dessous.

La doctrine de Bayes continue à nourir de nombreuses réflexions en philosophie de la croyance et des sciences. Il est courant de caractériser l’approche bayesienne par trois caractéristiques (qui ne sont pas tant énoncées par Bayes que par ses disciples) :

1. L’épistémologie (la question de la connaissance) doit être abordée en termes de degrés de croyance et non en termes de croyance entière (tout ou rien)
2. Les degrés de croyance peuvent (doivent) être soumis au calcul des probabilités (et une non conformité aux résultats du calcul serait signe d’irrationalité)
3. Quand un sujet acquiert une connaissance (par expérience) E relative à une certaine proposition p, la probabilité nouvelle qui mesure son degré de croyance en cette proposition se calcule à partir de la probabilité ancienne (avant l’expérience) par la règle de la probabilité conditionnelle : Prnouvelle (p) = Prancienne(p|E)

Ce dernier point nous renvoie à l’idée de probabilité conditionnelle, exprimée par Pr(X|Y) = Pr(X&Y)Pr(X) quand X≠0. Et il nous renvoie aussi à l’idée (exposée plus haut) que nous réactualisons nos degrés de croyance (donc les probabilités qui les mesurent) en fonction de l’acquisition de connaissances.

On peut évidemment dire que toute croyance est fonction de connaissances préalables, c’est pourquoi il est habituel de toujours noter la probabilité mesurant un degré de croyance de manière conditionnelle . On notera la probabilité d’une croyance relative à une connaissance d’arrière plan Pr (p|k), et la probabilité actualisée en fonction de l’acquisition d’une expérience E, Pr (p|E&k). Pour aller plus vite dans ce qui suit, j’éliminerai parfois la mention de la connaissance d’arrière-plan (k), mais l’idée est importante (nos degrés de croyance en une proposition sont toujours fonction de croyances préalables).

Ce qui a fait la fortune de l’approche bayesienne, c’est surtout le *théorème de Bayes* que l’on peut extraire de son *Essay* (bien qu’il ne s’y trouve pas explicitement). Avant de regarder l’énoncé du théorème, il sera utile d’en donner une idée très informelle. Ou plutôt, il sera utile de dire quelles sont les idées que le théorème formalise. La probabilité d’une hypothèse scientifique (la composition chimique d’une substance) ou policière (l’identité du criminel) pour expliquer un certain phénomène ou événement (une réaction chimique, un crime) peut être considérée *en elle-même* (ou en ne tenant compte que de connaissances générales d’arrière plan) ou *relativement à certaines informations*. Si on considère que certaines informations ont été données (par exemple, des pièces à conviction comme des traces, ou des témoignages), celles-ci sont supposées *élever* la probabilité de l’hypothèse (par rapport à sa probabilité intrinsèque). Mais dans quelle mesure ? Eh bien, sans donner une mesure quantitative, cette élévation est fonction de deux choses : 1) d’une part de la probabilité intrinsèque de ces informations, au sens où plus la donnée est probable *a priori*, moins elle contribue à élever la probabilité de l’hypothèse. Si des traces de M. X sont trouvées à côté de la maison où a été commis le crime, cela contribue à rendre plus probable l’hypothèse que M. X est coupable, mais cette contribution sera d’autant plus intéressante qu’il est *moins*  probable que M. X, *a priori*, passe par là et laisse des traces. Si au contraire, ces traces sont très probables *a priori*, parce qu’il passe tous les jours devant la maison, cette donnée, cet indice, contribuera peu à élever la probabilité de l’hypothèse. 2) D’autre part, la probabilité de l’hypothèse sera élevée par l’information donnée, si elle (l’hypothèse) contribue à rendre cette information plus probable qu’elle ne l’aurait été sans l’hypothèse. Ainsi l’hypothèse que M.X a commis le crime rend plus probable la présence de traces de M. X que s’il ne l’avait pas commis. Dans la mesure où l’hypothèse explique, rend compte de, la présence de l’indice, elle la rend plus probable (l’hypothèse non seulement explique le crime, mais elle explique aussi cette donnée). Un autre indice, comme la nature de l’arme du crime, pourrait ne pas être rendu probable par l’hypothèse que M.X est le criminel. En ce cas, cette donnée ne contribuerait pas non plus, en retour, à rendre probable l’hypothèse que M. X est le criminel. La probabilité de l’hypothèse en fonction de l’information est *directement* fonction de la probabilité de l’information en fonction de l’hypothèse : les deux probabilités conditionnelles (Pr(H|E) et Pr(E|H)) augmentent et diminuent ensemble[[24]](#footnote-24).

Voyons maintenant le théorème de Bayes. Il s’énonce ainsi (en reprenant le symbolisme précédent) :

Pr(H|E&k) = Pr(H|k)xPr(E|H&k) ou plus simplement Pr(H|E) = Pr(H)xPr(E|H)

Pr(E|k) Pr(E)

Le théorème dit que la probabilité d’une certaine hypothèse (proposition) H relativement à une certaine expérience E (et à la connaissance d’arrière plan k) est fonction de la probabilité de l’hypothèse en elle-même (ou relative seulement à une connaissance d’arrière-plan), qu’on appelle la *probabilité antérieure* de l’hypothèse (Pr(H|k)), de la probabilité de cette expérience relativement à la vérité de l’hypothèse (Pr(E|H&k)), et de celle de la probabilité de l’expérience en elle-même (probabilité antérieure de l’expérience Pr(E|k)), mais cette fois de manière inversement proportionnelle (plus l’expérience est probable en elle-même, moins elle contribue à probabiliser l’hypothèse, ou : l’expérience contribue d’autant plus à probabiliser l’hypothèse qu’elle est en elle-même moins probable).

Pour faire passer cette formule très abstraite, je propose de l’illustrer à nouveau par deux exemples simples (mais l’exemple initial de l’enquete policière peut aussi servir). Ne vous laissez pas désarçonner par le premier, au contraire, il devrait être facile à comprendre, moyennant un minimum de réflexion.

Exemple 1 : la probabilité de l’hypothèse (de l’existence) du Père Noël, étant donné qu’il y a des cadeaux dans la cheminée (probabilité conditionnelle : Pr(PN|E)), est fonction de (et augmente en fonction de) la probabilité intrinsèque de l’existence du Père Noël (indépendamment du fait qu’il y ait ou non des cadeaux, Pr(PN)) et de la probabilité qu’il y ait des cadeaux si le PN existe (Pr(E|PN)). Elle est aussi fonction, mais de manière inverse, de la probabilité qu’il y ait des cadeaux dans la cheminée (indépendamment du fait qu’il y ait ou non le PN, soit Pr(E)) – c’est-à-dire qu’elle augmente quand cette dernière diminue. Autrement dit, l’hypothèse (le PN existe) sera rendue d’autant plus probable par l’expérience (présence des cadeaux) que cette hypothèse sera intrinsèquement probable, qu’elle contribuera à rendre probable l’expérience, et que l’expérience sera *moins* probable en elle-même.

L’exemple permet d’illustrer deux points

1) Comme on l’avait vu avec l’exemple policier, la probabilité d’une hypothèse sous la condition qu’une expérience est réalisée (que le PN existe s’il y a des cadeaux, Pr(PN|E)) est fonction de la probabilité de l’expérience si l’hypothèse est vraie (qu’il y ait des cadeaux si le PN existe Pr(E|PN)). L’une augmente avec l’autre, même si la première peut être faible et la seconde élevée (comme on peut le supposer pour le PN). Le fait qu’une hypothèse rende probable une expérience (par exemple parce qu’elle en fournit une bonne explication) augmente donc la probabilité de l’hypothèse quand l’expérience est réalisée (le fait que le PN rende probable la présence des cadeaux nous indique que la présence des cadeaux élève la probabilité de l’hypothèse du PN).

2) Dans notre exemple, néanmoins, il est clair que, si l’hypothèse du PN rend probable la présence des cadeaux (Pr(PN|E) est élevée), sa probabilité intrinsèque (Pr(PN) ou Pr (PN|k)) est faible, alors que la probabilité intrinsèque qu’il y ait des cadeaux (Pr(E) ou Pr(E|k)) n’est pas si faible (en particulier parce que d’autres hypothèses, moins improbables que l’existence du PN, rendent probable la présence des cadeaux). Donc, bien que le fait que le PN rende probable la présence des cadeaux (c’est une très bonne explication de l’expérience : si elle est vraie, il est très probable que l’expérience – de la présence des cadeaux – se réalise), et contribue donc à montrer que la présence des cadeaux élève la probabilité de l’existence du PN, cela est parfaitement compatible avec le fait que la probabilité de l’existence du PN reste faible, même étant donnée la présence des cadeaux.

Exemple 2. Y a-t-il telle substance ABC dans le sol de mon jardin ? C’est peu probable *a priori* (probabilité antérieure), par exemple parce qu’il n’y en a pas dans la région. Y a-t-il telles fleurs F dans mon jardin ? C’est peu probable *a priori*, par exemple parce qu’il n’y en a pas chez les voisins. Or, la présence de la substance ABC rendrait (très) probable la présence des fleurs F, donc Pr(F|ABC) > Pr(F). Dans quelle mesure la probabilité de la présence de la substance ABC est-elle élevée par la présence des fleurs F dans mon jardin ? Le théorème de Bayes me dit que la probabilité (postérieure) qu’il y ait la substance ABC s’il y a les fleurs F (Pr(ABC|F)) est fonction des trois probabilités mentionnées : probabilité antérieure qu’il y ait ABC, probabilité antérieure qu’il y ait des F, probabilité postérieure qu’il y ait des F s’il y a de l’ABC.

Si la probabilité antérieure d’ABC est faible, pour que sa probabilité postérieure (étant donné la présence des fleurs) soit élevée, il faut (1) que la probabilité antérieure de la présence des fleurs soit faible et (2) que la probabilité de la présence des fleurs s’il y a la substance ABC soit élevée. (1) sera le cas, si seule la présence d’ABC pouvait expliquer celle des fleurs F, ou s’il était très difficile d’avoir une autre explication. En revanche, si la présence de fleurs F s’explique non seulement par la présence de la substance ABC, mais aussi par celle de la substance EFG qui est elle assez répandue dans la région, alors la probabilité antérieure de la présence de F ne sera pas si faible. (2) sera le cas, si la présence d’ABC entraîne systématiquement ou presque la présence de fleurs F. En revanche, même si ABC est nécessaire pour qu’il y ait des fleurs F, mais qu’il faut aussi une autre condition (par exemple une autre substance XYZ), dont la présence est très peu probable, alors la présence d’ABC élève la probabilité de F, mais pas autant. Bref, pour que la présence de F nous donne une bonne raison de penser qu’il y a de l’ABC dans le sol, il faut que l’ABC soit une explication (presque) *suffisante* (au sens où elle ne réclame comme autres conditions que des conditions assez probables) pour la présence de F, et une explication (presque) *nécessaire* au sens où il n’y aurait guère d’explication rivale plausible de la présence de F.

Revenons alors aux miracles. Ce que l’on cherche à savoir c’est si une certaine expérience, en l’occurrence un ou plusieurs témoignages, *élève*  la probabilité d’une hypothèse, en l’occurrence la production d’un miracle. Le théorème de Bayes appliqué à l’hypothèse d’un miracle M et du témoignage en faveur du miracle t(M) a la forme suivante (simplifiée) :

Pr(M|t(M)) = Pr(M)xPr(t(M)|M)

Pr(t(M))

Autrement dit, selon le théorème de Bayes, le témoignage en faveur du miracle élève sa probabilité si l’hypothèse du miracle rend plus probable le témoignage (qu’il ne l’est sans l’hypothèse). Cette probabilité sera également fonction de la probabilité antérieure du miracle, supposée très faible, et du caractère *improbable* a priori de l’occurrence du témoignage.

On peut se donner une idée plus précise de ce que demande la crédibilité du miracle en établissant le rapport entre la probabilité que le miracle ait lieu si le témoignage est donné (Pr(M|t(M)) et celle qu’il n’ait pas lieu si le témoignage est donné (Pr(~M|t(M)). Le miracle sera crédible (probable) si ce rapport est supérieur à 1, et sinon il ne devra pas être cru. La probabilité que le miracle n’ait pas lieu si le témoignage est donné peut également être exprimée selon théorème de Bayes

Pr(~M|t(M)) = Pr(~M)xPr(t(M)|~M)

Pr(t(M))

Et le rapport des deux probabilités donnera (moyennant une élimination élémentaire du facteur Pr(t(M))

Pr(M|t(M)) = Pr(M) x Pr(t(M)|M)

Pr(~M|t(M)) Pr(~M) Pr(t(M)|~M)

Donc ce rapport sera supérieur à 1 si la multiplication des deux facteurs à droite est supérieur à 1. Comme on suppose (par définition) que le miracle est intrinsèquement très peu probable, le premier facteur sera très petit (Pr(M) est très faible et Pr(~M) très élevée). Pour que le produit soit supérieur à 1, il faudra que le deuxième facteur soit très élevé, ce qui sera d’autant plus le cas que le miracle rend probable l’occurrence du témoignage (Pr(t(M)|M) est très grande) alors que la non-occurrence du miracle rendrait très improbable l’occurrence du témoignage (Pr(t(M)|~M) est très petite).

Evidemment, la perspective humienne est que, non seulement la probabilité antérieure du miracle est très faible (et donc très difficile à surmonter par l’autre facteur[[25]](#footnote-25)), mais qu’en plus la probabilité des témoignages en faveur des miracles est relativement élevée (passion du merveilleux, etc.) de sorte qu’ils peuvent facilement être produits alors qu’ils sont faux. Mais ces considérations relèvent de la deuxième partie de l’enquête sur les miracles. Si l’on s’en tient à la première partie, et à la maxime générale, elle ne semble pas dire plus que ce qui vient d’être énoncé au moyen du théorème de Bayes. A moins que Hume ne veuille dire qu’on ne peut jamais être dans une situation où le rapport de gauche (de l’équation) est supérieur à 1, parce que le deuxième facteur à droite (de l’équation) ne pourra jamais compenser (surmonter) le premier facteur, et cela sans considérer les arguments décridibilisant de la deuxième partie, mais simplement parce que l’improbabilité (‘incrédibilité’, ‘miraculosité’) d’un miracle est toujours au moins aussi forte que l’improbabilité (‘incrédibilité’, ‘miraculosité’) que le témoignage soit faux (formule qui revient à dire que le premier facteur ne peut pas être compensé par le second facteur).

Sur ce point, la majorité des commentateurs s’accordent à dire que, si c’était bien la pensée de Hume, elle serait erronée, et que le principe général de la maxime, ou sa reprise au moyen du théorème de Bayes, permet au contraire de montrer qu’une situation où il est rationnel de croire que le miracle a eu lieu sur la base du témoignage est possible. Deux considérations doivent notamment être prises en compte.

1) Théologie naturelle et téléologie. Nous avons vu que la probabilité d’une hypothèse était en fait toujours conditionnelle, puisqu’elle présupposait des connaissances d’arrière-plan (qu’on ne mentionne pas si elles sont communes à tous). Quand on parle de la probabilité antérieure du miracle, il s’agit de faire abstraction du témoignage, mais pas de toute connaissance d’arrière-plan (k). Or, il ne va pas de soi que la croyance qu’il y a un Dieu (le théisme) doive être tenue à l’écart quand on évalue la probabilité, même intrinsèque, d’un miracle. Autrement dit, la croyance qu’il y a un Dieu (qu’on peut noter D) pourrait faire partie de la connaissance d’arrière-plan (k). Au minimum, il faudrait dire que le théiste n’attribuera pas la même probabilité (improbabilité) au miracle que l’athée ou l’agnostique. Surtout s’il s’agit d’un miracle que Dieu (ou un dieu) aurait de bonnes raisons de produire (on en voit le but, la justification : se faire connaître, aider une personne en difficulté, etc.). Et quand bien même on voudrait limiter toute la discussion à des interlocuteurs agnostiques, ou qui ne veulent pas mettre la croyance (ou l’incroyance) en Dieu dans la balance, il faudrait encore envisager la *possibilité* qu’il y ait un Dieu, et qu’il ait des raisons de produire des miracles ou ce miracle en particulier. Cela pourrait encore modifier (en l’augmentant) la probabilité que l’on donne au miracle. Ces considérations pourraient non seulement conduire à élever la probabilité du miracle étant donné le témoignage (car Pr(M|t(M) & D) > Pr(M|t(M) & ~D) mais aussi à élever la probabilité du miracle sans considération du témoignage (Pr(M|D) > Pr(M|~D)). Si c’est le cas, le rapport Pr(M)/Pr(~M), qui est en fait Pr(M|k)/Pr(~M|k), pourrait ne pas être si petit que le veut Hume (si k inclut D ou même la possibilité que D soit vrai). On comprendrait alors qu’il puisse être plus facilement surmonté par l’autre facteur, si l’on estime que l’hypothèse que le miracle ait eu lieu explique très bien l’occurrence du témoignage alors que cette occurrence serait très difficile à expliquer si le miracle n’a pas eu lieu (direction argumentative souvent suivie dans les discussions apologétiques en faveur de la résurrection du Christ).

Cette stratégie, envisagée par plusieurs auteurs, et aujourd’hui, par Richard Swinburne, pourrait être contrée en ajoutant aux arguments de Hume contre les miracles, ceux qu’il produit contre l’hypothèse théiste ou du moins contre la rationalité de souscrire à l’hypothèse, comme de penser que l’on peut envisager avec plausibilité les desseins de Dieu. Il faudrait alors renforcer l’*Enquête* par les arguments des *Dialogues* (ou l’enquête sur les miracles, par celle sur la providence). Rien ne montre que cette stratégie de réponse soit dans l’esprit de Hume, dont la critique des miracles paraît plutôt autonome. Il faudrait alors reconnaître au moins une faiblesse dans sa discussion, un manque, qui n’est pas forcément dramatique, mais qui interdit de se contenter de son analyse telle qu’elle est : la téléologie possible de la part d’une divinité possible (le point est bien reconnu par Millican au §18) affaiblit l’argument ou doit être récusée par un autre argument.

2) Multiplication des témoins. Comme l’ont remarqué plusieurs philosophes ayant réfléchi sur la maxime de Hume, depuis au moins Charles Babbage, chaque témoin indépendant constitue un facteur indépendant en faveur du miracle, et contribue à élever la probabilité de l’hypothèse (que le miracle a eu lieu). Le théorème de Bayes permet de noter ainsi l’addition des témoins (chaque Ti est un témoin) :

Pr(M|T1(M) & ... & Tn(M)) = Pr(M) × Pr(T1(M) & ... & Tn(M)|M)

Pr(~M|T1(M) & ... & Tn(M)) Pr(~M) Pr(T1(M) & ... & Tn(M)|~M)

Si les témoins sont *indépendants* (quant au fait que le miracle a eu lieu), le deuxième facteur (à droite de l’équation) peut être réduit ainsi

Pr(T1(M)|M) × ... × Pr(Tn(M)|M)

Pr(T1(M)|~M) Pr(Tn(M)|~M)

Et si les témoins ont le même *poids*, ce sera plus simplement

[Pr(T1(M)|M) /Pr(T1(M)|~M)]n

Il est alors clair que si le rapport est supérieur à 1, on peut (toujours) trouver un *n* (nombre de témoins) suffisamment grand pour que, quelle que soit la grandeur du rapport Pr(M)/Pr(~M), le produit de ce rapport par celui ci-dessus soit supérieur à 1. Comme le disait Babbage :

[I]f independent witnesses can be found, who speak truth more frequently than falsehood, it is ALWAYS possible to assign a number of independent witnesses, the improbability of the falsehood of whose concurring testimony shall be greater than the improbability of the alleged miracle. (Babbage 1837: 202, emphasis original; cf. Holder 1998 and Earman 2000)

Cela ne veut évidemment pas dire qu’il y a des miracles dont le nombre de témoins est suffisamment grand pour qu’ils soient crédibles (ou même pour qu’ils doivent être crus), mais que cette possiblité ne peut pas être exclue en principe (y compris abstraction faite de toute théologie naturelle et de toute hypothèse sur les desseins possibles d’une divinité possible). Autrement dit, pour défendre l’irrationalité de la croyance aux miracles, Hume ne peut pas s’appuyer sur la seule maxime atteinte dans la première partie de l’*Enquête*. Et s’il la complète par des raisons montrant qu’*en fait* les récits de miracles sont peu crédibles (parce qu’ils sont assez probables, même s’ils sont faux), cela ne peut être qu’une défense de l’irrationalité *en fait* de la croyance aux miracles qui sont rapportés, pas de son irrationalité *en droit*.

Reste alors à savoir si les raisons apportées par Hume dans la deuxième partie de l’Enquête montrent bien qu’aucun récit d’aucun miracle apporté jusqu’ici (dans l’histoire de l’humanité) n’est crédible.

1. Pour Locke voir *Essai sur l’entendement humain* IV, ch 15 « De la probabilité » (intéressant en lui-même pour comprendre le texte de Hume) : « Comme la démonstration consiste à montrer la convenance ou la disconvenance de deux idées, par l’intervention d’une ou de plusieurs preuves qui ont entre elles une liaison constante; de même la Probabilité n’est autre chose que l’apparence d’une telle convenance ou disconvenance par l’intervention de preuves dont la connexion n’est point constante et immuable, ou du moins n’est pas aperçue comme telle, mais est ou paraît être ainsi le plus souvent, et suffit pour porter l’esprit à juger que la proposition est vraie ou fausse plutôt que le contraire » [↑](#footnote-ref-1)
2. Dans l’introduction de l’*Analogy of Religion*, Butler écrit : "PROBABLE evidence is essentially distinguished from demonstrative by this, that it admits of degrees; and of all variety of them, from the highest moral certainty, to the very lowest presumption.” [↑](#footnote-ref-2)
3. Au début de l’enquête sur les miracles, Hume utilise la formule plus générale selon laquelle c’est l’expérience qui est le guide de la vie. Mais la probabilité telle que l’entendent ces auteurs est justement le fruit de l’expérience, et désigne le poids que peuvent et doivent jouer les expériences (sensible, témoignage) en faveur de telle hypothèse. [↑](#footnote-ref-3)
4. Il y a tout un débat récent sur la distinction entre croyance et acceptation, largement influencée par le livre de J.L. Cohen, *Belief and Acceptance*, Oxford, 1989. Ceux que cela intéresse pourront consulter le collectif dirigé par L. Jaffro, *Croit-on comme on veut ?* Vrin 2013, consacré à une sorte d’histoire de cette distinction depuis l’Antiquité. Je peux accepter une propositon que je ne crois pas, par exemple en pariant que p et en misant sur p ; ou en supposant que p pour les besoins d’un raisonnement, etc. [↑](#footnote-ref-4)
5. Ministre anglican non conformiste, il a eu un rôle pionier dans la science des assurances. Par ailleurs ses *Observations on the Nature of Civil Liberty* (1776) auraient influencé certains leaders de la révolution américaine. [↑](#footnote-ref-5)
6. Thomas Bayes était également ministre anglican, non conformiste et touche à tout. Membre de la Royal Society, il avait écrit un pamphlet anonyme défendant le calcul de Newton contre les critiques de Berkeley. Il légua une maigre somme à Price à sa mort, ce qui témoigne d’un certain lien entre les deux, et c’est Price qui édita son *Essay* à titre posthume, avec un appendice de sa plume. Qui plus est, la critique que Price fait de Hume s’appuie sur les travaux de Bayes. [↑](#footnote-ref-6)
7. Ce qui est à son tour compris au sens de la formule (E/M) ci-dessus [↑](#footnote-ref-7)
8. On pourrait aussi dire qu’au lieu de comparer deux types de témoignages (témoins humains et expérience) on compare deux expériences (celle des témoignages, et celles des observations en faveur de la loi). C’est plutôt ainsi que s’exprime Hume, mais on perdrait l’idée du procès avec les témoins appelés à la barre. [↑](#footnote-ref-8)
9. Pour être plus précis,il faudrait donc dire : P(M|t(M)&k) > 0.5 si et seulement si Pr(M|k) > Pr(~M|t(M)&k). La formule exprime l’idée suivante : pour que la probabilité d’un miracle faisant l’objet d’un témoignage le rende crédible (plus probable que sa négation ou plus probable que 0.5) il faut et suffit que la probabilité intrinsèque du miracle soit supérieure à la probabilité que le miracle n’ait pas eu lieu alors qu’il fait l’objet d’un témoignage [↑](#footnote-ref-9)
10. pour que la probabilité d’un miracle faisant l’objet d’un témoignage le rende crédible (plus probable que sa négation ou plus probable que 0.5) il faut et suffit que la probabilité intrinsèque du miracle soit supérieure à la probabilité que le témoignage soit donné alors que le miracle n’a pas eu lieu. [↑](#footnote-ref-10)
11. Ce qui serait trivial serait de dire que Pr(M)>0.5 ssi Pr(M)>Pr(~M), mais ce n’est pas cela qui est dit. [↑](#footnote-ref-11)
12. George Campbell: What then shall be said of the conclusion which he [Hume] gives as the sum quintessence of the first part of the *Essay*? The best thing, for aught I know, that can be said is, that it contains a most certain truth, though at the same time the least significant, that ever perhaps was ushered into the world with such solemnity. . . . If any reader think himself instructed by this discovery, I should be loth to envy him the pleasure he may derive from it. (CDM, 55; 193) [↑](#footnote-ref-12)
13. On trouve notamment cette objection dès les lecture de Hume par Butler et Price. [↑](#footnote-ref-13)
14. Voir les travaux de Earman, Hajek et al. [↑](#footnote-ref-14)
15. Le témoignage est vrai ou faux. La probabilité qu’il soit vrai ou faux est de 1 (certitude absolue). Si la probabilité qu’il soit faux est *f*, la probabilité qu’il soit vrai est la probabilité qu’il soit vrai ou faux moins la probabilité qu’il soit faux (1-*f*). On peut aussi dire que la probabilité de la disjonction est la somme des probabilités des disjoints : Pr(p ou q) = Pr(p) + Pr(q). Dans notre cas, Pr (p ou non-p) = Pr(p) + Pr(non-p), et comme Pr(p ou non p) = 1, et que Pr(nonp) = Pr(p est faux)= f, alors Pr(p) = 1-*f.* [↑](#footnote-ref-15)
16. sachant que la probabilité de la conjonction (p&q) est égale au produit de la probabilité de chaque conjoint, si ceux-ci sont indépendants : (Pr(p&q) = Pr(p) x Pr(q))  [↑](#footnote-ref-16)
17. C’est un résultat acquis en psychologie du raisonnement probabiliste que nous évaluons mal les probabilités *à première vue*, en commettant de nombreux sophismes. Ce qui suit en illustre un assez naturel : l’ignorance de la *base rate* (fréquence de base) [↑](#footnote-ref-17)
18. Plus précisément on passe de la probabilité (avant le test) de 1 sur un million, à celle (après le test) de 1 sur 1002. Sur une population d’un milliard de personnes, 1000 ont la condition G et 999 999 000 ne l’ont pas. Si tous passaient le test, on devrait attendre 999 *Vrais Positifs* et 999 999 *Faux Positifs.* Ceux dont le test est positif seraient donc 1 000 998, alors que le test ne serait vrai que pour 999. La probabilité d’avoir la condition G serait de 999/1000998 = 1/1002 (0,000998) [↑](#footnote-ref-18)
19. Hajek répond se contente de cette réponse [↑](#footnote-ref-19)
20. Millican donne l’exemple de deux témoignages erronés, l’un qui consiste à prendre un tronc flottant sur le Loch Ness pour un monstre, et l’autre qui consiste à prendre le monstre du Loch Ness pour un tronc flottant. Il n’y aucune raison de penser que la fiabilité (ou le manque de fiabilité) du témoignage est la même dans les deux cas. [↑](#footnote-ref-20)
21. Je dirais (1-10-6)x(10-6), soit 10-6 – 10-12, donc environ 10-6 mais je ne suis pas sûr… [↑](#footnote-ref-21)
22. Autrement dit, la formule du calcul du *Faux Positif* donné plus haut

    Pr(~*M* & *t*(*M*)) = Pr(~*M*) x Pr(le témoignage est faux) = (1-*m*).f

    n’est généralement pas valide, car il ne faut pas se contenter du facteur « Pr(le témoignage est faux) », mais considérer au contraire quelque chose comme « Pr(le témoignage dit que M a eu lieu et est faux) ». Ainsi calculé, le *Faux Positif*  qui pouvait sembler assez probable (et plus probable que le *Vrai Positif*), peut se révèler en fait bien moins probable que le *Vrai Positif,* ce qui justifierait alors que l’on accorde du crédit au témoignage. [↑](#footnote-ref-22)
23. Dire que ce sont des énoncés *négatifs* et que les lois sont formulées par des énoncés *positifs* serait une fausse sortie : dire que l’on n’a jamais vu de saumon de 3m c’est dire que tous les saumons observés faisaient moins de 3m, mesuraient au maximum 1,5 m, etc. [↑](#footnote-ref-23)
24. Pr(H|k) est souvent désignée comme la *probabilité antérieure* d’une hypothèse, tandis que Pr(H|E&k) est sa *probabilité postérieure* (fonction d’informations nouvelles: expérience, etc.). La probabilité antérieure de l’expérience Pr(E|k) mesure la *surprise* de l’observation de E, on parle aussi de *la vraisemblance antérieure* de E. Quant à la probabilité postérieure de l’expérience, Pr(E|H&k), elle mesure combien H explique E. On parle alors de *vraisemblance (likelihood)* de l’hypothèse pour l’expérience E. [↑](#footnote-ref-24)
25. J’entends par ‘surmonter’ l’idée que le produit du premier facteur (supposé très inférieur à 1) par le second serait supérieur à 1 (parce que le second facteur serait suffisamment grand) [↑](#footnote-ref-25)