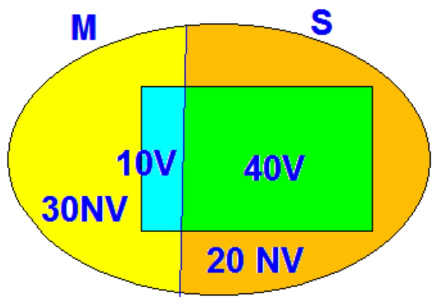
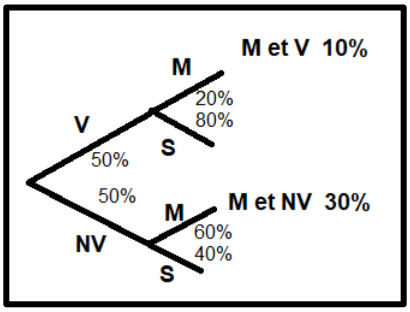
Tu me la Bayes belle.

Matheux, enseignants, chercheurs emploient la loi de Bayes dans deux types d'univers très différents.

Type 1



Supposons une population de 100 personnes où 40 sont malades (M) et 60 saines (S), 50 vaccinées (V), 50 non vaccinées (NV) et où la probabilité d'être malade quand on est vacciné (M quand V) est 20%.
L'épreuve : on tire une personne, elle est malade.
Quelle est la probabilité P pour qu'elle soit vaccinée ?
L'univers est la population.
L'énoncé suffit à déterminer la fréquence, autrement dit la probabilité des différents sous- ensembles dans l'univers :
M et NV → 30% , M et V → 10% , S et NV → 20% , S et V → 40%
On compare M et V (10%) et M et NV (30%) et on en déduit que
P =$\frac{10}{30+10} = \frac{1}{4} = 25\%$.



Si on parle de probabilité des causes c'est que le problème se pose souvent à travers un tel arbre où l'on connait les probabilités d'être malade ou sain quand on est vacciné ou malade ou sain quand on est non vacciné.
On considère que le vaccin (la cause) a une influence sur l'état de santé et connaissant la probabilité de toutes les combinaisons des deux modalités, on détermine la probabilité d'être vacciné quand on est malade comme le rapport
$\frac{P(M\ et\ V)}{P(M\ et\ V)+P(M\ et\ NV)} = \frac{10}{10+30} = 25\%$
Ici l'arbre ne donne que des probabilités mais il serait équivalent de se situer dans une population de 100 personnes où 10 seraient malades et vaccinées et 30 malades et non vaccinées. P serait alors la proportion des vaccinés parmi les malades.

Type 2

Nous nous trouvons devant 3 portes identiques.
Derrière l'une d'elles on trouve un trésor, derrière les 2 autres on ne trouve rien.
Le meneur de jeu, qui sait où se trouve le trésor, nous demande de choisir une porte puis en ouvre une autre sur une pièce vide et nous demande si au final pour localiser la porte derrière laquelle se trouve le trésor, nous voulons conserver notre choix initial où choisir la 3e porte, celle qu'il n'a pas ouverte.
Ce problème, initialement posé par un jeu télévisé américain (**paradoxe de Monty Hall**), a donné lieu à de grandes controverses pour savoir si la probabilité d'un changement opportun était $\frac{1}{2}$ ou $\frac{2}{3}$.
Pour trouver $\frac{2}{3}$ on fait appel à la populaire loi de Bayes qui est formidable pour démontrer que la probabilité reste constante même si la complexion de l'univers change. Dans ce qui suit 1M =porte 1 mauvaise, 1B = porte 1 bonne (cache le trésor).
Supposons que la porte 1 soit le choix initial (fréquence 1/3), et notons "o" la modalité "ouverte". Le meneur de jeu ouvre obligatoirement la porte 2 ou 3 si 1M (fréquence 1) et aléatoirement la porte 2 ou la porte 3 si 1B (fréquence ½).
<u>Dans l'ensemble des épreuves où la porte 1 a été choisie</u>, on peut avoir, en fonction de la configuration des 3 portes :

(**1B**,2M,3M) et **2o** (fréquence $\frac{1}{3} \text{x} \frac{1}{2} = \frac{1}{6}$)

(1B,2M,3M) et 3o (fréquence $\frac{1}{3} \text{x} \frac{1}{2} = \frac{1}{6}$)

(1M,2B,3M) et 3o (fréquence $\frac{1}{3} \text{x} 1 = \frac{2}{6}$)

(**1M**,2M,3B) et **2o** (fréquence $\frac{1}{3} \text{x} 1 = \frac{2}{6}$)

On en déduit qu'il y a 2 fois plus d'épreuves (fréquence $\frac{2}{6}$ contre fréquence $\frac{1}{6}$) où la porte 2 ayant été ouverte (2o) le trésor se situait derrière la porte 3. On a donc intérêt, sur un grand nombre d'épreuves du type de la nôtre à changer de porte.
Chose promise, chose due, la loi de Bayes a fait son travail et malgré qu'on ait ouvert une porte et dévoilé qu'elle ne cachait pas le trésor, les probabilités sont les mêmes qu'avant que le meneur de jeu n'ouvre la porte : elles sont favorables au changement de porte dans le rapport de 2 contre 1.

Quelle différence entre ces deux types d'univers ?

● **Dans les univers de type 1, tous les évènements (au sens probabiliste) existent <u>simultanément</u>. Au moment où on évalue les probabilités on se trouve devant un univers de personnes malades ou saines, vaccinées ou non vaccinées dont les sous-ensembles dénombrables constituent les évènements.**

● **Les univers de type 2 sont constitués d'ensembles d'épreuves qui reposent sur des actes (ouverture d'une porte plutôt qu'une autre, fourniture d'une carte plutôt qu'une autre) <u>mutuellement incompatibles</u>. Ils ne peuvent se produire simultanément. Autrement dit 2 évènements d'un univers de type 2 ne peuvent exister simultanément.**

D'autres exemples :

Type 1

2 usines A et B produisent chacune 50% des aspirateurs de la marque Susfor. Mais 2% des aspirateurs provenant de A sont défectueux contre 4% provenant de B. On achète un aspirateur défectueux.
Quelle est la probabilité qu'il provienne de B ?

Type 2 Les probabilités psychologiques d'Emile Borel (Théorie mathématique du bridge à la portée de tous).
On dispose de 6 cartes

♠2	♠3	♠4	♠5	♥V	♥R

On les distribue aléatoirement, face cachée, à deux joueurs appelés Est et Ouest. Chaque joueur a en sa possession un ensemble de 3 cartes qu'on appelle "main".
On peut calculer qu'il existe 20 combinaisons possibles pour la main d'Est dont 4 contiennent RV. La probabilité pour qu'Est ait RV est donc 4/20 = **20%**.
Demandons maintenant à Ouest et à Est de montrer chacun un pique de sa main. Est montre le ♠2 et Ouest le ♠3.

♠2	?	?
♠3	?	?

Quelle est la probabilité pour qu'Est ait RV ?
Du point de vue des mains possibles Est ne pouvait avoir en début de coup que
245, 24V, 24R, 25V, 25R, 2**RV**
Donc la probabilité pour qu'il ait RV est passée de **20%** à 1/6 soit **16,7%.**

Maintenant, laissons Borel nous expliquer pourquoi cette probabilité n'est pas la bonne :
« La faute de raisonnement (de l'évaluation précédente) provient de ce qu'on tient seulement compte de la probabilité d'entrée en jeu de la cause : Est recevra RV2, 5 fois sur 100. On oublie de tenir compte de la probabilité que, la cause étant entrée en jeu, c'est-à-dire Est ayant reçu RV2, Est et Ouest jouent tous deux comme ils l'ont fait. Voici maintenant le raisonnement correct qui consiste à appliquer la formule de Bayes. »
Dans la suite Borel nous explique qu'on va calculer une probabilité différente selon qu'Est et Ouest montrent leurs piques aléatoirement, ou qu'ils montrent le plus petit, ou qu'ils adoptent une autre stratégie quand ils ont le choix.

Dans le cas où Est et Ouest choisissent aléatoirement le pique qu'ils nous montrent, Borel nous explique que si Est a RV2 la probabilité pour qu'il montre le 2 est 1, tandis que la probabilité pour qu'Ouest montre le 3 est 1/3 puisqu'il a 3 petites cartes.

Si on passe en revue de la même façon les 6 mains possibles pour Est et qu'on s'intéresse à toutes les possibilités de montrer, on verra que sur les cas où le 2 et le 3 sont exposés, la fréquence de RV en Est sera 1/5 soit **20%** c'est-à-dire exactement ce qu'elle était au début du coup.
Détail du calcul combinant probabilité de présence et probabilité de montrer le 2 et le 3:

<table>
<tr><th>Est</th><th>Ouest</th><th>P(présence)</th><th>P(montrer 2 et 3)</th><th>P(présence et montrer)</th><th>TOTAL</th></tr>
<tr><td>245</td><td>3RV</td><td>1/6</td><td>1/3</td><td>1/18</td><td rowspan="5">4/18</td></tr>
<tr><td>24V</td><td>35R</td><td>1/6</td><td>1/4</td><td>1/24</td></tr>
<tr><td>24R</td><td>35V</td><td>1/6</td><td>1/4</td><td>1/24</td></tr>
<tr><td>25V</td><td>34R</td><td>1/6</td><td>1/4</td><td>1/24</td></tr>
<tr><td>25R</td><td>34V</td><td>1/6</td><td>1/4</td><td>1/24</td></tr>
<tr><td>2RV</td><td>345</td><td>1/6</td><td>1/3</td><td>1/18</td><td>1/18</td></tr>
</table>

Au total : probabilité d'avoir l'une des mains possibles et de fournir le 2 en Est et le 3 en Ouest = 5/18
Probabilité d'avoir RV en Est quand le 2 et le 3 sont fournis par les mains de l'exemple =
1/18 divisé par 5/18 = 1 / 5 = **20%**

Type X Le CNRS en folie. Extrait du "journal du CNRS".

Au tribunal comme ailleurs, il existe aussi de « mauvaises intuitions mathématiques » dues à des biais cognitifs…
L. S. : Oh que oui ! Je vais vous donner mon exemple préféré, car il rend tout le monde fou, y compris les mathématiciens et même les probabilistes. Imaginez que lors d'un voyage, votre voisin, en bavardant, vous fait savoir qu'il a deux enfants dont un fils qui, vous l'apprenez au passage, est né un mardi. La question est la suivante : quelle est la probabilité que l'autre enfant de cet homme soit une fille ?
Spontanément, on aurait envie de répondre 50 %, car savoir que l'un des enfants est un garçon ne préjuge en rien du sexe de l'autre. Eh bien c'est faux ! Car si on ne connaissait que le fait que l'un des enfants est un garçon, la probabilité que l'autre soit une fille serait de 2/3. En effet – sans compter les jumeaux –, les familles de deux enfants sont réparties en quatre types : gg, gf, fg et ff. Mais comme cette famille n'est pas du type ff (puisqu'il y a un fils), elle est forcément de type gg, gf ou fg ; ces trois types étant également répartis, il y a bien deux chances sur trois que l'autre enfant soit une fille.
Mais nous savons aussi que le garçon est né un mardi. Cela paraît incroyable, absurde, de penser que ce fait puisse changer la probabilité d'avoir une fille. Pourtant c'est bien le cas ! En effet, si l'on dresse la liste de toutes les possibilités d'avoir une famille de type gg, par exemple, avec tous les jours de la semaine, on arrive à 49 possibilités. Il en sera de même pour les familles de type gf et fg, soit un total de 147 possibilités. Si l'on écarte maintenant toutes celles qui ne contiennent pas un garçon né un mardi, il en reste exactement 27, dont 13 où l'autre enfant est un garçon, et 14 où il s'agit d'une fille. D'où une probabilité de 14/27 que l'autre enfant soit une fille… Il y a de quoi s'arracher les cheveux !

Effectivement ! Mais il ne devrait pas y avoir beaucoup de cervelle au bout des cheveux qu'ils se sont arrachés.
En fait j'aimerais beaucoup savoir où LS va acheter sa dope parce qu'elle a l'air de bonne qualité.

- Est-ce que quand on nous demande "quelle est la probabilité pour que l'autre enfant soit une fille ?" l'univers n'est pas constitué de deux évènements équiprobables : fille ou garçon. Point. Probabilité 50%.
- On ne voit pas ce que vient faire l'autre enfant dans le sac des possibles puisque la probabilité d'avoir un garçon ou une fille est indépendante du sexe des enfants qu'on a déjà eu. Ce serait le cas qu'on ne verrait pas pourquoi il faudrait distinguer fg de gf puisqu'il n'entre pas dans nos spéculations que (par exemple) l'enfant de la question soit né avant ou après l'autre. Ce serait le cas qu'il faudrait appeler G le garçon dont on connait l'existence et les couples possibles seraient Gf, fG, Gg et gG. Ou on pourrait arguer que la probabilité de (fg ou gf) serait égale à ½ puisque c'est en fait la probabilité pour que l'autre enfant soit une fille et cette probabilité serait égale à celle de gg puisque c'est la probabilité pour que l'autre enfant soit un garçon. Et donc, en somme, contrairement à ce qu'affirme L.S. la probabilité pour que l'autre enfant soit une fille ne serait pas 2/3 mais 1/2. Ouf.
- Et enfin si je compte toutes les possibilités de combiner un garçon né un mardi avec une fille née tous les jours de la semaine, puis avec un garçon né tous les jours de la semaine. Puis que j'inverse ces couples (le garçon né un mardi pouvant être né avant ou après l'autre enfant) j'en compte 28 et pas 27 et dans 14 d'entre eux l'autre enfant est une fille. Probabilité ½ .

De plus L.S a oublié qu'il fallait prendre également en compte que les enfants pouvaient avoir une chambre bleue ou une chambre rose paramètre essentiel dans le calcul qu'il (ou elle) se propose de faire.

En tous cas on dirait que l'univers onirique où se situe L.S. est de type 2 puisque par exemple né un mardi et né un lundi sont deux évènements mutuellement incompatibles. Mais greffer sur un univers simple des caractères qui n'ont rien à y faire (né avant, né tel jour) semble un procédé éprouvé quand on recherche le sensationnel. C'est bien ce qu'on fait quand par exemple on ajoute aux portes de Monty Hall un caractère "ouverte" ou "fermée" qui va permettre de nous brancher sur un univers d'épreuves alors que le seul caractère qui nous intéresse est en fait "cache le trésor" ou "ne cache pas le trésor".

Type 2. Un article paru sur "Le monde".

Pour illustrer la pertinence de leur point de vue, certains comparent le jeu de Monty Hall à une épreuve où l'on demanderait au joueur de choisir une carte au hasard parmi 52.
Puis on retournerait 50 cartes qui ne seraient pas l'as de pique parmi les 51 cartes restantes et on demanderait au candidat si pour trouver l'as de pique il préfère conserver son choix initial ou changer pour l'unique carte qui n'a pas été retournée parmi les 51 cartes qu'il n'a pas choisies.

> Les probabilités sont dans ce cas de 1/52 si vous conservez votre choix initial et de 51/52 (un peu plus de 98 %) si vous modifiez votre choix, parce que vous aurez retourné en tout 51 cartes sur 52.

Évidemment le candidat effrayé par la maigreur de ses chances initiales (une sur 52) s'empresse de changer de carte. Et ceux qui utilisent cette argumentation prétendent que le changement est bon à 51 contre 1 comme l'affirme un article du monde paru sur le sujet en 2013 (encadré ci - contre).
La loi de Bayes (employée à tort et à travers en utilisant des issues impossibles) a encore fait son travail et malgré qu'on connaisse 50 cartes qui ne peuvent pas être la nôtre, la probabilité que la nôtre soit l'as de pique est encore ce qu'elle était au début : 1 / 52.

Notre instinct nous souffle qu'il vaut mieux changer de carte parce que notre tendance casinomaniaque prend le dessus, et qu'instinctivement, on se dit que si l'on jouait de nombreuses fois à ce jeu, il faudrait miser sur l'autre carte pour gagner le plus souvent.
Mais d'un autre côté, il devrait nous faire considérer avec méfiance ce « une chance sur 52 » car une fois que nous avons éliminé 50 hypothèses que nous pouvions faire dans l'univers initial sur la carte que nous avons tirée (50 issues possibles), la probabilité qu'elle soit l'une des deux cartes restantes ne peut plus être ce qu'elle était au début c'est à dire 1/52.

●●●●●

Je pense que c'est la recherche du sensationnel qui justifie la publication de telles probabilités paradoxales. Il n'y a que 2 portes possibles mais l'une supporte les 2/3 de la probabilité. RV sont en Est une fois sur 6 mais quand on fournit aléatoirement 2 petits piques ils sont en Est une fois sur cinq. Je te montre 50 cartes qui ne sont pas l'as de pique et la probabilité que la carte que tu as choisie soit l'as de pique est encore 1/52, la probabilité que le second enfant soit une fille est 14/27.
On va voir que dans chaque cas on oublie l'axiomatique et la notion de possible, on quitte un univers stagnant désespérément banal pour immerger le lecteur dans une vision en fréquence parce qu'on sait que s'il se rend à l'évidence qu'en jouant souvent à ce jeu il gagnera avec telle fréquence, on l'aura convaincu que la probabilité, dans la situation où il se trouve est égale à cette fréquence. Et ça marche.

Alors ?

D'un point de vue axiomatique je ne trouve aucune irrégularité en ce qui concerne l'application de la loi de Bayes aux univers de type 1.
Par contre, je pense qu'il est impropre d'appliquer la loi de Bayes à un univers de type 2 pour calculer une probabilité des causes et je vais essayer de démontrer pourquoi.

1. Expérience aléatoire

D'après l'axiomatique, les probabilités ne s'appliquent qu'aux expérience aléatoires.
Commençons par poser une question innocente : Qu'est-ce qu'une expérience aléatoire ?
Une *expérience aléatoire* est une expérience dont l'issue (le résultat) dépend uniquement du hasard.
Pouvons-nous dire

- Qu'une expérience dans laquelle intervient un gus qui sait quelle porte il doit ouvrir quand le joueur n'a pas choisi celle derrière laquelle se cache le trésor est une expérience aléatoire ?
- Qu'une expérience où un gus cherche fébrilement l'as de pique parmi les 51 cartes que n'a pas choisi le joueur pour surtout le garder caché est une expérience aléatoire ?
- Qu'une expérience dans laquelle Est et Ouest montrent leurs piques en fonction d'une stratégie inconnue qui aurait une influence sur notre calcul est une expérience aléatoire ?

J'ai quelque réticence à admettre ce préalable qui pourtant est fondamental.
Parce que l'issue de l'expérience ne dépend pas uniquement du hasard, mais d'une personne qui manipule le hasard.
Et vous, qu'est-ce que vous en pensez ?
Bon laissons de côté cette petite réserve. Dans ce qui suit, nous effectuerons nos calculs dans l'univers des probabilité psychologiques d'Emile Borel où les petits piques sont fournis aléatoirement (ce qui exclut toute manipulation du hasard). Puisqu'après tout c'est un mathématicien éminent, mondialement connu pour ses travaux sur la mesure, professeur de probabilités à la Sorbonne et même brièvement député et ministre de la marine, dans une vie remplie d'hommages et de médailles, on devrait faire confiance à ses calculs.

2. Quel univers ?

D'abord, si Borel fait fournir les cartes à partir des mains possibles au moment du calcul et que RV se trouvent en Est dans 16,7% de ces mains, on se demande pourquoi 16,7% ne serait pas la probabilité de RV en Est.
Au bridge, les probabilités sont importantes et 99% des calculs réalisés par Borel lui-même s'appuient sur l'univers des mains possibles. On compte les mains possibles au moment du calcul, parmi elles celles qui constituent l'évènement auquel on s'intéresse et le rapport des deux nombres constitue la probabilité recherchée.

- Selon ce procédé, la probabilité de RV en Est serait 16,7%. Pourquoi Borel passe-t-il sans prévenir d'un univers de mains à un univers d'épreuves différenciées par le mode de fourniture des cartes ?
- Dans un autre cas on passe d'un univers de portes, qui nous permet d'évaluer la probabilité avant qu'aucune porte ne soit ouverte, à un univers d'épreuves où les portes sont ouvertes selon que le joueur a choisi la bonne ou non.
- Dans un autre cas pour trouver une probabilité de 1/52 pour le changement de cartes, il faut encore se situer dans un univers d'épreuves où le joueur choisit une carte et le maître du jeu lui en montre 50 qui ne sont pas l'as de pique.

Dans chaque cas, on passe d'un univers dont les évènements existent simultanément, à un univers d'épreuves qui serait façonné et quantifié par le déroulement d'évènements simultanément incompatibles.
Dans chaque cas, pour calculer une probabilité, il faudrait que le joueur imagine qu'il est dans un univers infini d'épreuves dont les fréquences déterminent la probabilité dans SON épreuve. Il se dit si je participais à de nombreuses épreuves, cet évènement se produirait avec telle fréquence, donc cette fréquence est la probabilité de mon évènement. Mais est-ce que cela a un sens si comme le joueur de Monty Hall on va participer à une épreuve unique avant d'aller se planquer dans un monastère thibétain et d'y passer le reste de sa vie avec le pognon qu'on a ramassé à ce putain de jeu en décidant de ne pas changer de porte ?

Le problème est que l'univers qu'on utilisait avant que la machine infernale ne soit lancée est toujours opérationnel au moment du calcul.
Monty hall : pourquoi ne dirait-on pas que la probabilité demandée est la probabilité pour que le trésor soit situé derrière la porte que j'ai choisie **sachant** qu'il n'est pas derrière la porte qu'on a ouverte (probabilités conditionnelles) ?
Si initialement les portes étaient équiprobables est ce que le fait d'imaginer (ce qui n'a aucun sens) qu'on participe à une série d'épreuves peut modifier cette équiprobabilité. Si la probabilité était 1/3 au début se pourrait-il qu'elle ait diminué alors qu'on a procédé à une restriction des possibles dans l'univers initial ?
Borel : pourquoi ne dirait-on pas que la probabilité de RV en Est est devenue probabilité de RV en Est sachant que le ♠2 est en Est et le ♠3 en Ouest ?
On a donc 2 procédés de calcul de la probabilité concurrents alors que le protocole de construction de l'univers est clairement défini avant l'expérience, par la situation du trésor la distribution ou le choix d'une carte.
Pourquoi l'univers initial dont la structure suffit à répondre aux questions qu'on se pose aurait-il besoin d'être brassé, modifié, reconstruit différemment dans de nombreuses épreuves pour apporter les réponses à ces questions ?
Et puis, mystérieusement, dans chaque cas, le calcul basé sur la loi de Bayes donne une probabilité qui est la probabilité initiale : avant qu'on ouvre une porte, avant qu'on montre 2 cartes, avant qu'on montre 50 cartes.
Il n'est pas vraiment étonnant que si un évènement a une probabilité initiale P en reproduisant l'expérience plusieurs fois on trouve que cet évènement se produit avec une fréquence P. Ce qui est étonnant c'est que la probabilité de cet évènement ne change pas alors qu'on a modifié l'univers initial.

3. L'univers à l'instant t.

Revenons au calcul de Borel

Est	Ouest	P(présence)	P(montrer 2 et 3)	P(présence et montrer)	TOTAL
245	3RV	1/6	1/3	1/18	4/18
24V	35R	1/6	1/4	1/24	
24R	35V	1/6	1/4	1/24	
25V	34R	1/6	1/4	1/24	
25R	34V	1/6	1/4	1/24	
2RV	345	1/6	1/3	1/18	1/18

Au total : probabilité d'avoir l'une des mains possibles et de fournir le 2 en Est et le 3 en Ouest = 5/18
Probabilité d'avoir RV en Est quand le 2 et le 3 sont fournis par les mains de l'exemple
= 1/18 divisé par 5/18 = 1 / 5 = **20%**
Borel nous explique que pour déterminer l'univers de notre calcul "on doit éliminer les hypothèses incompatibles avec ce que nous avons appris, mais qui auraient pu être envisagées comme possibles au début, dans les calculs faits avant la donne ".
Comprenez avant l'instant où on calcule la probabilité. Avant l'instant t.
Et il ajoute :" « La faute de raisonnement (de l'évaluation de la probabilité de RV en Est à 1/6 basée sur les mains possibles) provient de ce qu'on tient seulement compte de la probabilité d'entrée en jeu de la cause : Est recevra RV2, 5 fois sur 100. On oublie de tenir compte de la probabilité que, la cause étant entrée en jeu, c'est-à-dire Est ayant reçu RV2, Est et Ouest jouent tous deux comme ils l'ont fait. Voici maintenant le raisonnement correct qui consiste à appliquer la formule de Baye.»
Borel est donc soucieux de la compatibilité de son univers avec l'instant t puisqu'il fait fournir les petits piques de façon aléatoire **avec des mains qui, toutes, situent le 2 et le 3 là où on les a vus.** Effectivement, si on se situait avant l'instant t Ouest pourrait fournir le 2, par exemple, ce qui serait incompatible avec la certitude que le 2 est en Est. Un univers où le 2 serait en Ouest serait incompatible avec la logique (la notion de "possible" est universelle) et avec l'axiomatique.
Mais, de ce point de vue, Borel devrait réprouver le procédé de calcul qui consiste à imaginer que d'autres cartes que le 2 et le 3 soient fournis parce dans notre donne, au moment du calcul, la probabilité qu'Est et Ouest aient montré ou montreront d'autres cartes que le 2 et le 3 est NULLE.

Allons plus loin. Le propos de Borel est de démontrer que quand Est et Ouest fournissent 2 piques aléatoirement la probabilité de RV en Est (ou en Ouest) est 20%. Mais il y a un problème…

Pour que la fréquence de fourniture des piques ait une réalité, il faut se situer dans un grand nombre d'épreuves.
Par exemple on donne successivement à Est et Ouest chacune des 6 mains possibles et on leur demande de fournir aléatoirement un petit pique puis, quand ils ont terminé la série de 6 donnes, on recommence l'opération par exemple 3000 fois, ce qui fait 18000 donnes.
Sur ces 18000 donnes le 2 et le 3 ont été fournis 5000 fois et sur ces 5000 donnes RV étaient en Est 1000 fois.
Parfait. Mais il y a un problème. Sur ces 18000 donnes Est et Ouest ont fourni aléatoirement 2 piques, on a calculé 18000 fois que la probabilité de RV en Est était 20% et pourtant RV étaient en Est dans une donne sur 6 c'est à dire 3000 fois. D'après le calcul de Borel ne devraient-ils pas être en Est une fois sur 5 c'est-à-dire 3600 fois ?

Essayons de comprendre le principe de cette escroquerie.
Pour que la probabilité de fournir le 2 et le 3 avec par exemple 24R pour 35V soit 1/4, il faut que de temps en temps on fournisse le 4 et le 5 par exemple. Et quand c'est le cas, il est impossible que la probabilité de RV en Est soit 20% puisque si Est a le 2 et le 4, il ne peut avoir RV. Et quand ce sont le 2 et le 5 qui sont fournis c'est Ouest qui ne peut avoir RV alors qu'on calcule que la probabilité qu'il les ait est 20%. Or une chose est sûre cette probabilité est la même en Est et en Ouest.
Conclusion :
Borel essaie de démontrer que, quand Est et Ouest montrent aléatoirement chacun un pique, dans son univers la probabilité pour que RV soient situés en Est (ou en Ouest) est 20% alors que cette probabilité est nulle quand les joueurs fournissent n'importe quel couple de piques autre que le 2 et le 3.

Vous ne trouvez pas que ça fait beaucoup d'incohérences ?

Situez-vous dans un univers de type 1 et rien de tel ne se produit.
A l'instant t la probabilité d'être malade et la probabilité d'être vacciné ont un sens correspondant à la fréquence des malades et des vaccinés dans notre univers.
Donc la probabilité d'être malade et vacciné à un sens (fréquence de l'intersection des deux ensembles dans l'univers et produit des probabilités calculées).
Si Bayes vous a fait calculer que la probabilité d'être vacciné quand on est malade est 25%, faites défiler dans une série d'épreuves les 40 malades de notre univers et vous allez trouver que parmi eux 10 sont vaccinés.
Fréquence 25%.

4. Le test de la probabilité totale.

Revoyons le calcul de Borel.

Est	Ouest	P(présence)	P(montrer 2 et 3)	P(présence et montrer)	TOTAL
245	3RV	1/6	1/3	1/18	4/18
24V	35R	1/6	1/4	1/24	
24R	35V	1/6	1/4	1/24	
25V	34R	1/6	1/4	1/24	
25R	34V	1/6	1/4	1/24	
2RV	345	1/6	1/3	1/18	1/18

Si ce tableau permet de trouver la probabilité de RV quand le 2 et le 3 sont fournis, il permet aussi de calculer la probabilité de 4 en Est, de 5 en Est, de V en Est et de R en Est quand 2 et 3 sont fournis.

Pour chacune de ces probabilités on trouve $\frac{\frac{1}{24}+\frac{1}{24}+\frac{1}{18}}{\frac{5}{18}}=\frac{1}{2}$

Quand le 2 et le 3 sont fournis, ces probabilités sont toutes évaluées à $\frac{1}{2}$

Montrons maintenant les 2 mains, comportant chacune 2 cartes inconnues au moment où se pose le problème et inscrivons un numéro au dos de chaque carte inconnue.

Appliquons maintenant les probabilités élémentaires à cet univers :

Le calcul précédent nous ayant appris que chaque carte avait la même probabilité de se trouver en Est ou en Ouest, on peut calculer que la probabilité pour que la carte 1 soit le R est $\frac{1}{4}$ (la moitié de la probabilité de R en Est).

Et une fois que ce R est en place, la probabilité pour que la carte 2 soit le V est $\frac{1}{3}$ (il ne reste que 1 place vacante pour 3 cartes)

Et donc la probabilité pour que la carte 1 soit le R et la carte 2 soit le V est $\frac{1}{4} \text{x} \frac{1}{3} = \frac{1}{12}$.

Et on peut établir de même que la probabilité pour que la carte 1 soit le V et la carte 2 soit le R est aussi $\frac{1}{12}$.

D'où nous pouvons déduire que la probabilité de RV en Est est $\frac{1}{12} \text{ x } 2 = \frac{1}{6}$.

Or cette probabilité de $\frac{1}{6}$ trouvée à partir du procédé de calcul proposé par Borel lui-même, en exploitant les règles élémentaires du calcul de probabilité, c'est justement la probabilité qu'on trouve dans l'univers des mains possibles et que Borel conteste en disant qu'elle est égale à 20%.
N'est-ce pas la preuve que la probabilité calculée par Borel n'est pas vraiment pertinente dans notre situation ?

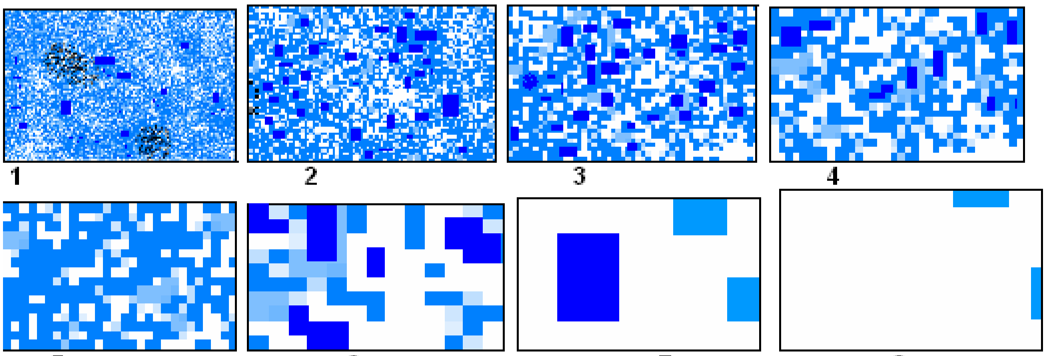
Un même raisonnement appliqué à d'autres univers de type 2 déboucherait sur les mêmes conclusions :

Il est impropre d'appliquer la loi de Bayes pour trouver une probabilité des causes à un univers de type 2. C'est-à-dire à un univers d'épreuves où se produisent des évènements mutuellement incompatibles comme la fourniture d'une carte plutôt qu'une autre, l'ouverture d'une porte plutôt qu'une autre, etc.

Ce qui justifie cette assertion est simple : quand un évènement A se produit (fourniture d'une carte, ouverture d'une porte) le seul univers dans lequel la production d'un évènement B simultanément incompatible avec A aurait une probabilité non nulle serait un univers précédant l'évènement A. Puisque A et B ne sont simultanément **possibles** qu'avant que l'un d'entre eux ne se produise.
L'univers de notre calcul n'est-il pas un univers où A s'est produit ? Un univers où la probabilité de A est 1 et la probabilité de B, 0 puisque A et B sont incompatibles. L'univers de notre calcul ne doit-il pas prendre en compte les évènements **possibles** au moment du calcul et seulement eux ? Et si c'est le cas qu'est-ce qui vous donne le droit d'utiliser pour votre calcul une probabilité de B non nulle et une probabilité de A différente de 1 comme vous le faites quand vous violez sauvagement la loi de Bayes sous le prétexte de trouver une probabilité des causes ?
En fait il vous est difficile d'imaginer que la probabilité dans une épreuve ne découlerait pas directement du protocole de tirage utilisé dans la série d'épreuves dont elle fait partie alors que pour simuler la probabilité dans une épreuve par une fréquence dans une série d'épreuves, il faut imposer à toutes les épreuves une absolue compatibilité avec le résultat qu'a produit la nôtre. Et ce n'est pas le cas dans votre univers de type 2.

Essayons d'aller plus loin et d'illustrer, à travers des exemples, la différence entre probabilité dans une épreuve et fréquence dans un univers d'épreuves de type 2.

◙ D'abord démontrons, à travers la parabole de la chute, qu'il est incorrect de faire l'amalgame entre fréquence initiale dans une série d'épreuves déformant l'univers initial et probabilité dans une épreuve une fois que l'univers a été déformé.
La situation est la suivante :
Mille fusées situées en orbite haute vont retomber sur la planète Uncououiuncounon qui est composée à 59% de bleu de Bresse et à 41% de blanc cabécou.
Les techniciens de cap carnaval suivent la chute erratique des engins à travers 8 photos, de plus en plus rapprochées, toutes ont en commun de contenir la zone où l'impact final est possible.


1 2 3 4 5 6 7 8

Il est incontestable que, la trajectoire de la fusée étant aléatoire, la fusée va s'écraser sur du bleu dans 59% des épreuves.
Mais dans l'épreuve dont vous voyez les photos, au vu de la huitième, diriez-vous que la probabilité pour qu'elle s'écrase sur du bleu est encore 59% ?
C'est donc que, comme le dit Borel, quand une épreuve nous apprend des choses qui modifient ou précisent l'univers initial, il faut oublier les probabilités initiales et refaire les calculs en fonction du zoom que nous avons opéré sur l'univers initial.

◙ Allons plus loin et essayons de définir le protocole de simulation en fréquence d'une probabilité par une série d'épreuves :
Vous prenez 10 cartes contenant la ♣D et vous les distribuez aléatoirement en deux tas de 5 cartes. Le tas A et le tas B.
Quelle est la probabilité que la dame soit dans le tas A ?
Intuitivement c'est 50%. Si on compte le nombre de combinaisons de cartes possibles pour composer le tas A et parmi elles celles qui contiennent la dame, le rapport du second nombre au premier évalue une probabilité égale à 0,5.
Et sur de nombreuses épreuves, la fréquence de la ♣D en A sera 50%.

Maintenant retournons une carte du tas A. C'est le ♥2. Quel est la probabilité que ce tas contienne la ♣D ?
Si pour calculer cette probabilité vous vous référez toujours à 1000 épreuves dans lesquelles vous distribuez les 10 cartes 5 par 5 de manière aléatoire il ne fait aucun doute que la ♣D se trouvera dans le tas A dans 50% des épreuves.
Mais cette probabilité n'est pas la probabilité pour que le tas A contienne la ♣D une fois qu'on sait qu'il contient le ♥2.

- Si parmi les combinaisons possibles pour les 4 cartes inconnues du tas A on fait le rapport entre le nombre de celles qui contiennent la ♣D et le nombre total des possibles, on trouve 4/9. La ♣D est plus souvent dans le tas B que dans le tas A.
- Si avant de retourner le ♥2 on fait le calcul de la probabilité pour que le tas qui contiendra le ♥2 contienne aussi la ♣D on trouve 4/9.
- Si on distribue aléatoirement les 9 cartes qui ne sont pas le ♥2 en un tas de 4 auquel on joint le ♥2 et un tas de 5, le tas de 5 qui contient le ♥2 contiendra la ♣D environ 4 fois sur 9.
- Si on donne les 10 cartes à un ordinateur et qu'on lui demande de tirer aléatoirement 9000 donnes où le ♥2 sera dans le tas A , il comptera que la ♣D se trouve dans le tas A environ 4000 fois sur 9000.

En fait quand vous évaluez à 50% la probabilité de la ♣D en A à travers une série d'épreuves où vous distribuez aléatoirement les cartes en deux tas de 5, vous commettez une erreur parce qu'au cours de vos épreuves notre ♥2 sera situé alternativement dans le tas A ou dans le tas B. Ce qui disqualifie cette série d'épreuves pour simuler une probabilité dans la nôtre puisque nous sommes dans une épreuve où la probabilité pour que le ♥2 se situe dans le tas B est nulle.

C'est donc que le procédé évaluant une probabilité à postériori à partir d'une série d'épreuves où il se produit des évènements incompatibles avec celui qui est avéré dans la nôtre, est incorrect. Si on veut simuler la probabilité une fois qu'un évènement s'est produit, par une fréquence dans une série d'épreuves il faut le faire dans une série d'épreuves où l'évènement qui s'est produit dans la nôtre se produit systématiquement, à l'exclusion de tous les évènements qui sont incompatibles avec lui.
Or, quand vous faites l'amalgame entre probabilité des causes dans un univers de type 2 et fréquence dans une série d'épreuves produisant des évènements incompatibles vous bafouez ce principe. Les fréquences produites par votre protocole ne peuvent servir de base à un calcul de probabilité une fois que l'un des évènements incompatibles s'est produit.

◙ Maintenant faisons appel à votre bon sens : Après un naufrage vous débarquez dans une île de Papouasie et sur la plage vous lisez un écriteau où il est écrit :
"Etat civil de l'île aujourd'hui 30 adultes de sexe féminin célibataires, 30 adultes de sexe masculin célibataires et 40 adultes de sexe féminin mariés à 40 adultes de sexes masculin".
Vous levez les yeux du panneau et vous voyez une adulte de sexe féminin. Quelle est la probabilité qu'elle soit mariée ?

- Pour calculer cette probabilité allez-vous faire le rapport de l'effectif des femmes mariées à l'effectif total des femmes (4/7) ?
- Ou allez-vous imaginer que vous vous situez dans une série d'épreuves où le hasard passe en revue, tour à tour, l'ensemble des femmes célibataires, l'ensemble des hommes célibataires, l'ensemble des couples et envoie à notre rencontre un élément pris au hasard dans chaque ensemble, ce qui fait que dans le cas d'un couple il ne nous envoie une femme qu'une fois sur deux ? Dans ce protocole de tirage quand vous rencontrez une femme elle est célibataire avec la fréquence de 1/3 et mariée avec la fréquence de 1/6 ce qui devrait donner d'après Bayes une probabilité de 1/3 pour qu'elle soit mariée ?

Bon quand je dis "d'après Bayes" je raconte n'importe quoi parce que seul le premier calcul (4/7) découle des lois de Bayes. Le second est le résultat d'un vague amalgame entre une probabilité et une fréquence dans une série d'épreuves obéissant à un protocole précis. Le calcul de cette fréquence n'a absolument rien à voir avec Bayes, mais bien évidemment s'il s'agit d'une probabilité elle est calculée avant que vous ne rencontriez une femme, quand la probabilité de rencontrer un homme n'était pas nulle. C'est la probabilité pour que la prochaine femme que produira ce protocole de tirage soit une femme mariée. Et cette probabilité n'est pertinente que parce la probabilité de rencontrer un homme n'est pas nulle.

Une fois que vous avez rencontré une femme et qu'on vous demande de calculer la probabilité pour qu'elle soit mariée, si vous imaginez une expérience en fréquence qui simule votre probabilité vous n'avez pas le droit de faire intervenir dans cette expérience un choix aléatoire en ce qui concerne les couples parce qu'en le faisant, en tirant une fois un homme, une fois une femme, vous ne respecteriez pas la conformité au protocole de simulation qui impose que vous ne tiriez que des femmes.
Il faut donc abandonner l'idée de se référer à cet univers d'épreuves pour évaluer la probabilité demandée et vous tourner tout simplement vers l'axiomatique.

Le seul univers concevable pour votre calcul est celui qui est formé des femmes mariées et célibataires. Quand vous imaginez l'univers dont peut provenir cette femme, il ne peut être composé que de femmes (puisque l'univers initial était composé d'hommes et de femmes et que la connaissance du sexe du tirage élimine le sous-ensemble des hommes de cet univers), mariées ou célibataires (puisque c'est le caractère qui définit la partition de notre univers en 2 sous-ensembles qu'il va falloir mesurer pour calculer la probabilité demandée). Et si les femmes des deux sous-ensembles ne sont pas équiprobables, on ne peut invoquer pour justifier cette rupture de l'équiprobabilité un processus qui produirait des épreuves nous mettant en présence d'un homme.

Par contre, on peut imaginer un protocole de tirage qui ne nous enverrait que des femmes et une femme célibataire 2 fois plus souvent qu'une femme mariée ou une femme célibataire aussi souvent qu'une femme mariée sans que les tirages soient simultanément incompatibles puisque dans tous les cas il s'agit de tirer une femme et qu'une femme mariée est indiscernable d'une femme célibataire.
Cela reviendrait, en rompant l'équiprobabilité des femmes à fonder la probabilité sur une fréquence. Mais en aucun cas cela ne justifierait l'utilisation de la loi de Bayes qui s'applique à un tirage unique dans des univers structurés par un croisement de plusieurs caractères (par exemple homme/femme + marié/célibataire ou sain/malade + vacciné/non vacciné).

◙ Suite à ce que nous venons de dire nous devons apporter une précision sur les univers de type 2.
Ils sont constitués d'épreuves différenciées par des actes mutuellement incompatibles <u>et discernables</u>.
Cela étant dit toutes les anomalies que nous avons relevées en contrôlant mathématiquement la pertinence de la probabilité calculée par Bayes dans l'univers de type 2 de Borel tiennent à un seul fait :
Si j'appelle "tirage" l'acte résultant de l'épreuve (nature de la porte ouverte, cartes fournies, femme vue, …).
Il est important de comprendre que pour simuler la probabilité dans une épreuve où un tirage s'est produit par une fréquence dans une série d'épreuves, il faut absolument que tous les tirages dans cette série soient compatibles avec le tirage produit dans la nôtre.
Ce n'est pas le cas dans un univers de type 2 puisqu'ils reposent sur une série d'épreuves où se produisent des tirages incompatibles avec le nôtre. Et en conclusion, les fréquences découlant des protocoles de tirage dans un univers de type 2 ne peuvent en aucun cas être utilisées pour calculer la probabilité dans notre épreuve une fois un tirage constaté.
Autrement dit :
Quand une porte a été ouverte, quand des cartes ont été fournies, pour calculer ou simuler une probabilité à ce stade par une fréquence dans une série d'épreuves on ne peut s'appuyer sur les fréquences d'une série d'épreuves ou d'autres portes que la nôtre auraient été ouvertes ou d'autres cartes que les nôtres fournies.

En conclusion : **Il est incorrect d'appliquer la loi de Bayes dans un univers de type 2.**

Note de l'auteur : je ne connais pas les mœurs de nos amis, habitants de la Papouasie, mais s'il y a parmi eux des hommes mariés à des hommes ou des femmes mariées à des femmes, je leur demande d'excuser le parti-pris de mon exposé sur la composition de l'état civil de l'ile. Il découle d'un souci de simplification et non de préjugés rétrogrades. En fait je n'ai absolument rien contre le mariage entre personnes du même sexe à condition qu'ils ne se séparent pas avec une fréquence immuable chaque fois que je débarque sur leur île. Cette remarque s'applique également aux couples du moindre choix.