HAL archives-ouvertes.fr

QUAND LA PROBABILITÉ CONDITIONNELLE CROISE LA STATISTIQUE

Léo Gerville-Réache

To cite this version:

Léo Gerville-Réache. QUAND LA PROBABILITÉ CONDITIONNELLE CROISE LA STATISTIQUE. CFIES 2015, Jan 2015, Bordeaux, France. <hal-01066880>

HAL Id: hal-01066880

https://hal.archives-ouvertes.fr/hal-01066880

Submitted on 22 Sep 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution| 4.0 International License

QUAND LA PROBABILITE CONDITIONNELLE CROISE LA STATISTIQUE

Léo Gerville-Réache[1]

[1]*Université de Bordeaux, CNRS, UMR 5251, France, leo.gerville-reache@u-bordeaux.fr*

Résumé. Dans cette communication, nous revenons sur la définition de la probabilité conditionnelle de Kolmogorov (1933) et, partant de la critique de Hajek (2003), nous développons le concept général de probabilité relative qui se décompose soit en probabilité conditionnelle soit en probabilité situationnelle. Nous illustrons ensuite notre propos sur l'un des plus délicats paradoxes de théorie de la décision de ces dernières décennies : les deux enveloppes.

Mots-clés. Probabilité conditionnelle, probabilité situationnelle, théorie de la décision.

Abstract. In this paper, we review the definition of the conditional probability of Kolmogorov (1933) and from the criticism of Hajek (2003), we develop the general concept of relative probability that decomposes either conditional probability or situational probability. We then illustrate our purpose on one of the most difficult paradoxes of decision theory in recent decades: the two envelopes paradox.

Keywords. Conditional probability, situational probability, decision theory.

1 Introduction

Kolmogorov (1933) introduit, page 6 de son ouvrage (*Foundation of the theory of probability*), la probabilité conditionnelle comme suit : *"If* $P(A) > 0$*, then the quotient*

$$P_A(B) = \frac{P(AB)}{P(A)}$$

is defined to be the conditional probability of the event B under the condition A."

NB: $P(AB)$ signifie $P(A \cap B)$ et $P_A(B)$ est indifféremment noté de nos jours: $P(B|A)$.

Hajek (2003) écrit à propos de cette définition, dans l'introduction de son papier (*What Conditional Probability Could Not Be)* : "*Call this the ratio analysis of conditional probability. It has become so entrenched that it is often referred to as the definition of conditional probability. I argue that it is not even an adequate.*"

Pour la défense de Kolmogov, celui-ci précisait clairement dans l'introduction de son ouvrage: *"The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry and Algebra. This means that after we have defined the elements to be studied and their basic relations are to be governed, all further exposition must be based exclusively on these axioms, independent of the usual concrete meaning of these elements and their relations."*

Mais alors, dans quelle mesure la définition mathématique de Kolmogorov correspond-t-elle au concept philosophique de probabilité conditionnelle? Après avoir rappelé quelques éléments de la critique d'Hajek, nous proposons une réflexion sur la nécessaire différenciation des vocables "si" et "sachant", associés, encore aujourd'hui, indifféremment au concept de probabilité conditionnelle. Nous développons le concept plus général de "probabilité relative" qui se décline soit en probabilité conditionnelle soit en probabilité situationnelle. Cette réflexion nous permet, par exemple, de proposer un regard original sur le très débattu paradoxe des deux enveloppes.

2 Hajek : une critique de la probabilité conditionnelle

Parmi les critiques d'Hajek (2003) sur la définition "par ratio" de Kolmogorov, se trouve la question de la probabilité conditionnelle à une valeur particulière d'une variable aléatoire continue. Ce problème, soulevé en son temps par Borel, avait conduit Kolomgorov, à consacrer, dans son ouvrage, un paragraphe intitulé "*Explanation of Borel Paradox*" dans lequel il écrit "*This shows that the concept of a conditional probability with regard to an isolate given hypothesis whose probability equal 0 is inadmissible*".

Hajek considère que c'est la définition par "ratio" de Kolmogorov qui génère cette inadmissible inadmissibilité des probabilités conditionnelles aux ensembles non-vides de mesure nulle. Il donne alors quelques exemples simples qui éclairent son point de vue.

"Continuous random variables give rise to non-trivial events of probability 0. Associated with such a random variable is a probability density function (think, for example, of the 'bell'-shaped normal distribution). The probability of a continuous random variable X, taking a particular value x, equals 0, being the integral from x to x of X's density function. Yet it would seem that various probabilities, conditional on this event, are well-defined, for example:

$$P(X = x, \text{ given } X = x) = 1.$$

Indeed, here and throughout the paper, we hold this truth to be self-evident: the conditional probability of any (non-empty) proposition, given itself, is 1. I think that this is about as basic a fact about conditional probability as there can be, and I would consider giving it up to be a desperate last resort. Some others are comparably basic; in the case before us they include:

$$P(X \neq x, \text{ given } X = x)=0,$$

[...] Less trivially, various conditional probability assignments based on judgments of independence are compelling – for instance:

$$P(\text{this coin lands heads, given } X = x) = 1/2."$$

Les remarques d'Hajek comme celle qui veut que *P(this coin lands heads, given X = x) = 1/2*, semblent particulièrement pertinentes. Il est effectivement difficile de comprendre pourquoi une telle probabilité ne serait pas mathématiquement définissable et donc non-quantifiable. Dans la section suivante, nous tentons de réconcilier le mathématicien et le statisticien-philosophe par une approche duale du concept plus général de *probabilité relative*.

3 Probabilité conditionnelle ou situationnelle?

La définition de Kolmogorov parle de "*the conditional probability of the event B under the condition A"*. Il s'agit littéralement d'une "probabilité conditionnelle" : si A alors on aurait $P(B)$=... Dans les propos d'Hajek, on trouve le terme "*given"*. S'agirait-il d'une "probabilité situationnelle" : sachant A, on a $P(B)$=... La distinction semblera plus qu'anodine à la plupart des probabilistes et des statisticiens qui utilisent indifféremment les deux vocables $P(B$ "si" $A)$ et $P(B$ "sachant" $A)$ via l'écriture $P(B|A)$. Pourtant, la pertinence de la critique d'Hajek passe peut-être par une distinction essentielle, entre une probabilité basée sur une hypothèse et une probabilité basée sur une réalité.

En logique il existe une différence indéniable entre (1) Si A alors B et (2) Sachant A alors B. En effet, de (1) on ne peut rien de conclure sur la véracité de A ou de B alors que de (2) on conclut que A et B sont vraies (A n'est plus une hypothèse mais une réalité et donc B l'est également). Si l'on envisage que le concept de probabilité de B relativement à A: $P(B|A)$, diffère selon qu'il s'agit d'une probabilité conditionnelle: $P(B$ si $A)$, ou d'une probabilité situationnelle: $P(B$ sachant $A)$, les valeurs de $P(B$ si $A)$ et $P(B$ sachant $A)$ peuvent-elles différer dans des cas très particuliers?

La différence d'évaluation entre probabilités conditionnelles et situationnelles se pose pour les "événements" possibles de probabilités nulles. En effet, si on considère par exemple une variable aléatoire uniforme continue sur [0,1], notée X, et la réalisation présentement : X=0,17. Alors on doit être en mesure de parler et de quantifier la probabilité situationnelle d'un événement *B*, sachant *X*=0,17. Mais on ne peut pas considérer ou quantifier la probabilité conditionnelle d'un événement *B*, sous l'hypothèse *X*=0,17. Dans le premier cas, *X*=0,17 bien que de probabilité nulle, est un événement réalisé et on peut donc le considérer (c'est notre interprétation de la critique d'Hajek). Dans l'autre cas, *X*=0,17 est une hypothèse inadmissible (comme le précise Kolmogorov).

Aussi, en reprenant l'exemple d'Hajek, il nous semble que l'on ne puisse pas définir la probabilité conditionnelle: *P*("la pièce tombe sur face" si *X*=0,17) mais que l'on puisse définir et quantifier la probabilité situationnelle : *P*("la pièce tombe sur face" sachant *X*=0,17), qui sera de 1/2.

Hajek, critiquant la définition de Kolmogorov, ne distingue pas les concepts de probabilité conditionnelle et de probabilité situationnelle. Il n'y a, de notre point de vue, aucun reproche à faire à la définition par ratio de Kolmogorov qui est très clair dans les termes : $P_A(B)$ est la probabilité conditionnelle de *B* sous la condition *A*. On peut également noter que ce n'est aucunement à la théorie mathématique des probabilités de définir le concept de probabilité situationnelle. En effet cette définition est du ressort de la science statistique dans la mesure où celle-ci traite de probabilités relatives à la *réalisations* de variables aléatoires.

Le mathématicien et le statisticien-philosophe ne traite donc pas du même concept et il est donc normal que dans certains cas, liés par exemple au fait que dans la mathématique de Kolmogorov, certains événements de probabilité nulle sont non-vides, les valeurs des deux concepts divergent. Gerville-Réache et Couallier (2013) proposaient le concept "statistique" de probabilité conditionnelle et ses conséquences sur l'analyse de certains paradoxes. Dans la section qui suit, nous utilisons cette dualité du concept de probabilité relative pour analyser un paradoxe, non résolu, qui anime de longue date, mathématiciens, statisticiens, logiciens et philosophes.

4 Le paradoxe des deux enveloppes

Les énoncés diffèrent sensiblement selon les auteurs. Nous présentons ici celui d'Emery (2001), publié dans la revue de l'IREM de Strasbourg.

Je mets sous enveloppe deux chèques à votre nom, le montant de l'un étant le double de l'autre. Vous tirez au sort l'une des deux enveloppes, vous l'ouvrez, puis vous devez choisir celui des deux chèques que vous allez toucher. Vous ne connaissez donc que le montant m de l'un des chèques, et vous savez que l'autre vaut le double ou la moitié de m. Quel choix faites-vous?

La présentation faite de ce problème est souvent accompagnée de la discussion suivante, en deux points :

*1) Appelons C le chèque de valeur m que vous avez trouvé en ouvrant l'enveloppe, et I, le chèque de montant inconnu 2m ou m/2 qui est dans l'autre enveloppe. Si vous choisissez I, vous avez une chance sur deux d'encaisser le montant 2m et une chance sur deux d'encaisser m/2. Votre espérance de gain est alors de 1/2*2m+1/2*m/2, c'est à dire 5/4*m. Cette quantité est toujours supérieure à m; ainsi, en choisissant I plutôt que C, vous augmentez votre espérance de gain, ceci quelle que soit la valeur de m lue.*

2) La conclusion ci-dessus est manifestement absurde. La stratégie préconisée au point précédent - tirer au sort l'une des enveloppes, l'ouvrir, puis, quel que soit son montant (donc, en fait, sans même avoir à l'ouvrir), choisir systématiquement l'autre enveloppe - n'est qu'un procédé un peu détourné pour tirer au sort, avec chances égales, celle des enveloppes qui sera choisie. Mais c'est aussi ce que fait la stratégie qui consiste à toujours choisir C; la stratégie proposée en 1, systématiquement choisir I, n'est donc finalement pas meilleure que de choisir systématiquement C.

Ces deux arguments démontrent une chose et son contraire. Où est le vrai? A quel endroit y a-t-il une erreur?

La littérature sur ce paradoxe est très abondante, plus de 100 articles depuis les années 80 (voir Gerville-Réache (2014) pour un point de vue sur les différentes propositions de résolution). Dans l'énoncé qui précède, il existe, selon nous, un abus dans le raisonnement 1 et un dans le raisonnement 2. Dans le raisonnement 1, il est écrit: " *Si vous choisissez I, vous avez une chance sur deux d'encaisser le montant 2m et une chance sur deux d'encaisser m/2*". Vous n'avez pas *une chance sur deux,* ce "1/2" est précisément le degré de croyance, via le principe de raison insuffisante de Laplace (voir De Scheemaekere X. (2012)), qu'il semble épistémiquement rationnel d'attribuer aux deux possibilités 2*m* et *m*/2 des montants de l'autre enveloppe, dans l'état de connaissance dans lequel vous êtes après ouverture. Dans le raisonnement 2, il est écrit "*tirer au sort l'une des enveloppes, l'ouvrir, puis, quel que soit son montant (donc, en fait, sans même avoir à l'ouvrir)..*. Il est abusif de dire qu'un raisonnement construit après ouverture de l'enveloppe, qui ne dépend pas du montant découvert, est équivalent à faire le raisonnement sans même l'ouvrir. Aucun des deux raisonnements n'est ici correct! Aussi, reprenons l'analyse du problème. Il est clair que quelle que soit l'enveloppe que vous avez en main, en changeant :

- soit vous gagnez et vous repartez avec le double du montant que vous avez en main.
- soit vous perdez, vous repartez avec la moitié de la somme que vous avez en main.

Pourtant cette conséquence ne se traduit pas de la même façon si vous connaissez ou pas le montant de l'enveloppe que vous avez en main.

- Avant d'ouvrir l'enveloppe en votre possession, vous pouvez seulement dire que l'une des enveloppes contient un certain montant inconnu *N* et l'autre 2*N* et ajouter :

  - Si en changeant je gagne, alors c'est que j'avais en main le plus petit montant inconnu *N* et que l'autre contient 2*N*. Je gagne donc *N* et repars avec le double du montant que j'avais.
  - Si en changeant je perds, alors c'est que j'avais en main le plus grand montant inconnu 2*N* et que l'autre contient *N*. Je perds donc *N* et repars avec la moitié du montant que j'avais.

  N'ayant aucune raison de croire que vous avez plus de chance d'avoir en main le plus petit que le plus gros montant, vous n'avez aucune raison de changer d'enveloppe. En effet, si vous changez, la somme inconnue que vous pourriez gagner est égale à la somme inconnue que vous pourriez perdre; ces deux possibilités étant épistémiquement équiprobables.

- Après ouverture de l'enveloppe en main, et découvrant le montant *m*, vous pouvez dire :

  - Si en changeant je gagne, alors c'est que *m*, le montant connu que j'avais en main, était le plus petit montant et que l'autre contient 2*m*. Je gagne donc *m* et repars avec le double du montant que j'avais.
  - Si en changeant je perds, alors c'est que *m*, le montant connu que j'avais en main, était le plus gros montant et que l'autre contient *m*/2. Je perds donc *m* /2 et repars avec la moitié du montant que j'avais.

  N'ayant aucune raison de croire que vous avez plus de chance avoir en main le plus petit que le plus gros montant, vous avez maintenant une raison de changer d'enveloppe. En effet, si vous changez, la somme connue que vous pourriez gagner est égale au double de la somme connue que vous pourriez perdre; ces deux possibilités étant épistémiquement équiprobables.

Il y ici a une différence fondamentale entre le fait de prendre connaissance ou pas du montant de l'enveloppe en main. La source du paradoxe réside dans l'amalgame de ces deux états de connaissance aux conséquences distinctes. Ce qui est redoutablement troublant, c'est que la décision qui consiste à changer sachant le montant *m* de l'enveloppe en main ne dépende pas de la valeur de *m* mais seulement de l'état de connaissance et de la décision qu'il engendre. En effet, si quelle que soit la valeur découverte, vous avez intérêt de changer alors pourquoi auriez-vous besoin de connaitre cette valeur pour avoir intérêt à changer?

Dietrich et List (2005) propose une analyse précise de cette question. Selon eux, la contradiction apparente proviendrait d'un "axiome tacite" de la théorie des jeux : *"Event-wise dominance principle: Let P be a partition of the set of all possible states of the world into non-empty events. For any two lotteries L_1 and L_2, if you strictly prefer L_1 to L_2 conditional on observing event E for* *every E in P, then you strictly prefer L_1 to L_2 unconditionally."* Principe que les auteurs rejettent au profit du suivant : *"State-wise dominance principle: For any two lotteries L_1 and L_2, if you strictly prefer L_1 to L_2 conditional on every possible state of the world, then you strictly prefer L_1 to L_2 unconditionally."*. Cette analyse axiomatique ferait disparaitre l'apparente contradiction du paradoxe des deux enveloppes. Pour autant, nous lui préférons l'analyse basée sur la différentiation entre un raisonnement conditionnel et un raisonnement situationnel.

Le raisonnement conditionnel (avant ouverture de l'enveloppe) consiste à dire "Si il y a *m* dans mon enveloppe, alors...". Ce raisonnement nécessitera alors l'utilisation d'une probabilité conditionnelle pour les deux montants conséquemment possibles de l'autre enveloppe. Mais comme nous l'avons vue précédemment, la définition de la probabilité conditionnelle nécessite de définir $P(m)$, et cela quel que soit *m*. Or, n'ayant aucune information sur les montants possibles des deux enveloppes (en particulier, pas de montant maximum connu), la seule loi de probabilité "non-informative" pour *m* serait une loi uniforme sur $\mathbf{R}_+^*$. Mais, nous savons bien qu'une telle loi est inadmissible. La modélisation de notre ignorance nous contraindrait, quoi qu'il arrive, à considérer comme possible que $P(m)=0$. Aussi, aucun raisonnement conditionnel n'est totalement admissible et c'est la symétrie des enveloppes qui impose l'indifférence au changement d'enveloppe.

En revanche, connaissant le montant de l'enveloppe en main, on est en mesure de construire un raisonnement situationnel. Celui-ci ne nécessite pas de probabiliser l'espace des valeurs possibles de *m*. En effet, la probabilité situationnelle que l'enveloppe *C* contienne *m* sachant que l'enveloppe *C* contient *m* vaut incontestablement 1. L'espace des possibles pour l'autre enveloppe (*I*), ici réduit aux deux valeurs connues $2m$ et $m/2$, est alors probabilisable par une loi équiprobable qui correspond précisément à notre nouvel état de connaissance. Aussi, on ne peut ni croire ni écrire, $\forall m$ de l'ensemble des possibles :

$$P(I = 2m\ si\ C = m) = P(I = m/2\ si\ C = m) = 1/2.$$

Mais il est cohérent et non contradictoire de croire et d'écrire, $\forall m$ de l'ensemble des réalisés :

$$P(I = 2m\ sachant\ C = m) = P(I = m/2\ sachant\ C = m) = 1/2.$$

Pour résumer : avant l'ouverture de l'enveloppe, le raisonnement qui conduit à conclure que l'on aurait intérêt à changer est inadmissible. En effet, il s'agit d'un raisonnement hypothétique, produisant des probabilités épistémiques conditionnelles en contradiction avec la théorie mathématique des probabilités de Kolmogorov et sa définition de la probabilité conditionnelle. En revanche, après ouverture de l'enveloppe, il s'agit d'un raisonnement basé sur la réalité connue du montant de l'enveloppe en main, produisant des probabilités épistémiques situationnelles, certes non définies par Kolmogorov mais conforme à notre interprétation de la critique d'Hajek.

Cependant, une fois le montant de *C* connu, je peux toujours me dire "si j'avais choisi au départ *I*, après l'ouverture de *I*, j'aurais attribué d'autres probabilités situationnelles et j'aurais changé pour l'enveloppe *C*". Mais en l'état, j'ai choisi et ouvert *C* et c'est maintenant que j'ai une décision à prendre. C'est la comparaison "mentale" avec la décision que j'aurais prise si j'avais choisi et ouvert *I* qui génère alors un dilemme du type "OUI... MAIS... SI... ALORS... DONC... NON.".

Dans ce dilemme, il semble que la théorie des probabilités et/ou la théorie de la décision rationnelle nous laissent sans réponse, comme dans le dilemme des prisonniers. Ross (2010) affirme qu'il existe des situations où la décision d'un agent ne peut être pleinement rationnelle. Aussi, une fois l'enveloppe en main ouverte, nous trouvons-nous devant un de ces possibles dilemmes rationnels?

Cozic (2007) précise que la théorie de la décision formalise l'intuition pré-théorique suivante : *"Le choix d'une action par un agent est rationnel si, étant donné ce qu'il croit et étant donné ce qu'il peut choisir, l'action qu'il choisit est celle dont les conséquences satisfont le mieux ses désirs"*. Serait-ce la théorie de la décision rationnelle et non son concept pré-théorique qui rendrait irrationnel le choix de changer d'enveloppe, une fois le montant de l'enveloppe en main connu?

5 Discussion

La théorie des probabilités de Kolmogorov a axiomatisé la mathématique des probabilités en introduisant l'infini via la sigma-additivité. Une conséquence de cet axiome est que la possibilité de réalisation d'un évènement de probabilité nulle est concevable. Le dilemme est que cet axiome engendre un autre paradoxe pointé par Borel puis Hajek. Il traite, selon nous, de la probabilité relative à une valeur hypothétique d'une variable aléatoire continue (i.e. probabilité conditionnelle). Kolmogorov convient alors que le concept de probabilité conditionnelle à un ensemble non-vide de probabilité nulle est inadmissible.

Nous n'avons trouvé aucun paradoxe dans la définition de Kolmogorov. Celui-ci définit le concept probabiliste de probabilité conditionnelle et pas le concept statistique de probabilité situationnelle dont parle, selon notre interprétation, Hajek. Ce dernier produit des remarques parfaitement fondées mais raisonne en réalité d'un point de vue statistique. Nous pensons donc que peuvent coexister deux variantes du concept de probabilité relative, l'une conditionnelle et l'autre situationnelle. Le vocable se devrait alors d'être précisé. On parlerait d'une part de $P(B$ si A) et d'autre part de $P(B$ sachant $A)$ ou encore de $P(B$ connaissant $A)$. Dans la plupart des problèmes statistiques, cette distinction sera sans effet et semblera inutile. Cependant, certains problèmes retors souffrent possiblement de cet amalgame ; Delabre et Gerville-Réache (2014) reviennent, pour le paradoxe de la Belle au bois dormant, sur les concepts de "révision" et de "mise à jour" d'une probabilité.

Si l'introduction de la sigma-additivité dans l'axiomatique de Kolmogorov et la possibilité de construire mathématiquement des variables aléatoires continues constituent un point de tension entre la théorie de probabilités et la statistique, ne devons-nous pas prendre un grand soin à distinguer la dualité de cette "probabilité relative"? Borel écrivait en 1924 : "*Des probabilités évaluées à des moments différents, sur la base d'informations différentes, doivent être considérées comme totalement nouvelles, et non comme des ajustements par rapport à des jugements précédents*". La statistique, traitant par essence, de *réalisations* de variables aléatoires peut-elle faire l'impasse sur le concept de probabilité situationnelle?

Bibliographie

[1] Cozic M. (2007). *Un paradoxe de la décision : l'induction rétrograde*, Ulm, Journée "L'action".

[2] De Scheemaekere X. (2012). Fondements philosophiques du concept de probabilité. *Eme (Modulaires Européennes).*

[3] Delabre L., Gerville-Réache L. (accepté). *Insaisissable Belle au bois dormant*, Philosophia Scientiae.

[4] Dietrich F., List C. (2005). *The Two-Envelope Paradox: An Axiomatic Approach*, Mind New Series, Vol. 114, No. 454 , pp. 239-248.

[5] Emery M. (2001). *Quelques phénomènes curieux en probabilités et statistiques*. *L'Ouvert*, N°104, pp. 37-47.

[6] Gerville-Réache L. (2014). *Why do we change whatever amount we found in the first envelope: the Wikipedia two envelopes problem commented*, Preprint arXiv:1402.3311, 17pp.

[7] Gerville-Réache L., Couallier V. (2013). *Tentative d'identification de paradoxes latents dans l'application des probabilités conditionnelles*, 45ème Journées de statistique, Toulouse, France 6pp.

[8] Hájek A. (2003), *What Conditional Probability Could Not Be*, Synthese, Volume 137, Issue3, pp. 273-323.

[9] Kolmogorov, A. N. (1933). Grundbegriffe der Wahrscheinlichkeitrechnung, Ergebnisse Der Mathematik; *Foundation of the theory of probability*, Chelsea Publishing Company, 1950.

[10] Ross J. (2010). *Sleeping Beauty, Countable Additivity, and Rational Dilemmas*, Philosophical Review, Vol. 119, No. 4, pp. 411-474.