Atelier C 06

Premières modélisations d'expériences aléatoires :

Comment éviter les solutions hasardeuses ?

Annie Fauconnet

CMI - Université de Provence Marseille

L'atelier a regroupé une quarantaine de collègues et a été très animé, de façon sympathique d'ailleurs.

On ne rendra compte ici que de certains des exemples qui ont été traités (avec leur numéro initial).

Remarque : les énoncés étaient prévus pour susciter le travail des participants, ce ne sont pas des exemples de sujets à proposer en contrôle !

I. Utilisation d'une hypothèse d'équiprobabilité

1. *On prend au hasard 5 cartes dans un jeu de 32. Calculer les probabilités des événements:*

   a) il y a exactement 3 as parmi les cartes tirées,

   b) il y a au moins un valet,

   c) il y a au moins un as et un valet, ...

L'hypothèse naturelle est que les C_{32}^{5} lots de 5 cartes parmi 32 que l'on peut obtenir sont équiprobables (on ne s'intéresse pas ici au procédé de tirage des cartes qui réalise cette hypothèse).

Soit Ω l'ensemble de tous ces lots de 5 cartes : à toute proposition e qui définit un événement on peut faire correspondre la partie E de Ω formée des éléments pour lesquels la proposition e est vraie, on notera respectivement A, B, C les parties de Ω qui représentent les événements indiqués en a), b), c).

Détermination de la probabilité P sur les parties de Ω :
Les propriétés des fréquences entraînant en particulier que P doit être additive et telle que $P(\Omega) = 1$, l'hypothèse d'équiprobabilité conduit à $P(E) = \text{Card}(E)/\text{Card}(\Omega)$.

a) Calcul de Card(A)

Pour le faire on regarde comment on peut fabriquer un élément quelconque de A :

1°) on choisit un lot de 3 as parmi les 4 du jeu : il y a C_4^3 résultats différents possibles,

2°) on choisit un lot de 2 cartes parmi les 28 qui ne sont pas des as : il y a C_{28}^2 résultats différents possibles, à chaque lot de 3 as on peut adjoindre un lot quelconque de 2 cartes non as, d'où $\text{Card}(A) = C_4^3 \times C_{28}^2$ (card d'un ensemble produit)

b) Calcul de P(B)

La tentative de calcul de card(B) par la méthode précédente présentant des difficultés on s'intéresse à l'événement contraire ("pas de valet") représenté par $\overline{B}$, complémentaire de B dans Ω.

c) Calcul de P(C)

La proposition qui définit l'événement fait apparaître la conjonction "et", elle s'explicite sous la forme : "il y a au moins un as" et "il y a au moins un valet", d'où $C = A' \cap B$ si on note A' la partie qui représente l'événement "il y a au moins un as". Mais pour calculer P(C) il vaut mieux comme ci-dessus passer au complémentaire : $\overline{C} = \overline{A'} \cup \overline{B}$, $P(\overline{C}) = P(\overline{A'}) + P(\overline{B}) - P(\overline{A'} \cap \overline{B})$

On sait calculer $P(\overline{A'}) = P(\overline{B})$ et $P(\overline{A'} \cap \overline{B})$ grâce à l'expression commode de card $(\overline{A'} \cap \overline{B})$: C_{24}^5.

2. *Une boîte contient 10 boules indiscernables au toucher, dont 2 blanches, 5 rouges et 3 vertes. On prend 4 boules au hasard dans cette boîte, quelle est la probabilité pour que trois des boules tirées soient rouges et la quatrième blanche ?*

Manifestement on s'intéresse aux couleurs des boules tirées. Le résultat d'une expérience peut être représenté par un triplet d'entiers positifs ou nuls (x,y,z), de somme quatre, x étant le nombre de boules blanches, y le nombre de boules rouges et z le nombre de boules vertes dans le lot tiré. L'événement qui nous intéresse est alors représenté par le singleton {(1,3,0)} : c'est un événement élémentaire.

Détermination de la probabilité : l'expression "on prend 4 boules au hasard dans une boîte..." conduit à faire une hypothèse d'équiprobabilité analogue à celle de l'exercice 1, mais pour l'expliciter il faut individualiser les 10 boules présentes : on imagine qu'on les numérote de 1 à 10. Notons Ω l'ensemble des lots de 4 boules numérotées, et E la partie de Ω qui représente l'événement dont on cherche la probabilité, P(E) se calcule selon la même méthode que pour P(A) dans l'exercice 1.

Remarques

- Ici Ω n'est pas l'ensemble des issues possibles, il est plus riche, mais il reste que tout événement peut être représenté par une partie de Ω, et **les opérations logiques sont traduites en opérations ensemblistes**.
- Notons respectivement X_b le nombre de boules blanches, X_r le nombre de boules rouges et X_v le nombre de boules vertes dans le lot de 4 boules qu'on va tirer : on ne connaît pas à l'avance les valeurs qui seront prises par ces caractères, mais elles seront bien déterminées une fois l'expérience faite : X_b, X_r, X_v **sont des variables aléatoires** associées à la situation, et on peut noter $E = [X_b = 1] \cap [X_r = 3]$, ce qui a l'avantage d'être facile à décoder.

4. *On lance une pièce de monnaie neuf fois consécutives. Sachant qu'à chaque coup la pièce présente "pile" ou "face" et que toutes les séquences possibles sont équiprobables,*
   a) calculer la probabilité pour que "face" apparaisse au deuxième et au cinquième rangs,
   b) calculer la probabilité pour qu'il apparaisse plus de "pile" que de "face".

On peut enregistrer le résultat d'une expérience sous la forme d'une suite de lettres (p ou f) L'univers est alors l'ensemble Ω des mots de 9 lettres p ou f. On peut aussi enregistrer seulement les rangs des lancers où l'on obtient "face", l'univers est alors l'ensemble Ω' des parties de {1, 2,..., 9} et Card(Ω) = Card(Ω') = 2^9. Toutes les issues possibles étant supposées équiprobables la probabilité d'un événement est déterminée comme dans l'exercice 1.

– Posons $A = F_2 \cap F_5$ (la notation parle d'elle-même), le calcul de card(A) se fait par l'intermédiaire d'un procédé de fabrication d'un élément quelconque de A (partie de Ω, ou de Ω' selon le choix fait).

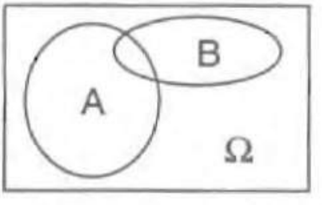
On pourra, plus tard, vérifier que ce modèle assure l'indépendance des événements F_2 et F_5 , etc..., ce qui est très satisfaisant (mais les programmes en vigueur au lycée ne permettent pas de construire l'espace probabilisé à partir de l'hypothèse d'indépendance mutuelle des 9 lancers...) .

– Soit $B = [\, X_p > X_f \,]$, en notant X_p et X_f les nombres aléatoires de "pile" et de "face" dans une suite de neuf lancers.

Il est clair que $\Omega = [\, X_p > X_f \,] + [\, X_p < X_f \,]$ (+ note l'union disjointe) et que les deux parties ont le même cardinal (par la bijection obtenue en permutant p et f), d'où P(B) = 1/2.

II. Utilisation d'hypothèses issues d'expériences préalables

2. *Soit une expérience aléatoire dans laquelle deux événements, "a" et "b", peuvent éventuellement se produire. Sur cent reproductions qui ont été faites de cette expérience on a observé 74 fois l'événement "a", 65 fois l'événement "b", et 42 fois la conjonction des deux.*
Les événements "a" et "b" étant représentés par des parties A et B d'un ensemble Ω peut-on choisir de poser $P(A) = 3/4$, $P(B) = 2/3$ et $P(A \cap B) = 4/10$?



Ce sont des fractions simples proches des fréquences* observées. Si l'on admet ces valeurs on peut calculer

$P(A \cap \overline{B}) = P(A) - P(A \cap B)$,

$P(\overline{A} \cap B) = P(B) - P(A \cap B)$ et

$P(\overline{A} \cap \overline{B}) = 1 - P(A \cup B)$, mais on trouve ici $-1/60$ ce qui est impossible et disqualifie le choix fait.

En résumé, une fois que les événements ont été représentés par des parties d'un ensemble Ω, il faut faire un certain nombre d'hypothèses, vérifiant certaines conditions, pour définir une probabilité P telle que pout tout événement "e" (représenté par une partie E de Ω), P(E) soit une valeur "convenable" de la probabilité que cet événement a de se produire.

---

*En effet quand on reproduit n fois une situation aléatoire, on constate que la fréquence de réalisation de chaque événement tend à se stabiliser quand n devient très grand, on prend cette fréquence limite supposée comme probabilité de l'événement.

Notions de probabilité conditionnelle et d'indépendance

Soit une situation aléatoire S et deux événements "a" et "b" relatifs à cette situation. Comment définir la probabilité que "b" se réalise sachant que "a"est réalisé ("a" étant de probabilité non nulle) ?

Etude statistique.

Si on reproduit N fois la situation S, dans la suite des N épreuves obtenues il y en a N_a qui réalisent "a" et parmi elles il y en a $N_{a,b}$ qui réalisent l'événement "a et b". $N_{a,b} / N_a$ est la fréquence de réalisation de "b" sur une suite de N_a reproductions de S_a (situation S dans laquelle l'événement "a" est réalisé).

Or $N_{a,b} / N_a = (N_{a,b} / N) : (N_a / N)$, d'où l'idée de définir la probabilité de "b" dans la situation S_a par $P(A \cap B) / P(A)$ si on a déjà modélisé la situation S (cela suppose $P(A) \neq 0$).

Remarque. Dans le cas où la modélisation a été faite à l'aide d'une hypothèse d'équiprobabilité de sorte que $P(A) = \text{Card}(A) / \text{Card}(\Omega)$, alors $P(A \cap B)/P(A) = \text{Card}(A \cap B) / \text{Card}(A)$, tout est cohérent.

Etude mathématique

- Si on a modélisé la situation S et défini une application P convenable sur un ensemble de parties de Ω, alors, si $P(A) \neq 0$, l'application PA définie par $PA(B) = P(A \cap B)/ P(A)$ (que l'on note habituellement $P(B / A)$) est aussi une probabilité, et sa restriction à des parties de A permet de modéliser la situation S_a.
- Inversement, pour certaines situations aléatoires, il arrive que l'on définisse l'application P en déterminant d'abord les probabilités d'un système complet d'événements "a_i"(représentés par les éléments Ai d'une partition de Ω), puis en modélisant les situations Si (S dans laquelle "a_i " est réalisé), et en appliquant les formules

$$P(B) = \sum_i P(A_i \cap B) = \sum_i P(A_i) \times P(B / A_i)$$

Indépendance de deux événements

- Définition mathématique : étant donné un espace probabilisé (Ω, Φ, P) deux "événements" A et B sont indépendants ssi $P(A \cap B) = P(A) \times P(B)$ (alors A et $\overline{B}$ ont aussi indépendants, etc...).
- Inversement si dans une situation aléatoire S deux événements, représentés par des parties A et B de Ω sont tels que la réalisation de l'un n'apporte

aucune information sur les chances de réalisation de l'autre, alors on détermine l'application P de telle sorte que $P(A \cap B) = P(A) \times P(B)$.

Les notions de probabilité conditionnelle et d'indépendance permettent de préciser le sens de l'adjectif "convenable" utilisé plus haut.

III. Exercices d'utilisation des probabilités conditionnelles

2. *Une boîte U contient 3 boules blanches et 2 boules noires, et une boîte V contient 2 boules blanches et 8 boules noires. On tire une boîte au hasard,*

a) et dans cette boîte on tire une boule au hasard, calculer la probabilité pour qu'elle soit blanche,

b) et dans cette boîte on fait successivement deux tirages d'une boule, avec remise. Quelle est la probabilité de tirer une boule blanche au deuxième coup sachant qu'on a tiré une blanche au premier ?

On peut enregistrer le résultat d'une expérience sous la forme d'une suite de lettres et poser pour a) : $\Omega_a = \{ (u,b), (u,n), (v,b), (v,n)\}$

pour b) : $\Omega_b = \{ (u,b,b), (u,n,b),\ldots (v,b,n), (v,n,n)\}$

Il est plus important de noter les parties qui représentent les événements constituants de chaque étape :

• U (resp. V) les parties (respectivement de Ω_a et de Ω_b) qui représentent "la boîte tirée est U (resp. V)"

• B (resp. N) la partie de Ω_a qui représente "la boule tirée est blanche (resp. noire)", B_i (resp. N_i) la partie de Ω_b qui représente "la i-ième boule tirée est blanche (resp. noire)".

Quand on réalise l'expérience a), ou bien c'est l'événement U qui se réalise, ou bien c'est l'événement V, ce qui se traduit par $\Omega_a = U + V$, et ou bien c'est l'événement B qui se réalise, ou bien c'est l'événement N, ce qui se traduit par $\Omega_a = B + N$, en conjugant les deux aspects on obtient : $\Omega_a = (U + V) \cap (B + N)$, qu'on développe en $\Omega_a = U \cap B + U \cap N + V \cap B + V \cap N$. (NB. Tous les "ou bien" sont exclusifs).

On retrouve ici les quatre événements élémentaires possibles dont il faut déterminer les probabilités.

Détermination de la probabilité P (on pourrait faire une présentation en tableau carré)

a) On a envie de définir P telle que $P(U) = P(V) = 1/2$, $P(B/U) = 3/5\ldots$, $P(N/V) = 8/10$, ce qui donne $P(U \cap B) = 3/10$ ..., $P(V \cap N) = 4/10$, et aucune contradiction n'apparaît (on pourrait le démontrer).

Or $B = (U + V) \cap B = U \cap B + V \cap B$, donc $P(B) = P(U) \times P(B / U) + P(V) \times P(B / V) = 2/5$.

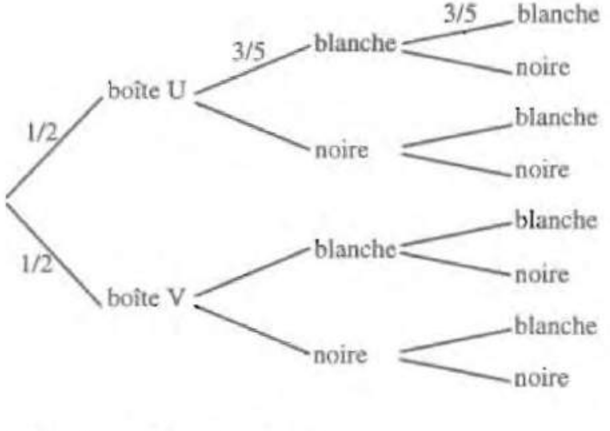
b) Ici il y a trois partitions et les événements élémentaires sont du type $U \cap B_1 \cap B_2$, mais pour déterminer l'application P, à partir de $P(U) = P(V) = 1/2$ on peut utiliser deux méthodes,

- soit $P(U \cap B_1 \cap B_2) = P(U) \times P(B_1 \ / U) \times \ P(B_2 / U \cap B_1)$
- soit $P(U \cap B_1 \cap B_2) = P(U) \times P(B_1 \cap B_2 / U) = 1/2 \times (3/5 \times 3/5)$, car dans la boîte U on fait deux tirages avec remise, donc indépendants (cela suppose qu'on a déjà modélisé ce type de situation).
  Cette deuxième méthode est la plus agréable ici et pourrait aussi être utilisée si on faisait des tirages sans remise (à condition d'avoir déjà modélisé la suite de deux tirages sans remise dans une boîte donnée).

Utilisation d'un arbre pondéré

Inconvénients :

- les probabilités sont écrites au-dessus de traits qui ne représentent pas des événements (mais peuvent dans certains cas évoquer des passages d'un état à un autre, ce qui est source d'erreurs),
- rien ne permet de manifester la signification exacte de ce qui est écrit (probabilité de quel événement, dans quelle situation conditionnelle ?).



Il faut donc absolument l'accompagner

- de la notation des événements, (par exemple comme ci-dessus par U, V, B_1 , B_2, ...),

- de la définition des valeurs écrites : $P(B_1 / U) = 3/5$, etc...

Ensuite seulement on peut formuler le problème posé dans b) : il s'agit de calculer $P(B_2 / B_1)$.

On sait que $P(B_2 / B_1) = P(B_1 \cap B_2) / P(B_1)$.

Or $B_1 = (U + V) \cap B_1 = U \cap B_1 + V \cap B_1$,

donc $P(B_1) = P(U) \times P(B_1 / U) + P(V) \times P(B_1 / V) = 2/5$.

et $B_1 \cap B_2 = (U + V) \cap (B_1 \cap B_2) = U \cap (B_1 \cap B_2) + V \cap (B_1 \cap B_2)$,

donc $P(B_1 \cap B_2) = P(U) \times P(B_1 \cap B_2 / U) + P(V) \times P(B_1 \cap B_2 / V) = 1/5$.

La recherche d'un espace probabilisé pour représenter une situation aléatoire et l'analyser est une démarche analogue à celle qui est utilisée dès le collège pour traiter certains problèmes "par l'algèbre" (choix et notation des inconnues, mise en équation ...).

La notation + pour une union disjointe a des avantages :

si $A = B + C$ alors $P(A) = P(B) + P(C)$,

si $A = B + C$ alors $E \cap A = E \cap B + E \cap C$.

Ces égalités sont généralisables à une union de plus de deux parties deux à deux disjointes.