EPFL Physique, 2ème année
Département de mathématiques

PROBABILITÉ ET STATISTIQUE

SÉRIE 1

Exercice 1 : Donner un argument combinatoire pour vérifier les identités suivantes.

(1)
$$\binom{m+n}{r} = \sum_{j=0}^{r} \binom{n}{j} \binom{m}{r-j} .$$

Indication: considérer un ensemble de n boules noires et de m boules blanches.

(2) (à faire avant (3))
$$\binom{n}{r} = \sum_{j=r}^{n} \binom{j-1}{r-1} \quad , \quad n \geq \cancel{k}\,r .$$

Indication: compter les sous-ensembles de $\{1, \ldots, n\}$ ayant r éléments tels que le plus grand élément est j.

(3)
$$\sum_{j=1}^{n} j \binom{n}{j} = n \cdot 2^{n-1} .$$

Indication: compter de deux façons différentes le nombre de comités que l'on peut former parmi un ensemble de n personnes si le nombre de personnes du comité est quelconque et si chaque comité a un président.

(4)
$$\sum_{j=1}^{n} j^2 \binom{n}{j} = n(n+1) \cdot 2^{n-2} .$$

Indication: compter de deux façons différentes le nombre de comités que l'on peut former parmi un ensemble de n personnes si le nombre de personnes du comité est quelconque et si chaque comité a un président et un secrétaire, le cumul des charges étant autorisé.

(5)
$$\binom{n}{n_1, \ldots, n_r} = \sum_{j=1}^{r} \binom{n-1}{n_1, \ldots, n_j - 1, \ldots, n_r} .$$

Exercice 2 : On considère un système de n particules pouvant chacune être dans r états d'énergie E_k différents. Le nombre de particules dans l'état E_k est n_k ($\sum_{k=1}^{r} n_k = n$). On demande de déterminer la proportion des configurations d'énergie telles que n_k est fixé $\forall k$, lorsque les particules sont classiques, resp. des bosons, resp. des fermions. Exemple numérique: $n = 3$, $r = 5$, $n_1 = n_3 = n_4 = 1$ et $n_2 = n_5 = 0$. Si les particules sont des bosons combien y a-t-il de configurations d'énergie telles que $n_k \geq 1\ \forall k$?

Exercice 3 : Considérer les vecteurs $(x_1, \ldots, x_r)$ avec composantes x_i non-négatives et entières. Combien y a-t-il de tels vecteurs

(1) si $\sum_{i=1}^{r} x_i \leq n$;
(2) si $\sum_{i=1}^{r} x_i \leq n$ et $x_i \geq 1$, $\forall i$;
(3) si $\sum_{i=1}^{r} x_i = n$ et il y a exactement k composantes égales à zéro.

BIBLIOGRAPHIE

Le premier ouvrage donne une bonne introduction à la théorie des probabilités; il est accessible à quelqu'un n'ayant aucune connaissance préalable de ce sujet. De nombreux exemples sont traités (la plupart des exemples sont répétés plusieurs fois). Le niveau théorique de cet ouvrage est un peu moins élevé que celui du cours. Les autres ouvrages couvrent ensemble la matière du cours et naturellement d'autres sujets non abordés dans le cours. Le niveau de ces ouvrages correspond assez bien au niveau du cours. L'ouvrage de Feller est un classique. La liste est très sommaire.

(1) Ross S.: A first course in Probability, Prentice-Hall 5^{e} éd. (1998). Une version française existe aux PPR. .

(2) Chung K.L.: Elementary Probability Theory, Springer (1974).

(3) Brémaud P.: Introduction aux probabilités, Springer (1988).

(4) Feller W.: An introduction to probability theory and its applications vol.1 Wiley (1957).

PROBABILITÉ ET STATISTIQUE

SÉRIE 2

Exercice 1 : Le but de l'exercice est de refaire ses gammes avec la manipulation des ensembles. Soit Ω un ensemble non vide. On désigne par $\mathcal{P}(\Omega)$ l'ensemble de tous les sous-ensembles de Ω muni des opérations usuelles d'*intersection* $A \cap B$, d'*union* $A \cup B$ et de *complémentation* $A^c := \Omega \backslash A$. On pose $A \backslash B := A \cap B^c$. Dans la suite de l'exercice les lettres capitales comme A, B, ..., désignent des éléments de $\mathcal{P}(\Omega)$, i.e. des sous-ensembles de Ω. Les opérations $\cap$ et $\cup$ sont commutatives, associatives et distributives,

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C) \quad , \quad (A \cup B) \cap C = (A \cap C) \cup (B \cap C) .$$

Vérifier les formules de De Morgan

$$C \backslash (\cup_n A_n) = \cap_n (C \backslash A_n) \quad , \quad C \backslash (\cap_n A_n) = \cup_n (C \backslash A_n) .$$

Vérifier l'affirmation suivante

$$A \subset B \iff A \cap B = A \iff A \cup B = B \iff A^c \cup B = \Omega \iff A \cap B^c = \emptyset .$$

Soit $\{A_n\}_n$ une famille d'éléments de $\mathcal{P}(\Omega)$. Montrer que cette famille possède un suprémum, i.e. qu'il existe $B \in \mathcal{P}(\Omega)$ tel que $A_n \subset B$ $\forall n$ et si $A_n \subset C$ $\forall n$ alors $B \subset C$. De même montrer que cette famile possède un infimum, i.e. qu'il existe $B \in \mathcal{P}(\Omega)$ tel que $A_n \supset B$ $\forall n$ et si $A_n \supset C$ $\forall n$ alors $B \supset C$.

Soit X un ensemble et $f : X \to \Omega$ une application définie sur X. Cette application induit une application notée f^{-1}, $f^{-1} : \mathcal{P}(\Omega) \to \mathcal{P}(X)$,

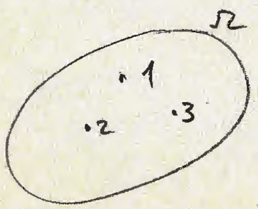
$$f^{-1}(A) := \{ x \in X : f(x) \in A \} .$$

Vérifier les affirmations suivantes

$$f^{-1}(A \backslash B) = f^{-1}(A) \backslash f^{-1}(B) \ , \ f^{-1}(\cup_n A_n) = \cup_n f^{-1}(A_n) \ , \ f^{-1}(\cap_n A_n) = \cap_n f^{-1}(A_n).$$

Une famille $\mathcal{A}$ de sous-ensembles de Ω est une **algèbre de Boole** si $\Omega \in \mathcal{A}$; si $A \in \mathcal{A}$, alors $A^c \in \mathcal{A}$; si $A, B \in \mathcal{A}$ alors $A \cup B \in \mathcal{A}$. Montrer que $\emptyset \in \mathcal{A}$; si $A_1, \ldots, A_k \in \mathcal{A}$ alors $\cap_{j=1}^k A_j \in \mathcal{A}$. Si l'algèbre possède un nombre fini d'éléments, alors il existe une famille d'éléments de $\mathcal{A}$, $C_1, \ldots, C_r$, appelés les **atomes** de $\mathcal{A}$, tels que: 1) $C_i \cap C_j = \emptyset$, $i \neq j$; 2) $\cup_{j=1}^r C_j = \Omega$; chaque $A \in \mathcal{A}$ est la réunion de C_j.
Indication: Soit $\mathcal{A} = \{A_1, \ldots, A_n\}$. Poser $A_i^1 := A_i$ et $A_i^{-1} := A_i^c$; considérer les sous-ensembles C^b, $b = (b_1, \ldots, b_n)$, $b_i = \pm 1$, définis par

$$C^b := \cap_{i=1}^n A_i^{b_i} .$$



$\mathcal{A} = \{\Omega, \emptyset, \{1,2\}, \{3\}\} = \{A_1, A_2, A_3, A_4\}$

atomes: $\{1,2\}$, $\{3\}$, $\{\emptyset\}$

CORRIGÉ DE LA SÉRIE 1

Exercice 1 : (1). Il s'agit de déterminer le nombre de sous-ensembles à r éléments d'un ensemble à $m+n$ éléments.

$$\binom{n}{j}\binom{m}{r-j}$$

est le nombre de tels sous-ensembles qui ont j boules noires et $r-j$ boules blanches. (2). Le nombre de sous-ensembles à r éléments de $\{1,\ldots,n\}$, dont le plus grand élément est j, est égal à

$$\binom{j-1}{r-1}.$$

(3). On choisit un comité avec j membres, puis on choisit son président. Le nombre de tels comités est

$$j\binom{n}{j}.$$

On peut aussi choisir d'abord un président, puis choisir les autres membres; comme le nombre des autres membres varie entre zéro et $n-1$, le nombre total de comités est $n \cdot 2^{n-1}$. Même démarche pour (4). On choisit un comité avec j membres, puis on choisit son président et son secrétaire, qui peut être la même personne. Le nombre de tels comités est

$$j^2\binom{n}{j}.$$

On peut choisir d'abord un président qui est aussi secrétaire, puis choisir les autres membres; comme le nombre des autres membres varie entre zéro et $n-1$, le nombre total de comités est $n \cdot 2^{n-1}$. On peut ensuite choisir un président et un secrétaire différent, ce qui est possible de $n(n-1)$ façons, puis choisir les autres membres; le nombre de tels comités est $n(n-1) \cdot 2^{n-2}$. (5). On marque un point d'un ensemble de n objets. On fait un rangement de ces objets dans r boîtes. L'objet marqué peut se trouver dans n'importe quelle boîte; s'il se trouve dans la boîte j il y a

$$\binom{n-1}{n_1\,,\ \ldots\,,\ n_j-1\ \,,\ \ldots\,,\ n_r}$$

rangements possibles.

Exercice 2 : Dans le cas de particules classiques, i.e. distinguables, la proportion des configurations avec $n_1,\ldots,n_r$ fixés est

$$\frac{n!}{n_1!\cdots n_r!}\cdot\frac{1}{r^n}.$$

Dans le cas des bosons, resp. fermions, pour autant que $n_j \leq 1$, cette proportion est

$$\frac{n!(r-1)!}{(n+r-1)!} \quad \text{resp.} \quad \frac{(r-n)!n!}{r!}.$$

Lorsqu'un des $n_j \geq 2$ cette proportion est nulle dans le cas des fermions. Ex. numérique: on a resp. $6/125$, $1/35$ et $1/10$. Enfin il y a $\binom{n-1}{r-1}$ configurations de n bosons avec $n_j \geq 1$, $\forall j$.

Exercice 3 : Posons $x_{r+1} := n - \sum_{i=1}^{r} x_i$. Le nombre de solutions pour (1) est $\binom{n+(r+1)-1}{(r+1)-1} = \binom{n+r}{r}$; pour (2), $\sum_{j=r}^{n}\binom{j-1}{r-1} = \binom{n}{r}$; pour (3), $\binom{r}{k}\binom{n-1}{r-k-1}$.

EPFL Physique, 2ème année
Département de mathématiques

PROBABILITÉ ET STATISTIQUE

SÉRIE 3

Exercice 1 : Ω est un ensemble et $\mathcal{A} \subset \mathcal{P}(\Omega)$, $\mathcal{B} \subset \mathcal{P}(\Omega)$ sont deux algèbres de Boole. La famille de sous-ensembles de Ω, $\mathcal{A} \cup \mathcal{B}$ est-elle une algèbre de Boole? Si $\mathcal{A}_n \subset \mathcal{P}(\Omega)$ sont des algèbres de Boole et si $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ $\forall n$, est-ce que $\cup_{n\geq 1}\mathcal{A}_n$ est une algèbre de Boole? On suppose que les algèbres $\mathcal{A}$ et $\mathcal{B}$ sont finies. Par définition l'algèbre de Boole $\mathcal{A} \vee \mathcal{B}$ engendrée par les algèbres $\mathcal{A}$ et $\mathcal{B}$ est la plus petite algèbre de Boole contenant $\mathcal{A}$ et $\mathcal{B}$. Décrire l'algèbre de Boole $\mathcal{A} \vee \mathcal{B}$.

intersection de toutes les algèbres qui contiennent $\mathcal{A}$ et $\mathcal{B}$

Exercice 2 : Vérifier l'identité

$$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C).$$

Donner la formule analogue pour $P(A_1 \cup A_2 \cup \cdots \cup A_n)$.

Exercice 3 : Une mesure de probabilité définit une (pseudo)-distance entre les événements de $\mathcal{F}$. En effet, considérons la différence symétrique de A et B,

$$A\Delta B := (A\setminus B) \cup (B\setminus A).$$

Donner la signification de cet événement. Exprimer $A\Delta B$ uniquement à l'aide des opérations "union", "intersection", "complémentation". Montrer que l'application

$$(A, B) \in \mathcal{F} \times \mathcal{F} \to P(A\Delta B)$$

satisfait les propriétés d'une distance, sauf une. Indiquer laquelle.

CORRIGÉ DE LA SÉRIE 2

Exercice 1 : Supposons que $x \in C\backslash(\cup_n A_n)$. Cela signifie que

$$x \in C \cap (\cup_n A_n)^c,$$

i.e. $x \in C$ et $x \notin \cup_n A_n$, donc en particulier $x \notin A_n$ $\forall n$. Par conséquent $x \in C\backslash A_n$ $\forall n$, ce qui donne

$$x \in \cap_n(C\backslash A_n).$$

Inversément, supposons que $x \in \cap_n(C\backslash A_n)$. Cela signifie que $x \in C\backslash A_n$ $\forall n$; donc $x \in C$ et $x \notin A_n$ $\forall n$, ce qui entraine $x \notin \cup_n A_n$ et par conséquent

$$x \in C \cap (\cup_n A_n)^c = C\backslash(\cup_n A_n).$$

Supposons que $x \in C\backslash(\cap_n A_n)$. Cela signifie que $x \in C$ et $x \notin \cap_n A_n$; par conséquent il existe m tel que $x \in A_m^c$, i.e. $x \in C \cap A_m^c = C\backslash A_m$. Donc

$$x \in \cup_n(C\backslash A_n).$$

Inversément, supposons que $x \in \cup_n(C\backslash A_n)$. Cela signifie que $x \in C\backslash A_m$ pour au moins un m; en particulier $x \notin A_m$, et donc $x \notin \cap_n A_n$, i.e. $x \in (\cap_n A_n)^c$, ce qui donne

$$x \in C\backslash(\cap_n A_n).$$

(a): pour tout A, B nous avons $A \cap B \subset A \subset A \cup B$. (b): si $A \subset B$, alors pour tout C nous avons $A \cap C \subset B \cap C$.
Supposons que $A \subset B$. Alors (a) et (b) donnent $A = A \cap A \subset A \cap B \subset A$, i.e. $A = A \cap B$. De même, $B \subset A \cup B = (A \cap B) \cup B = (A \cup B) \cap B \subset B$, i.e. $A \cup B = B$. $A^c \cup B = A^c \cup (A \cup B) = (A^c \cup A) \cup B = \Omega$. Par la formule de De Morgan $A^c \cup B = \Omega$ implique $(A^c \cup B)^c = \Omega^c = \emptyset = A \cap B^c$. Finalement $A = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c) = A \cap B \subset B$.

Le suprémum de la famille $\{A_n\}_n$ est $\cup_n A_n$ et l'infimum est $\cap_n A_n$. En effet $\forall m$

$$\cap_n A_n \subset A_m \subset \cup_n A_n.$$

Supposons que l'on ait un C et un B tels que $\forall m$

$$C \subset A_m \subset B.$$

Alors, nécessairement $C \subset \cap_m A_m$ et $\cup_m A_m \subset B$.

Nous démontrons une seule des relations. $x \in f^{-1}(\cap_n A_n)$ signifie $f(x) \in \cap_n A_n$, i.e. $f(x) \in A_n$ $\forall n$. Ceci signifie que $x \in f^{-1}(A_n)$ $\forall n$, i.e. $x \in \cap_n f^{-1}(A_n)$.

$\emptyset \in \mathcal{A}$ car $\Omega^c \in \mathcal{A}$. On a par induction $\cup_{j=1}^k A_j \in \mathcal{A}$. Donc aussi $\cup_{j=1}^k(\Omega\backslash A_j) \in \mathcal{A}$, i.e. par la formule de De Morgan $\Omega\backslash(\cap_{j=1}^k A_j) \in \mathcal{A}$, ce qui entraine $\cap_{j=1}^k A_j \in \mathcal{A}$.
On peut avoir $C^b = \emptyset$, mais on va montrer que chaque $\omega \in \Omega$ est contenu dans un ensemble C^b. En effet, $\forall i$ $\omega \in A_i$ ou $\omega \in A_i^c$. Montrons $C^b \cap C^{b'} = \emptyset$; si $b \neq b'$, alors il existe au moins un j tel que $b_j = 1$ et $b'_j = -1$ ou vice-versa. Enfin, $\forall i$, $C^b \subset A_i$ si $b_i = 1$ et sinon $C^b \cap A_i = \emptyset$. D'autre part si $\omega \in A_i$ il existe C^b tel que $\omega \in C^b$; donc

$$A = \bigcup_{\substack{b: \\ C^b \cap A \neq \emptyset}} C^b.$$

PROBABILITÉ ET STATISTIQUE

SÉRIE 4

Exercice 1 : Pendant l'examen de MMP on pose la question:

- Parmi les m affirmations suivantes, blablabla ..., une et une seule est vraie. Quelle est cette affirmation?

On suppose que l'élève connait la réponse avec probabilité p, $0 < p < 1$. Si l'élève ne connait pas la réponse, l'élève répond "au hasard". Calculer la probabilité que l'élève connaissait la réponse sachant que l'élève a donné la réponse correcte.

Exercice 2 : Loterie "6 de 36". Chaque participant choisit 6 nombres différents parmi les 36 nombres $\{1, 2, \ldots, 36\}$. On tire "au hasard" 6 nombres différents. On gagne si 3 des nombres au moins de son choix coincident avec la série des 6 nombres tirés. Calculer la probabilité de gagner.

Exercice 3 : Trois joueurs a, b et c jouent tour à tour l'un contre l'autre une partie d'échecs, de la façon suivante: c ne joue pas et a joue contre b; le gagnant joue contre c et le perdant ne joue pas et ainsi de suite. On suppose que chaque participant gagne sa partie d'échecs avec probabilité $1/2$ et qu'il n'y a pas de partie nulle. Les parties s'arrêtent dès qu'un des joueurs gagne successivement deux fois. Calculer les probabilités que les parties s'arrêtent parce que a a gagné deux fois de suite, b a gagné deux fois de suite, respectivement c a gagné deux fois de suite. Calculer la probabilité que l'on joue au moins k parties.

CORRIGÉ DE LA SÉRIE 3

Exercice 1 : $\mathcal{A} \cup \mathcal{B} = \{ C \subset \Omega : C \in \mathcal{A}$ ou $C \in \mathcal{B} \}$. Si A est un atome de $\mathcal{A}$ et si B est un atome de $\mathcal{B}$ en général $A \cap B \notin \mathcal{A} \cup \mathcal{B}$; la réponse est non. Par contre $\cup_{n \geq 1} \mathcal{A}_n$ est une algèbre de Boole. $\Omega \in \cup_{n \geq 1} \mathcal{A}_n$; si $A \in \cup_{n \geq 1} \mathcal{A}_n$ il existe m tel que $A \in \mathcal{A}_m$, et donc $A^c \in \cup_{n \geq 1} \mathcal{A}_n$; finalement si $A \in \cup_{n \geq 1} \mathcal{A}_n$ et $B \in \cup_{n \geq 1} \mathcal{A}_n$, alors à cause de la monotonicité il existe m tel que $A \in \mathcal{A}_m$ et $B \in \mathcal{A}_m$. Donc $A \cup B \in \cup_{n \geq 1} \mathcal{A}_n$.

Notons les atomes de $\mathcal{A}$ par A_i, $i = 1, \ldots, p$ et ceux de $\mathcal{B}$ par B_j, $j = 1, \ldots, q$. $\mathcal{A} \vee \mathcal{B}$ doit contenir les ensembles $C_{i,j} := A_i \cap B_j$, $i = 1, \ldots, p$ et $j = 1, \ldots, q$. Par conséquent $\mathcal{A} \vee \mathcal{B}$ est l'algèbre dont les atomes sont exactement ces ensembles $C_{i,j}$.

Exercice 2 : On a $P(A \cup B) + P(A \cap B) = P(A) + P(B)$. Utilisant cette identité nous avons

$$\begin{aligned}
P((A \cup B) \cup C) &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \\
&= P(A) + P(B) - P(A \cap B) + P(C) - P((A \cap C) \cup (B \cap C)) \\
&= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).
\end{aligned}$$

Posons

$$S_k(n) := \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}).$$

Nous avons l'identité

$$P(A_1 \underset{\cup}{\cancel{\cap}} A_2 \underset{\cup}{\cancel{\cap}} \cdots \underset{\cup}{\cancel{\cap}} A_n) = \sum_{j=1}^{n} (-1)^{j+1} S_j(n).$$

Exercice 3 : On a

$$A \setminus B = A \cap B^c \implies A \Delta B = (A \cap B^c) \cup (B \cap A^c).$$

$A \Delta B$ est réalisé si et seulement si A est réalisé et B ne l'est pas, ou si B est réalisé et A ne l'est pas. Ou encore, $A \Delta B$ est réalisé si et seulement si A ou B est réalisé, mais pas les deux simultanément. On a

$$P(A \Delta A) = P(\emptyset) = 0.$$
$$P(A \Delta B) = P(B \Delta A).$$
$$A \Delta B \subset (A \Delta C) \cup (B \Delta C).$$

En effet,

$$\begin{aligned}
A \Delta B = &\left((A \cap (C^c \cup C)) \cap B^c \right) \cup \left((B \cap (C^c \cup C)) \cap A^c \right) = \\
&\left([(A \cap C^c) \cup (A \cap C)] \cap B^c \right) \cup \left([(B \cap C^c) \cup (B \cap C)] \cap A^c \right) = \\
&(A \cap C^c \cap B^c) \cup (A \cap C \cap B^c) \cup (B \cap C^c \cap A^c) \cup (B \cap C \cap A^c) \subset \\
&\left(\underbrace{(A \cap C^c) \cup (C \cap A^c)}_{= A \Delta C} \right) \cup \left(\underbrace{(B \cap C^c) \cup (C \cap B^c)}_{= B \Delta C} \right).
\end{aligned}$$

Ceci entraîne

$$P(A \Delta B) \leq P(A \Delta C) + P(B \Delta C).$$

En général ce n'est pas vrai que $P(A) = 0$ implique $A = \emptyset$.

Série 4, exercice 1 : Axiomatisation

Soit l'espace $(\Omega, \mathcal{F}, p)$.

Soit les événements : $C =$ " l'élève connaît la réponse " , $C \in \mathcal{F}$, $p(C) = p$

$R =$ " donne la réponse correcte " , $R \in \mathcal{F}$

$\rightarrow$ on vérifie bien que : $C^c, R^c \in \mathcal{F}$, $C \cup R \in \mathcal{F}$

On utilise la formule de Bays pour calculer l'événement $C|R$ (C sachant R).

$$P(C|R) = \frac{P(C \cap R)}{P(R)}$$

On forme une partition de Ω avec : $\{C, C^c\}$; $\begin{cases} C \cup C^c = \Omega \\ C \cap C^c = \emptyset \end{cases}$

$$P(C|R) = \frac{P(R|C) \cdot P(C)}{P(R|C) \cdot P(C) + P(R|C^c) \cdot P(C^c)}$$

$$= \frac{1 \cdot p}{1 \cdot p + p(R|C^c) \cdot (1-p)}$$

Avec $R|C^c$ qui représente l'événement : donne la réponse correcte au hasard. Pour calculer cet proba. d'événement on pose $\Omega = \{m_1 \dots m_m\}$ les épreuves qui sont les réponses , $A \in \mathcal{F}$ l'événement : donne la réponse correcte. Comme $card(\Omega) < \infty$, $card(\Omega) = m$, alors on a un espace de proba. discret : donc :

$$P(R|C^c) = \frac{card(A)}{card(\Omega)} = \frac{\sum_{\omega \in A} P(\omega)}{\sum_{\omega \in \Omega} P(\omega)} = \frac{1}{m}$$

Conclusion : $\boxed{P(C|R) = \dfrac{p}{p + \frac{1-p}{m}}}$

#

Exercice 2, série 4 : axiomatisation : soit l'esp. de proba. $(\Omega, \mathcal{F}, p)$

Ω = ensemble des possibles = $\{(1,2,3,4,5,6);(1,2,3,4,5,7);\dots\} = \{\omega_1,\dots\omega_n\}$

$\mathcal{F}$ = alg. de Boole / σ-algèbre représentés par des événements $A_i \in \mathcal{F}$, A_i sont représentés par des épreuves élémentaires ω_i

p = mesure de probabilité : application : $p : \Omega \to \mathbb{R}$ t.q. i) $P(\Omega) = 1$

ii) $P(A) \in [0,1] \quad \forall A \in \mathcal{F}$

iii) $A \cap B = \emptyset \Rightarrow P(A \cap B) = P(A) + P(B) \quad \forall A, B \in \mathcal{F}$

iv) $A_0 \supset A_1 \supset \dots \quad \bigcap_{n \geq 0} A_n = \emptyset$, alors : $\lim_{n\to\infty} P(A_n) = 0$

(la condition iv) est la continuité de la proba, qui est équivalente à la σ-additivité)

Ici : – $\mathrm{card}(\Omega) = \binom{36}{6}$ = nb. de tirages de 6 boules parmi 36 sans compter l'ordre

– comme $\mathrm{card}(\Omega) < \infty$, si on prend $\mathcal{F} = \mathcal{P}(\Omega)$, alors on a un espace de probabilité discret.

– soit $A \in \mathcal{F}$ l'événement "gagné : 3 nb. au moins coïncident avec les 6 boules"

– hypothèse d'équiprobabilité

$$\Longrightarrow \quad P(A) = \frac{\sum_{\omega \in A} P(\omega)}{\sum_{\omega \in \Omega} P(\omega)} = \frac{\mathrm{card}(A)}{\mathrm{card}(\Omega)} = \frac{1}{\binom{36}{6}} \cdot \mathrm{card}(A)$$

Calcul de card(A) : – K = nb. de boules parmi les 6 choisies. Soit : $\bar{\omega} = (\bar{a}_1, \bar{a}_2, \bar{a}_3, \bar{a}_4, \bar{a}_5, \bar{a}_6)$ le tirage des 6 boules choisies, $\omega_i = (a_{i1}, a_{i2}, a_{i3}, a_{i4}, a_{i5}, a_{i6})$ un tirage qqe.

– Soit l'événement :

$$\mathcal{F} \supset E_k = \{\omega_i \in \Omega \text{ t.q. } \mathrm{card}(\bar{\omega} \cap \omega_i) = k\}$$

$\to$ E_k représente le nb. de Boules correctes tirées.

$\to$ E_k sont des événements ~~indépendants~~ disjoints : $E_i \cap E_j = \emptyset$

$\to$ $\mathrm{card}(A) = \mathrm{card}(E_3 \cup E_4 \cup E_5 \cup E_6) = \sum_{i=3}^{6} \mathrm{card}(E_i)$

$$\mathrm{card}(E_3) = \binom{6}{3} \cdot \binom{30}{3}$$
$$\mathrm{card}(E_4) = \binom{6}{4} \cdot \binom{30}{2}$$
$$\mathrm{card}(E_5) = \binom{6}{5} \cdot \binom{30}{1}$$
$$\mathrm{card}(E_6) = \binom{6}{6} \cdot \binom{30}{0}$$

Conclusion :

$$\boxed{P(A) = \frac{\sum_{i=3}^{6} \binom{6}{i} \cdot \binom{30}{6-i}}{\binom{36}{6}}}$$

Exercice 3, série 4 : Axiomatisation

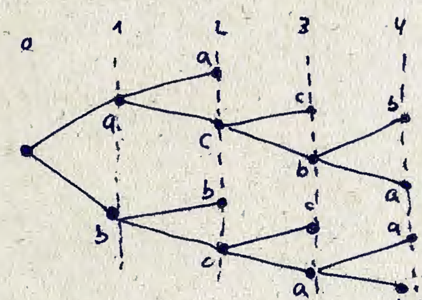
Soit : $(\Omega, \mathcal{F}, P)$, avec :

$\Omega = \{(a,a), (a,b,b), \dots\}$ = "ensemble des suites possibles" , a, b, c : épreuves élémentaires a, b, c resp. a gagné 1 fois d'affilée.

$\mathcal{F}$ = σ-algèbre contenant des évènements $\in \mathcal{F}$ qui sont des sous-ens. de Ω.

P = mesure de probabilité

Illustration :



Soit : - $A, B, C \in \mathcal{F}$ les évènements A, B, C resp. gagnent 2 fois de suite.

→ on va voir que : $\{A,B,C\} \equiv$ partition de Ω , i.e. $\nexists$ de suite infinie $\in \Omega$

→ $|\Omega|$ dénombrable $\Longrightarrow$ esp. prob. discret :

$$P(A) = \frac{\sum_{\omega \in A} P(\omega)}{\sum_{\omega \in \Omega} P(\omega)} = \frac{|A|}{|\Omega|}$$

— indépendance des parties et équiprobabilité $\Rightarrow$ la proba. d'une chaîne de longueur k est : $\left(\frac{1}{2}\right)^k$

$$\sum_{\omega\, \text{fini} \in \Omega} p(\omega) = 2 \cdot \sum_{k \geq 2} \left(\frac{1}{2}\right)^k = 2 \cdot \left\{ -(+1 + 1/2) + \sum_{k \geq 0} \left(\frac{1}{2}\right)^k \right\} = 1$$

$\Longrightarrow p(\omega\, \text{infini}) = 0$ (ens. négligeable)

Conclusion : → $\Omega = A \cup B \cup C$ de manière disjointe

→ $P(\Omega) = 1 = P(A) + P(B) + P(C)$

→ équiprobabilité : $P(A) = P(B)$

$$\Longrightarrow P(A) = \underbrace{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^3 + \dots}_{\text{arbre supérieur}} + \underbrace{\left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^3 + \dots}_{\text{arbre inférieur}}$$

$$= \sum_{k \geq 0} \left(\frac{1}{8}\right)^k \cdot \left[\frac{1}{4} + \frac{1}{16}\right] = \frac{5}{16} \cdot \frac{8}{7} = \frac{5}{14}$$

$\Longrightarrow P(A) = P(B) = \frac{5}{14}$ #

$\Longrightarrow P(C) = 1 - 2P(A) = 1/7$

• E = "joue au minimum k parties" , $E \in \mathcal{F}$, $P(E) = \sum_{\omega \in E} p(\omega)$

$$P(E) = 2 \cdot \left(\frac{1}{2}\right)^{k-1} = \frac{1}{2^{k-2}} \quad \#$$

PROBABILITÉ ET STATISTIQUE

SÉRIE 5

√ **Exercice 1 :** On considère n boules numérotées $1, 2, \ldots, n$, que l'on aligne par ordre croissant de $i = 1, 2, \ldots, n$ de gauche à droite. On fait une permutation "au hasard" de ces boules (les permutations sont équiprobables). Quelle est la probabilité p_n de l'événement E, "aucune boule ne se retrouve à sa place initiale"? Déterminer $\lim_{n\to\infty} p_n$. Quelle est la probabilité que k et seulement k boules soient non permutées?
Indication. Calculer la probabilité de E_k, "la boule k se retrouve à sa place initiale". Calculer la probabilité de l'événement E^c en utilisant l'ex. 2, série 3.

Exercice 2 : On lance n fois de façon "indépendante" une pièce de monnaie équilibrée, i.e. à chaque fois la probabilité d'obtenir pile est $1/2$. Le résultat de l'expérience est k piles et $n - k$ faces. Par définition une série de k piles consiste en une succession de k piles exactement. Le résultat de l'expérience est décrit complètement en indiquant successivement la longueur des séries en commençant par celle de faces, puis de piles, de faces etc, la première série de faces pouvant être de longueur zéro. Quelle est la probabilité d'avoir observé exactement $r \geq 1$ séries de piles ?

CORRIGÉ DE LA SÉRIE 4

Exercice 1 : A est l'événement "l'élève connait la réponse"; B est l'événement "l'élève donne la bonne réponse".

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(B|A)\,P(A)}{P(B|A)\,P(A) + P(B|A^c)\,P(A^c)}$$

= formule de Bayes.

$$= \frac{1\cdot p}{1\cdot p + (1/m)\cdot(1-p)}\,.$$

Exercice 2 : On pose

car on ne tient pas compte de l'ordre ici

l'espace Ω n'a même pas été introduit.

$$\Omega := \{\,\omega = (\omega_1, \dots, \omega_6) : \omega_j \in \{1, 2, \dots, 36\}\,,\ \omega_j \neq \omega_i\,,\ j \neq i\,\}\,.$$

Ω est isomorphe à l'ensemble des sous-ensembles à 6 éléments de $\{1, 2, \dots, 36\}$. La cardinalité de Ω est $|\Omega| = \binom{36}{6}$. Chaque $\omega \in \Omega$ est équiprobable. Notons $\overline{\omega} = (\overline{\omega}_1, \dots, \overline{\omega}_6)$ mon choix. L'événement E "je gagne" est la réunion disjointe des événements E_3, E_4, E_5 et E_6, où E_k est l'événement

$$E_k := \{\,\omega \in \Omega : |\{\omega_1, \dots, \omega_6\} \cap \{\overline{\omega}_1, \dots, \overline{\omega}_6\}| = k\,\}\,.$$

card

La probabilité de E_k est

$$P(E_k) := \frac{\binom{6}{k}\binom{30}{6-k}}{\binom{36}{6}}\,.$$

On obtient $P(E) = \sum_{k=3}^{6} P(E_k) \simeq 0.0416 + 0.0033 + 0.000092 + 0.0000005 \simeq 0.0451$.

Exercice 3 : Pour Ω on peut prendre l'ensemble des suites <u>finies</u> de lettres

axiomatisation

$$aa,\ acc,\ acbb,\ acbaa,\ acbacc,\ \dots$$
$$bb,\ bcc,\ bcaa,\ bcabb,\ bcabcc,\ \dots$$

et des deux suites <u>infinies</u> → doit passer à la σ-algèbre.

$$acbacbacbacb\cdots\,,\quad bcabcabcabca\cdots\,.$$

La suite $acbaa$ signifie: a gagne, c gagne, b gagne, a gagne, a gagne et par conséquent les parties s'arrêtent. D'après les règles admises la probabilité de cette suite $acbaa$ est $1/32$. De manière générale la probabilité d'une suite ω de k lettres est $p(\omega) = 2^{-k}$. Nous avons

branche supérieure + branche inférieure

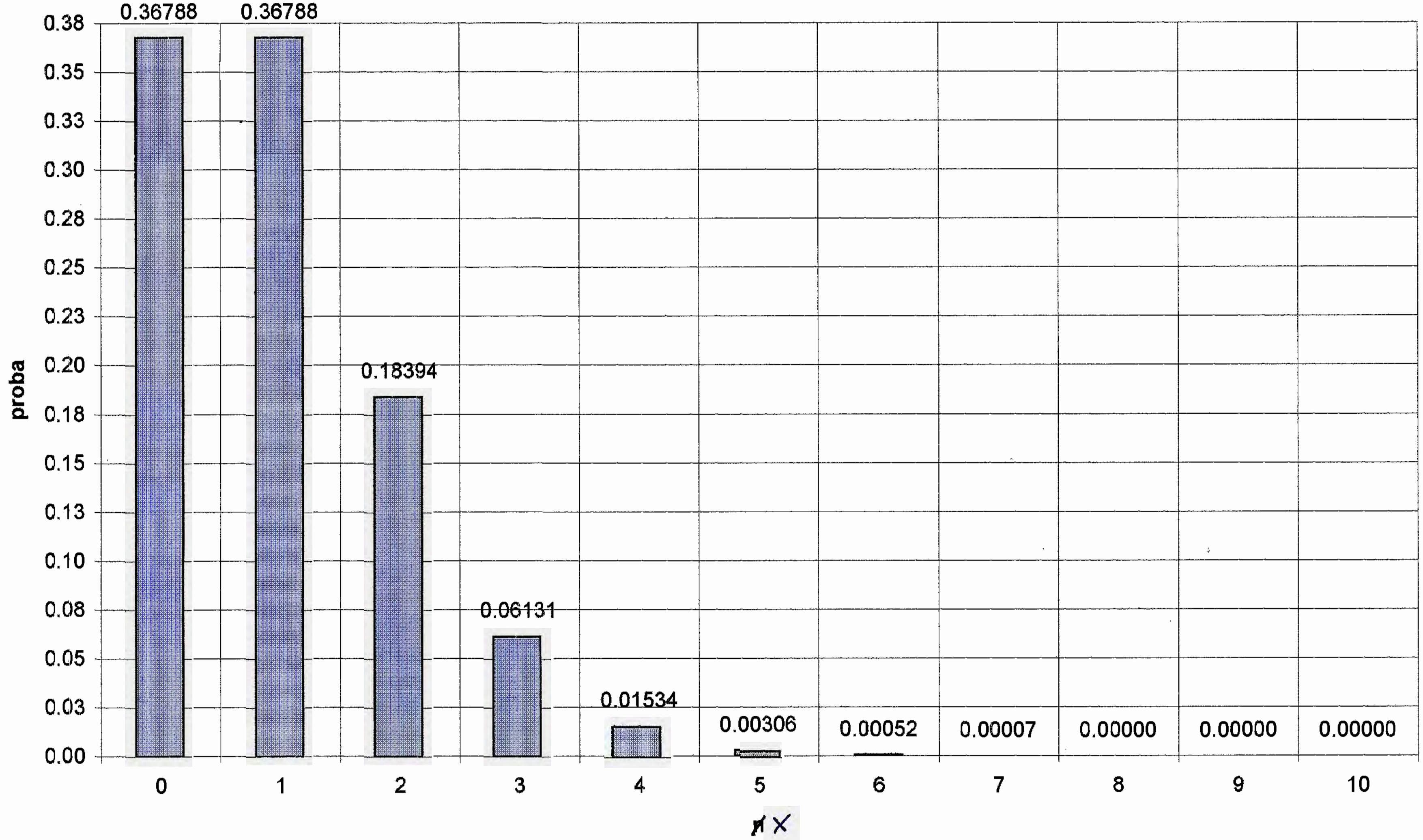
$$\sum_{\substack{\omega:\\ \text{s. finie}}} p(\omega) = 2\sum_{k\geq 2} 2^{-k} = 1\,.$$

Si ω est une des deux suites infinies $p(\omega) = 0$. La probabilité que les parties s'arrêtent parce que a a gagné deux fois de suite est la même que la probabilité que les parties s'arrêtent parce que b a gagné deux fois de suite. Cette probabilité est

$$\frac{1}{4} + \frac{1}{4}\sum_{k\geq 1} 8^{-k} + \frac{1}{16} + \frac{1}{16}\sum_{k\geq 1} 8^{-k} = \frac{5}{14}\,.$$

La probabilité que c gagne deux fois successivement est $2/7$. La probabilité de jouer au moins k parties, si $k \geq 2$, est égale à la somme des probabilités de jouer exactement j parties, $j \geq k$, i.e. $2\cdot 2^{-k}\sum_{j\geq 0} 2^{-j} = 2^{-k+2}$. Cette probabilité est nulle si $k < 2$.

Soit une permutation de N boules (N = 10). Le graphe montre la probabilité pour que x boules soient non permutées. La somme des probabilités donne bien 1.



Exercice 1, série 5 : en rapide

Soit l'espace $(\Omega, \mathcal{F}, P)$.

$\Omega = \{\omega_1, \dots, \omega_{n!}\}$, $\omega_i = (\omega_{i1}, \dots, \omega_{in})$ une permutation de boules

Ω = ensemble des possibles (épreuves)

Soit les événements de $\mathcal{F}$:

$$\begin{cases} A_k = \text{"la boule } k \text{ reste à sa place"}, \; k = 1, \dots, n \\ \left(\bigcup_{k=1}^{n} A_k\right) = \text{"au moins 1 boule reste à sa place"} \\ \left(\bigcup_{k=1}^{n} A_k\right)^c = E = \text{"aucune boule ne se retrouve à sa place initiale"} \end{cases}$$

Comme $\operatorname{card}(\Omega) < \infty$, alors on est dans un esp. de proba. discret, $\operatorname{card}(\Omega) = \operatorname{card}(S_n) = n!$. Donc :

$$\begin{cases} P(A_k) = \dfrac{\operatorname{card}(A_k)}{\operatorname{card}(\Omega)} = \dfrac{\sum_{\omega \in A_k} p(\omega)}{\sum_{\omega \in \Omega} p(\omega)} = \dfrac{1}{n!} \sum_{\omega \in A_k} p(\omega) = \dfrac{|A_k|}{n!} = \dfrac{(n-1)!}{n!} \\ P\left(\bigcap_{j=1}^{k} A_k\right) = \dfrac{(n-k)!}{n!} \end{cases}$$

Avec :

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{k=1}^{n} (-1)^{k+1} \cdot \sum_{1 \leq j_1 < \dots < j_k < \dots \leq n} P(A_{j_1} \cap \dots \cap A_{j_k})$$

$$= \sum_{k=1}^{n} P(A_k) - \sum_{j_1 < j_2}^{n} P(A_{j_1} \cap A_{j_2}) + \sum_{j_1 < j_2 < j_3}^{n} P(A_{j_1} \cap A_{j_2} \cap A_{j_3}) - \dots$$

$$\binom{n}{1} \frac{(n-1)!}{n!} - \binom{n}{2} \frac{(n-2)!}{n!} + \binom{n}{3} \frac{(n-3)!}{n!} - \dots$$

par la série 1, exercice 2

$$\binom{n}{r} = \sum_{j=r}^{n} \binom{j-1}{r-1}$$

$$= \sum_{k=1}^{n} (-1)^{k+1} \frac{(n-k)!}{n!} \binom{n}{k} = -\sum_{k=1}^{n} \frac{(-1)^k}{k!}$$

$$P(E) = 1 + \sum_{k=1}^{n} \frac{(-1)^k}{k!} = \sum_{k \geq 0}^{n} \frac{(-1)^k}{k!}$$

- $\lim_{n \to \infty} P(E) = \frac{1}{e}$
- $\tilde{E}$ = "k et seulement k boules non permutées." , $\tilde{E} \in \mathcal{F}$

  $\tilde{E} \iff$ "(n-k) boules permutées)"

  $\operatorname{Card}(\Omega) = n < \infty \implies$ espace de probabilité discret $\implies P(\tilde{E}) = \dfrac{\operatorname{card}(\tilde{E})}{\operatorname{card}(\Omega)} = \dfrac{1}{n!} |\tilde{E}|$

Calcul :
- $\binom{n}{k}$ choix de k boules non permutées parmi n
- déjà calculé : $P(\text{"n-k boules permutées"})^{(*)} = \dfrac{|U|}{(n-k)!} \longrightarrow |U| = (n-k)! \cdot P(\text{"..."})$

  ~~parmi n-k~~ ~~avec~~ $\approx U$

$$\implies P(\tilde{E}) = \frac{1}{n!} \binom{n}{k} \cdot (n-k)! \cdot P(\text{"..."}) = \frac{\cancel{n!}}{k! \cancel{(n-k)!}} \cdot \frac{1}{\cancel{n!}} \cdot \cancel{(n-k)!} \cdot P(\text{"..."}) = \frac{P(\text{"..."})}{k!}$$

$$P(\tilde{E}) = \frac{1}{k!} \sum_{u=0}^{n-k} \frac{(-1)^u}{u!}$$

#

(*) $P(\text{"n-k boules permutées"})$ revient à calculer $P(E)$ avec dans la formule $n-k$ au lieu de n.

Exercice 2, série 5 : "axiomatisation"

Soit l'espace $(\Omega, \mathcal{F}, P)$, avec :

$\Omega = \{\omega_1, \dots, \omega_{\binom{n}{k}}\}$, $\omega_i = (\omega_{i1}, \dots, \omega_{in})$, $\omega_{ij} \in \{0, 1\}$

Ω = ensemble des possibles = toutes les épreuves possibles après avoir lancé n fois 1 pièce de monnaie

ω_{ij} : 1 lancé de pièce

$$\begin{cases} card(\Omega) = n! < \infty \implies \text{espace de prob. discret} \\ \mathcal{F} : \text{algèbre de Borel} \\ E_r = \text{" } r \geq 1 \text{ séries de piles "}, \ E_r \in \mathcal{F} \end{cases}$$

$$\Rightarrow P(E_r) = \frac{\sum_{\omega \in E_r} P(\omega)}{\sum_{\omega \in \Omega} P(\omega)} = \frac{card(E_r)}{\binom{n}{k}}$$ (fin de parenthèse introductive)

Reformulation : – pour résoudre plus facilement notre exercice : (reformulation pour calculer card(E_r))
– on observe :

$$\{f_1, p_1, f_2, p_2, \dots, p_r, f_{r+1}\}, \quad \sum_{i=1}^{r} (p_i + f_i) + f_{r+1} = n$$

→ ~~dans cet espace de prob.~~ ainsi, il est plus facile de calculer card(E_r) : on observe donc k piles, donc :

Calc. : $|E_r|$: $\sum_{i=1}^{r} p_i = k$; $p_i > 0$ (i)

$\sum_{i=1}^{r+1} f_i = n-k$; cas à distinguer (ii)

(i) Connaît ce # : $\#_p = \binom{k-1}{r-1}$

(ii) $$\begin{cases} f_1 = f_{r+1} = 0 : \sum_{i=2}^{r} f_i = n-k \iff \sum_{i=1}^{r-1} f_i = n-k \implies \#_{f_1} = \binom{n-k-1}{r-2} \\ f_1 \neq 0, f_{r+1} \neq 0 : \sum_{i=1}^{r+1} f_i = n-k \iff \#_{f_2} = \binom{n-k-1}{r} \\ \{f_1 = 0 ; f_{r+1} \neq 0\} ; \{f_1 \neq 0 ; f_{r+1} = 0\} \quad 2 \cdot \sum_{i=1}^{r} f_i = n-k \implies \#_{f_3} = 2\binom{n-k-1}{r-1} \end{cases}$$

$\Rightarrow |E_r| = \#_p(\#_{f_1} + \#_{f_2} + \#_{f_3})$

Calcul de : card(Ω) : – c'est le nb. de possibilités de tirer k piles parmi n : card$(\Omega) = \binom{n}{k}$

$$\boxed{P(E_r) = \frac{\#_p(\#_{f_1} + \#_{f_2} + \#_{f_3})}{\binom{n}{k}}}$$

~~On peut encore expliquer $\#_p(\#_{f_1} + \#_{f_2} + \#_{f_3})$~~ :

EPFL
Département de mathématiques

Physique, 2ème année

PROBABILITÉ ET STATISTIQUE

SÉRIE 6

$\omega \in \infty\ A_m$

$\omega \in A_m\ \forall m$ sauf du nb. $< \infty$ de A_m

Exercice 1 :Soient $A_n \subset \Omega$, $n \geq 1$. On pose

$$\boxed{\limsup_n A_n := \bigcap_{n\geq 1} \bigcup_{m\geq n} A_m \quad , \quad \liminf_n A_n := \bigcup_{n\geq 1} \bigcap_{m\geq n} A_m .}$$

À quel événement correspond $\limsup_n A_n$, resp. $\liminf_n A_n$? L'**indicatrice** I_A de l'événement A est la fonction

$$I_A(\omega) := \begin{cases} 1 & \omega \in A, \\ 0 & \omega \notin A. \end{cases}$$

Donner les indicatrices de $\limsup_n A_n$ et $\liminf_n A_n$. Supposons que $A_n = A$ si n est pair et $A_n = B$ si n est impair. Déterminer $\liminf_n A_n$ et $\limsup_n A_n$. Démontrer les inégalités

$$P(\liminf_n A_n) \leq \liminf_n P(A_n) \leq \limsup_n P(A_n) \leq P(\limsup_n A_n) .$$

→ permet de comprendre que convergence p.s $\Rightarrow$ convergence (*) en probabilité

p.s.: $P(\limsup_n \{|X_n - X| \geq \varepsilon\}) = 0$

$\Rightarrow$ c.p. $\lim_n P(|X_n - X| \geq \varepsilon) = 0$

Démontrer l'affirmation

$$\sum_{n \geq 1} P(A_n) < \infty \Longrightarrow P(\limsup_n A_n) = 0 .$$

(Borel-Cantelli 1)

donc avec ces inégalités, comme $\limsup$ est tjs $\geq$ aux autres $\lim$, on a (*)

Exercice 2 : Calculer l'espérance de la v.a. réelle $Y = X^2$ si la fonction de répartition de X est

$$F(t) := \begin{cases} 0 & t < 0, \\ 1/4 & t \in [0,2), \\ 2/3 & t \in [2,7), \\ 4/5 & t \in [7,12), \\ 1 & t \geq 12. \end{cases}$$

Même question pour $Y = X$, $Y = X^2$ et $Y = X^3 + 1$, si la fonction de répartition de X est

$$F(t) := \begin{cases} 0 & t < -2, \\ (t+2)/3 & t \in [-2,-1), \\ 1/2 & t \in [-1,0), \\ t^3 + 1/2 & t \in [0,1/2), \\ 1 & t \geq 1/2. \end{cases}$$

Exercice 3 : Soit X une variable aléatoire prenant les valeurs $0, 1, 2, \ldots$. Vérifier l'identité

$$\mathrm{E}(X) \equiv \sum_{n\geq 0} n \cancel{X(n)} = \sum_{k\geq 0} P(\{X > k\}) .$$

$P(X = n)$

CORRIGÉ DE LA SÉRIE 5

Exercice 1 :On doit calculer $1 - P(\cup_{k\geq 1} E_k)$; ceci peut se faire par la formule de l'ex. 2 série 3. Calculons E_k. La k^e boule se retrouve à sa place initiale, la place des autres boules est arbitraire. Il y a $(n-1)!$ permutations qui laissent k invariant. Donc

$$P(E_k) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

Pour utiliser la formule de l'ex. 2 série 3 il faut calculer

$$P(E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_k}).$$

Le nombre de permutations qui laissent k indices déterminés invariants est $(n-k)!$. Par conséquent

$$P(E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_k}) = \frac{(n-k)!}{n!} = \frac{1}{n(n-1)\cdots(n-k+1)}.$$

Nous obtenons (cf ex. 2 série 3)

$$P(E^c) = \sum_{k=1}^{n} (-1)^{k+1} S_k(n) \quad , \quad S_k(n) = \binom{n}{k} \frac{1}{n(n-1)\cdots(n-k+1)} = \frac{1}{k!},$$

et

$$P(E) = \sum_{k=0}^{n} (-1)^k \frac{1}{k!} \overset{n\to\infty}{\to} \frac{1}{\mathrm{e}} \simeq .36788.$$

Notons F_k l'événement "k et seulement k boules restent invariantes". Choisissons ces k boules. Ceci peut se faire de $\binom{n}{k}$ manières différentes. Une fois ce choix fait, aucune des autres $n-k$ boules ne revient à sa place initiale. D'après le calcul ci-dessus il y a

$$(n-k)!\Big(\sum_{k=0}^{n-k} (-1)^k \frac{1}{k!}\Big)$$

configurations possibles. Par conséquent

$$F_k = \binom{n}{k} (n-k)!\Big(\sum_{k=0}^{n-k} (-1)^k \frac{1}{k!}\Big) = \frac{1}{k!}\Big(\sum_{k=0}^{n-k} (-1)^k \frac{1}{k!}\Big) \simeq \frac{\mathrm{e}^{-1}}{k!}.$$

Exercice 2 : Ω est l'ensemble des mots de longueur n ayant k zéros et $n-k$ uns. Chaque mot est équiprobable. $|\Omega| = \binom{n}{k}$. Notons $a_1, a_2, \ldots a_{r+1}$ les longueurs des séries de faces, $a_1 \geq 0$, $a_j \geq 1$, $j = 1, \ldots r$, $a_{r+1} \geq 0$ telles que $\sum_{j=1}^{r+1} a_j = n-k$. Notons $b_1, \ldots, b_r$ les longueurs des séries de piles, $b_j \geq 1$, $j = 1, \ldots r$, telles que $\sum_{j=1}^{r} b_j = k$. Il y a $\binom{k-1}{r-1}$ suites $b_1, \ldots, b_r$ différentes. Le nombre des suites $a_1, \ldots, a_{r+1}$ différentes est égal au nombre de solutions positives de l'équation

$$\sum_{j=1}^{r+1} \overline{a}_j = n - k + 2.$$

(Poser $\overline{a}_1 = a_1 + 1$, $\overline{a}_{r+1} = a_{r+1} + 1$, sinon $\overline{a}_j = a_j$.) Le nombre de ces solutions est $\binom{n-k-1}{r}$, ce qui donne pour la probabilité cherchée

$$\frac{\binom{n-k+1}{r}\binom{k-1}{r-1}}{\binom{n}{k}}.$$

EPFL
Département de mathématiques

Physique, 2ème année

PROBABILITÉ ET STATISTIQUE

SÉRIE 7

résultat du cours à savoir

Exercice 1 : Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité. X et Y sont deux variables aléatoires réelles, étagées. La **fonction de répartition** de X et Y est définie sur R^2 par la formule

$$F_{XY}(t,s) := P(\{X \leq t, Y \leq s\}).$$

Donner et vérifier les propriétés de F.

p.10.

pas fait

Exercice 2 : On considère la même situation que dans l'exercice 1. Montrer que

$$\sigma^2(X) = \mathrm{E}\big((X - \mathrm{E}(X))^2\big).$$

idem.

Soit $\alpha > 1$ et β tel que $1/\alpha + 1/\beta = 1$. Vérifier les inégalités de Hölder et Minkowski

$$|\mathrm{E}(XY)| \leq \mathrm{E}(|X|^\alpha)^{1/\alpha}\,\mathrm{E}(|Y|^\beta)^{1/\beta}$$

$$\mathrm{E}(|X+Y|^\alpha)^{1/\alpha} \leq \mathrm{E}(|X|^\alpha)^{1/\alpha} + \mathrm{E}(|Y|^\alpha)^{1/\alpha}.$$

Exercice 3 : Vérifier l'identité

$$\sigma^2(X) = \mathrm{E}\big((X - \mathrm{E}(X))^2\big).$$

Dans quel cas la variance de X est nulle? Si $0 \leq X(\omega) \leq c$, montrer que

$$\sigma^2(X) \leq \frac{c^2}{4}.$$

Soit X ayant une variance finie. Montrer pour $t > 0$ que

$$P(\{\omega : |X(\omega) - \mathrm{E}(X)| \geq t\}) \leq \frac{\sigma^2(X)}{t^2}.$$

Exercice 4 : Est-ce vrai ou faux,

$$\lim_{a_n \to a} P(\{X < a_n\}) \stackrel{?}{=} P(\{X < a\}).$$

$\Omega = \{0,1\}$ $\quad \mathcal{F} = \mathcal{P}(\Omega)$ $\quad 1 - \frac{1}{n}$

$P(\{0\}) = \frac{1}{2}$, $P(\{1\}) = \frac{1}{2}$

$X(\omega) = \omega$

$P(\{X(\omega) < \frac{1}{n}\}) = \frac{1}{2}\ \forall n \quad \rightarrow \quad \lim P(\cdot) = \frac{1}{2}$

$P(\{X(\omega) < 0\}) = 0$

CORRIGÉ DE LA SÉRIE 6

Exercice 1 : Si $\omega \in \limsup_n A_n$, $\omega \in \cup_{m\geq n} A_m$ $\forall n$. En particulier il existe $m_1 \geq 1$ tel que $\omega \in A_{m_1}$; soit $n = m_1 + 1$; alors il existe $m_2 > m_1$ tel que $\omega \in A_{m_2}$. Nous obtenons: si $\omega \in \limsup_n A_n$ alors il existe une suite (dépendant de ω) $m_1 < m_2 < \ldots$, $m_k \to \infty$ lorsque $k \to \infty$, telle que $\omega \in A_{m_k}$ $\forall k$. Inversément si cette condition est satisfaite pour ω', alors $\omega' \in \limsup_n A_n$. On exprime ceci par: $\limsup_n A_n$ *est l'ensemble des* ω *contenus dans une infinité de* A_n. De façon similaire on a: $\liminf_n A_n$ *est l'ensemble des* ω *contenus dans tous les* A_n*, à l'exception d'un nombre fini d'entre eux.* Nous avons $I_{\cup_n A_n} = \sup_n I_{A_n}$, $I_{\cap_n A_n} = \inf_n I_{A_n}$, ce qui donne $I_{\liminf_n A_n} = \liminf I_{A_n}$, $I_{\limsup_n A_n} = \limsup I_{A_n}$. Si $A_n = A$ si n est pair et $A_n = B$ si n est impair, $\liminf_n A_n = A \cap B$, $\limsup_n A_n = A \cup B$. En utilisant la continuité de P pour les suites monotones,

$$\begin{aligned} P(\liminf_n A_n) &= \lim_n P(\cap_{m\geq n} A_m) \leq \liminf_n P(A_n) \\ &\leq \limsup_n P(A_n) \leq \lim_n P(\cup_{m\geq n} A_m) = P(\limsup_n A_n)\,. \end{aligned}$$

Enfin, si $\sum_{N\geq 1} P(A_n) < \infty$, nous avons

$$P(\limsup_n A_n) = \lim_n P(\cup_{m\geq n} A_m) \leq \lim_n \sum_{m\geq n} P(A_m) = 0\,.$$

Exercice 2 : $\mathrm{E}(X^2) = 0 \cdot 1/4 + 2^2 \cdot \overset{5}{\cancel{1}}/12 + 7^2 \cdot 2/15 + 12^2 \cdot 1/5 = \overset{37}{\cancel{107/3}}$. Dans le deuxième cas on peut décomposer $F = F_1 + F_2$, F_1 est constante par morceaux et F_2 est continue pour tout t.

$$F_1(t) := \begin{cases} 0 & t < -1 \\ 1/6 & t \in [-1, 1/2) \\ 13/24 & t \geq 1/2 \end{cases} \quad , \quad F_2(t) := \begin{cases} 0 & t < -2 \\ (t+2)/3 & t \in [-2, -1) \\ 1/3 & t \in [-1, 0) \\ t^3 + 1/3 & t \in [0, 1/2) \\ 11/24 & t \geq 1/2 \end{cases}$$

Si $Y = g \circ X$, nous avons

$$\mathrm{E}(Y) = g(-1) \cdot 1/6 + g(1/2) \cdot 3/8 + 1/3 \int_{-2}^{-1} g(s)\,ds + 3 \int_0^{1/2} g(s) s^2\,ds\,.$$

Exercice 3 : Posons $A_{i,j} := P(\{X = i\})$, $i, j \geq 0$. Nous avons

$$\sum_{n\geq 0} nX(n) = \sum_{n\geq 1} \sum_{k=1}^{n} A_{n,k} = \sum_{k\geq 1} \sum_{n\geq k} A_{n,k} = \sum_{k\geq 1} \sum_{n\geq k} P(\{X = n\})\,.$$

Par conséquent

$$\sum_{n\geq 0} nX(n) = \sum_{k\geq 0} \sum_{n>k} P(\{X = n\}) = \sum_{k\geq 0} P(\{X > k\})\,.$$

$$\sum_{k\geq 0} P(\{X>k\}) = \sum_{k\geq 0} \sum_{j>k} P(X=j) = \sum_{k,j\geq 0} \mathbb{I}_{\{k<j\}} P(X=j) = \sum_{j\geq 0} P(X=j) \sum_{0\leq k<j} 1$$
$$= \sum_{j\geq 0} j \cdot P(X=j)$$

PROBABILITÉ ET STATISTIQUE

SÉRIE 8

Exercice 1 : Soit $X_1, \ldots, X_n$ n variables aléatoires réelles. La **covariance** de X_1 et X_2 est définie par

$$\mathrm{Cov}(X_1, X_2) := \mathbb{E}(X_1 X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2)$$

si $\mathbb{E}(|X_1 X_2|) < \infty$, $\mathbb{E}(|X_1|) < \infty$ et $\mathbb{E}(|X_2|) < \infty$. La matrice de covariance de $X_1, \ldots, X_n$ est par définition la matrice $n \times n$ $(\mathrm{Cov}(X_i, X_j))_{i,j}$. Vérifier que cette matrice est non-négative. $\mathrm{Cov}(X_1, X_2)$ est aussi applée fonction de corrélation en mécanique statistique.

Exercice 2 : On considère un gas libre de n particules classiques (distinguables) dans un cube $V \subset \mathbb{R}^3$ de côté R. Une configuration de ce gas est donnée par $\omega = (\omega_1, \cdots, \omega_n) \in \mathbb{R}^{3n}$, t.q. $\forall i$, $\omega_i \in V$. (On ne s'intéresse ici qu'à la position de ces particules dans V). L'hypothèse que le gas est libre se traduit par le fait que chaque configuration est équiprobable. Cela signifie que la distribution de probabilité sur l'espace des phases $\Omega := \{ (\omega_1, \cdots, \omega_n) \in \mathbb{R}^{3n} : \omega_i \in V \ , \forall i \}$ est la distribution uniforme de densité φ,

$$\varphi(\omega) := \begin{cases} \overset{1/R^{3n}}{\cancel{1\mathbb{R}^{3n}}} & \text{si } \omega \in \Omega\,, \\ 0 & \text{sinon.} \end{cases}$$

Soit $Q \subset V$ un sous-ensemble de volume $|Q| > 0$. On définit la variable aléatoire N_Q, $N_Q(\omega) := |\{ j : \omega_j \in Q \}|$, qui compte le nombre de particules dans Q pour la configuration ω. Déterminer la distribution de probabilité de N_Q et calculer son espérance. i) Déterminer la distribution de cette même variable aléatoire lorsque l'on ii) prend la *limite thermodynamique*, $R \to \infty$, $n \to \infty$ de sorte que $n/R^3 \to \lambda > 0$. Calculer son espérance. iii)

Calculer $\mathbb{E}(\exp(tN_Q))$, $t \in \mathbb{R}$, dans la limite thermodynamique. iv) Utiliser ce calcul pour estimer la probabilité que l'on observe dans Q une grande fluctuation du v) nombre de particules, i.e. une fluctuation proportionnelle à $|Q|$,

$$P(\{ |N_Q - \mathbb{E}(N_Q)| \geq a|Q| \}) \quad , \quad a > 0\,.$$

Indication: décomposer en $P(\{ N_Q \geq (\lambda + a)|Q| \})$ et en $P(\{ N_Q \leq (\lambda - a)|Q| \})$.

Exercice 3 : Soit X une variable aléatoire réelle ayant une densité de probabilité f_X, supposée continue. Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction strictement monotone, de classe C^1. On pose $\boxed{Y := g \circ X.}$ Déterminer la densité de probabilité f_Y de Y.

chgt. de var. 1D.

Ex. 2: Idée: proba. d'une grande fluctuation : $P(\{N_Q \geq \lambda \cdot |Q| + a|Q|\}) = P(\{e^{tN_Q} \geq e^{t(\lambda|Q| + a|Q|)}\})$, $\forall t > 0$
puis: markov, puis: optimiser l'inégalité sur t.

CORRIGÉ DE LA SÉRIE 7

Exercice 1 : Notons F_{XY}, F_X et F_Y les fonctions de répartition de (X, Y), X et Y. Les propriétés de F_{XY} sont: $0 \leq F_{XY}(t, s) \leq 1$; F_{XY} est monotone non-décroissante en chaque variable; elle est continue à droite en (t, s),

$$\lim_{t_n \downarrow t} \lim_{s_m \downarrow s} F_{XY}(t_n, s_m) = F_{XY}(t, s)\,;$$

$$\lim_{s \uparrow \infty} F_{XY}(t, s) = F_X(t) \quad , \quad \lim_{t \uparrow \infty} F_{XY}(t, s) = F_Y(s)\,;$$

enfin

$$\lim_{t \downarrow -\infty} F_{XY}(t, s) = \lim_{s \downarrow -\infty} F_{XY}(t, s) = 0\,.$$

La vérification de ces propriétés est semblable à celle des propriétés analogues de F_X. Nous vérifions

$$\lim_{s \uparrow \infty} F_{XY}(t, s) = F_X(t)\,.$$

La variable Y étant finie, $Y < \infty$, nous obtenons par continuité séquentielle de P,

$$P(\{X \leq t\}) = P(\{X \leq t, Y < \infty\}) = \lim_{s \uparrow \infty} P(\{X \leq t, Y \leq s\}) = \lim_{s \uparrow \infty} F_{XY}(t, s)\,.$$

Exercice 2 : La variable X définit une partition finie de Ω, $\{C_1, \ldots, C_n\}$; X est constante sur C_i, $X = x_i$. Y définit une partition finie de Ω, $\{D_1, \ldots, D_m\}$; Y est constante sur D_j, $Y = y_j$. Posons $p(x_i, y_j) := P(\{X = x_i, Y = y_j\})$. Remarquons que

$$\sum_{j=1}^{m} p(x_i, y_j) = \sum_{j=1}^{m} P(\{X = x_i, Y = y_j\}) = P(\{X = x_i\}) \equiv p(x_i)\,.$$

La variable aléatoire XY est associée à la partition $\{C_i \cap D_j : 1 \leq i \leq n, 1 \leq j \leq m\}$; sur $C_i \cap D_j$ XY est constante et prend la valeur $x_i y_j$. Par définition de $\mathbb{E}(XY)$ nous obtenons

$$|\mathbb{E}(XY)| = |\sum_{i,j} x_i y_j p(x_i, y_j)| \leq \sum_{i,j} (|x_i|^\alpha p(x_i, y_j))^{1/\alpha} (|y_j|^\beta p(x_i, y_j))^{1/\beta}\,,$$

ce qui donne par les inégalités classiques de Hölder, resp. de Minkowski,

$$|\mathbb{E}(XY)| \leq \Big(\sum_{i,j} |x_i|^\alpha p(x_i, y_j)\Big)^{1/\alpha} \Big(\sum_{i,j} |y_i|^\beta p(x_i, y_j)\Big)^{1/\beta} = \mathbb{E}(|X|^\alpha)^{1/\alpha} \mathbb{E}(|Y|^\beta)^{1/\beta}\,.$$

$$\Big(\sum_{i,j} p(x_i, y_j)|x_i + y_j|^\alpha\Big)^{1/\alpha} = \Big(\sum_{i,j} |p(x_i, y_j)^{1/\alpha} x_i + p(x_i, y_j)^{1/\alpha} y_j|^\alpha\Big)^{1/\alpha}$$
$$\leq \Big(\sum_{i,j} p(x_i, y_j)|x_i|^\alpha\Big)^{1/\alpha} + \Big(\sum_{i,j} p(x_i, y_j)|y_j|^\alpha\Big)^{1/\alpha}\,.$$

Exercice 3 : $\sigma^2(X) = 0$ si et seulement si $\boxed{P(\{\,\omega \in \Omega : X(\omega) \neq \mathbb{E}(X)\,\}) = 0,}$ ce qui s'exprime en disant que X est presque sûrement constante. Si $0 \leq X(\omega) \leq c$, posons $X \equiv c \cdot Y$ et $\alpha := \mathbb{E}(Y)$; nous avons $0 \leq Y \leq 1$ et $Y^2 \leq Y$ et donc $\mathbb{E}(Y^2) \leq \mathbb{E}(Y)$.

$$\sigma^2(X) \leq c^2 \sigma^2(Y) = c^2(E(Y^2) - \alpha^2) \leq c^2(\alpha - \alpha^2) \leq \frac{c^2}{4}\,.$$

$Y := |X - \mathbb{E}(X)|^2 \geq 0$, $\mathbb{E}(Y) = \sigma^2(X)$ et $\{\,Y \geq t^2\,\} = \{\,|X - \mathbb{E}(X)| \geq t\,\}$. Nous avons

$$P(\{\,|X - \mathbb{E}(X)| \geq t\,\}) = P(\{\,Y \geq t^2\,\}) \leq \frac{\mathbb{E}(Y)}{t^2} = \frac{\sigma^2(X)}{t^2}\,.$$

Exercice 4 : C'est faux: considérer X t.q. $P(\{X = a\}) > 0$ et choisir $a_n \downarrow a$.

PROBABILITÉ ET STATISTIQUE

SÉRIE 9

espérance de Poisson

par déf. c'est l'espérance : $\overbrace{\mathbb{E}}^{\text{de Bernoulli}} = n \cdot p$

Exercice 1 : On considère la distribution binomiale $Bi(n,q)$ avec $q = q(n)$ dépendant de n. Posons $\lambda(n) := q(n) \cdot n$; soient $\alpha(n) \geq 0$ et $\beta(n) \geq 0$, t.q. $\lim_{n\to\infty} \alpha(n) = 0$ et $\lim_{n\to\infty} \beta(n) = 0$ lorsque $n \to \infty$. Montrer que si $\underline{q(n)\sqrt{n} \leq \alpha(n)}$ et $\underline{k \leq \beta(n)\sqrt{n}}$, alors pour n grand

⚠ ≠ =

$$Bi(n, q(n))(k) \simeq \frac{\lambda(n)^k}{k!} \mathrm{e}^{-\lambda(n)} .$$

Donner une estimation de l'erreur. Indication: $\forall t$, $|t| \leq a$ et $\forall a$, $a \geq 1$,

$$0 \leq \mathrm{e}^{-t} - \left(1 - \frac{t}{a}\right)^a \leq \frac{t^2}{a} \mathrm{e}^{-t} .$$

Exercice 2 : Une variable aléatoire X est dite exponentielle de paramètre $\lambda > 0$, si elle a une densité de probabilité

$$f(x) := \begin{cases} c\mathrm{e}^{-\lambda x} & \text{si } x \geq 0 , \\ 0 & \text{sinon } x < 0 . \end{cases}$$

Donner la valeur de c; calculer l'espérance et la variance de X. $t \in \mathbb{R}$ est une **médiane** pour X si et seulement si

$$P(\{X \leq t\}) \geq 1/2 \quad \text{et} \quad P(\{X \geq t\}) \geq 1/2 .$$

Déterminer la médiane de X.

Exercice 3 : Soit X une variable aléatoire gaussienne réduite. Calculer la densité des variables $|X|$ et X^2.

Exercice 4 : On considère un schéma de Bernoulli défini par $P(\{ X_i = 1 \}) = P(\{ X_i = 0 \}) = 1/2$, $\forall i$. Déterminer la probabilité que $\sum_{i=1}^{2n} X_i = n + k$ lorsque k est fixé et n tend vers l'infini. Indication: utiliser la formule de Stirling.

CORRIGÉ DE LA SÉRIE 8

Exercice 1 : On remarque que $\mathrm{Cov}(X_1, X_2) = \mathbb{E}((X_1 - \mathbb{E}(X_1))(\mathbb{X}_2 - \mathbb{E}(X_2)))$. Soit $C_{i,j}$ la matrice de covariance. Soit $u = (u_1, \cdots, u_n) \in \mathbb{R}^n$;

$$\sum_{i=1}^{n}\sum_{j=1}^{n} u_i u_j C_{i,j} = \sum_{i=1}^{n}\sum_{j=1}^{n} u_i u_j \mathbb{E}((X_i - \mathbb{E}(X_i))(\mathbb{X}_j - \mathbb{E}(X_j)) =$$

$$\mathbb{E}(\sum_{i=1}^{n} u_i(X_i - \mathbb{E}(X_i)) \sum_{j=1}^{n} u_j(X_j - \mathbb{E}(X_j))) \geq 0\,.$$

Exercice 2 : Calculons la probabilité que les particules $\omega_{i_1} \in Q, \ldots, \omega_{i_k} \in Q$, $\omega_j \notin Q$ si $j \neq i_1, \ldots, i_k$. L'événement correspondant est le sous-ensemble $A \subset \Omega$,

$$A = \{\, \omega \in \Omega : \omega_j \in Q,\ j = i_1, \ldots, i_k\,; \omega_j \notin Q,\ j \neq i_1, \ldots, i_k \,\}\,.$$

Nous avons

$$P(A) = \int_A \varphi(\omega)\, d\omega = \left(\frac{|Q|}{R^3}\right)^k \left(1 - \frac{|Q|}{R^3}\right)^{n-k}\,.$$

par conséquent la distribution de N_Q est binomiale de paramètres n, $q = \frac{|Q|}{R^3}$,

$$P(\{\, N_Q = k \,\}) = \binom{n}{k} \left(\frac{|Q|}{R^3}\right)^k \left(1 - \frac{|Q|}{R^3}\right)^{n-k}\,.$$

$\mathbb{E}(N_Q) = n \cdot q = \rho \cdot |Q|$, ρ la densité (physique) de particule. Dans la limite thermodynamique $n \cdot q \to \lambda |Q|$. La distribution de N_Q est une distribution de Poisson d'intensité $\lambda|Q|$. $\mathbb{E}(N_Q) = \lambda \cdot |Q|$, λ densité du gas. Posons $\overline{\lambda} := \lambda|Q|$ et $\overline{a} := a|Q|$. Un calcul simple donne

$$\mathbb{E}(\exp(tN_Q)) = \exp(-\overline{\lambda}(1 - \mathrm{e}^t))\,.$$

Par l'inégalité de Markov nous obtenons

$$P(\{\, N_Q \geq \overline{\lambda} + \overline{a} \,\}) = P(\{\, \mathrm{e}^{tN_Q} \geq \mathrm{e}^{t(\overline{\lambda}+\overline{a})} \,\}) \leq \exp(-\overline{\lambda}(1 - \mathrm{e}^t) - (\overline{\lambda} + \overline{a})t)\,.$$

On optimalise par rapport à t, i.e. on choisit $t = \ln \frac{\overline{\lambda}+\overline{a}}{\overline{\lambda}}$.

$$P(\{\, N_Q \geq \overline{\lambda} + \overline{a} \,\}) \leq \exp(\overline{a} - (\overline{\lambda} + \overline{a}) \ln(1 + \frac{\overline{a}}{\overline{\lambda}}))\,.$$

De même

$$P(\{\, N_Q \leq \overline{\lambda} - \overline{a} \,\}) = P(\{\, \mathrm{e}^{-tN_Q} \geq \mathrm{e}^{-t(\overline{\lambda}-\overline{a})} \,\}) \leq \exp(-\overline{a} - (\overline{\lambda} - \overline{a}) \ln(1 - \frac{\overline{a}}{\overline{\lambda}}))\,.$$

Si $\overline{a}/\overline{\lambda} \ll 1$, par ex. $a = u/\sqrt{|Q|}$, $\ln(1 + x) \simeq x$ et $P(\{\, |N_Q - \mathbb{E}(N_Q)| \geq a|Q| \,\}) \leq 2\exp(-\frac{a^2|Q|}{\lambda})$.

Exercice 3 : Si X a une densité de probabilité continue f_X,

$$f_X(t) = \frac{d}{dt} F_X(t) = \frac{d}{dt} \int_{-\infty}^{t} f_X(s)\, ds\,.$$

Supposons que g est strictement décroissante.

$$F_Y(t) = P(\{\, Y \leq t \,\}) = P(\{\, g \circ X \leq t \,\}) = P(\{\, X \geq g^{-1}(t) \,\}) = 1 - F_X(g^{-1}(t))\,.$$

En dérivant par rapport à t nous obtenons

$$f_Y(t) = f_X(g^{-1}(t)) \left| \frac{d}{dt} g^{-1}(t) \right|\,.$$

EPFL
Département de mathématiques

Physique, 2ème année

PROBABILITÉ ET STATISTIQUE

SÉRIE 10

Exercice 1 : Une variable aléatoire réelle X a une distribution gamma de paramètres (t, λ), $t > 0$ et $\lambda > 0$, si sa densité de probabilité est

$$\varphi(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{t-1}}{\Gamma(t)} & x \geq 0\,, \\ 0 & x < 0\,. \end{cases}$$

$\Gamma(t)$ est la fonction gamma, qui est définie par   retrouver la fct.

$$\Gamma(t) := \int_0^\infty e^{-y} y^{t-1}\, dy\,.$$

Cette fonction satisfait l'identité $\Gamma(t) = (t-1)\Gamma(t-1)$; si $t = n$, un entier positif, $\Gamma(n) = (n-1)!$.

a) Calculer $\mathbb{E}(X)$ et $\sigma(X)$.

b) Si X et Y sont deux variables aléatoires indépendantes de distributions gamma, de paramètres (s, λ) resp. (t, λ), alors $X + Y$ a une distribution gamma de paramètres $(s + t, \lambda)$.

c) Considérons une source radioactive; le nombre d'émissions de cette source pendant l'intervalle de temps $[0, x]$ est décrit par une variable aléatoire $N(x)$, dont la distribution est de Poisson $\pi_{\lambda x}$. Désignons par T_n la variable aléatoire indiquant le temps de la n^{ieme} émission. Montrer

$$P(\{\, T_n \leq x \,\}) = P(\{\, N(x) \geq n \,\})\,.$$

Exemple de loi Gamma:
V.A. qui désigne le temps de la n-ème émission d'une source radioactive

d) Déterminer la densité de la variable T_n.

Exercice 2 :
On considère une suite de variables aléatoires indépendantes X_n, $n \geq 1$, sur un espace de probabilité $(\Omega, \mathbb{F}, P)$ telles que $P(X_n = 1) = 1 - P(X_n = 0) = p$, $0 < p < 1$. On pose $S_n = \sum_{j=1}^{n} X_j$.

a) Établir l'estimée

$$P(\left| \frac{S_n}{n} - p \right| \geq \varepsilon) \leq \frac{1}{4\varepsilon^4 n^2}\,. \quad \Longleftrightarrow \quad P\left(\frac{S_n}{n} \geq p \pm \varepsilon \right) \leq \frac{1}{4\varepsilon^4 n^2}$$

b) Utilisant l'estimée ci-dessus démontrer la loi forte des grands nombres de Borel,

$$P(\{\, \lim_{n\to\infty} \frac{S_n}{n} = p \,\}) = 1\,.$$

~~Cantell-Borell~~ Borel-Cantelli

$\xrightarrow{n\to\infty}$) = 0

CORRIGÉ DE LA SÉRIE 9

Exercice 1 : Posons $p(n) := 1 - q(n)$, $\lambda(n) := nq(n)$ et $x(n)(j) := q(n)/p(n) - j/np(n)$.

$$\frac{n!}{k!(n-k)!}q(n)^k(1-q(n))^{n-k} = \frac{\lambda(n)^k}{k!}(1-q(n))^n p(n)^{-k}\prod_{j=0}^{k-1}(1-\frac{j}{n})$$

$$= \frac{\lambda(n)^k}{k!}\mathrm{e}^{-\lambda(n)} \cdot \big((1-\frac{\lambda(n)}{n})^n \mathrm{e}^{\lambda(n)}\big)\prod_{j=0}^{k-1}(1+x(n)(j))$$

Supposons que $|x(n)(j)| \leq a < 1$; nous avons

$$-\frac{1}{1-a}|x| \leq \ln(1+x) \leq |x| \quad \text{si } |x| \leq a\,.$$

Par conséquent,

$$\exp\big(-\frac{1}{(1-a)np(n)}(\frac{k^2}{2}+nq(n)k)\big) \leq \prod_{j=0}^{k-1}(1+x(n)(j)) \leq \exp\big(\frac{1}{np(n)}(\frac{k^2}{2}+nq(n)k)\big)\,,$$

et en utilisant l'indication donnée dans la série,

$$1 - nq(n)^2 \leq \big((1-\frac{\lambda(n)}{n})^n \mathrm{e}^{\lambda(n)}\big) \leq 1\,.$$

Nous obtenons

$$\exp\big(-\frac{\beta(n)[\beta(n)+2\alpha(n)]}{2(1-a)p(n)}\big)(1-\alpha(n)^2)\cdot\frac{\lambda(n)^k}{k!}\mathrm{e}^{-\lambda(n)} \leq Bi(n,q(n))(k)$$

$$\leq \exp\big(\frac{\beta(n)[\beta(n)+2\alpha(n)]}{2p(n)}\big)\cdot\frac{\lambda(n)^k}{k!}\mathrm{e}^{-\lambda(n)}\,.$$

Exercice 2 : $c = \lambda$; $\mathbb{E}(X) = 1/\lambda$; $\sigma^2(X) = 1/\lambda^2$. La médiane est $\ln 2/\lambda$. Noter que la médiane existe toujours, par contre elle n'est pas nécessairement unique. Ne pas confondre médiane et moyenne.

Exercice 3 : Notons $\varphi(t)$ la densité de la gaussienne réduite. La densité $g(t)$ de $|X|$ est nulle $t < 0$; $g(t) = \varphi(\sqrt{t})/\sqrt{t}$ si $t \geq 0$. Ce résultat s'obtient à partir de $F_{X^2}(t) = P(\{-\sqrt{t} \leq X \leq \sqrt{t}\}) = F_X(\sqrt{t}) - F_X(-\sqrt{t})$.

Exercice 4 : La distribution de $S_{2n} := \sum_{i=1}^{2n} X_i$ est la distribution binomiale $Bi(2n, 1/2)$. Nous avons donc

$$P(\{S_{2n} = n+k\}) = \frac{(2n)!}{(n+k)!(n-k)!}\frac{1}{2^{2n}}\,.$$

En utilisant $n! \simeq \sqrt{2\pi n}\,\mathrm{e}^{-n}n^n$ nous obtenons

$$P(\{S_{2n} = n+k\}) \simeq \left(\frac{n}{n+k}\right)^{n+k}\left(\frac{n}{n-k}\right)^{n-k}\sqrt{\frac{n}{\pi(n^2-k^2)}}\,.$$

Le dernier facteur tend vers $1/\sqrt{\pi n}$ lorsque $n \to \infty$. D'autre part

$$\lim_{n\to\infty}\left(\frac{n}{n+k}\right)^{n+k} = \lim_{n\to\infty}\left(1-\frac{k}{n+k}\right)^{n+k} = \mathrm{e}^{-k}\,,$$

$$\lim_{n\to\infty}\left(\frac{n}{n-k}\right)^{n-k} = \lim_{n\to\infty}\left(1+\frac{k}{n-k}\right)^{n-k} = \mathrm{e}^{k}\,.$$

La valeur asymptotique de $P(\{S_{2n} = n+k\})$ est $1/\sqrt{\pi n}$, indépendante de k.

EPFL
Département de mathématiques

PROBABILITÉ ET STATISTIQUE

SÉRIE 11

Exercice 1 : On considère une suite de variables indépendantes de Bernoulli, X_n, $n \geq 1$, telles que $P(X_n = {}^{1}\cancel{0}) = 1 - P(X_n = 0) = p$, $0 < p < 1$. On définit deux variables aléatoires S_n, $n \in \mathbb{N}$, et T_r, $r \in \mathbb{N}$:

$$S_n(\omega) := \sum_{j=1}^{n} X_j \quad , \quad T_r(\omega) := \min\{\, m : S_m \geq r \,\} .$$

Vérifier l'identité

$$P(T_r > n) = P(S_n < r) .$$

Déterminer la distribution de ces deux variables aléatoires; montrer en particulier que $P(\{\, T_r < \infty \,\}) = 1$.

Exercice 2 : On considère deux urnes contenant chacune N boules. On tire une boule d'une des urnes à chaque unité de temps en choisissant l'urne au hasard, sans enregistrer l'urne qui a été choisie. Lorsque l'on découvre pour la première fois que l'une des urnes est vide, quelle est la probabilité que l'autre urne contienne k boules?

⚠ pas tjs $(\mathbb{R}, \mathcal{B}, \mu_x)$, on peut avoir $(J, \mathcal{B}_J, \mu_x)$

Exercice 3 : Développement décimal.
Soit l'espace de probabilité $(\Omega, \mathbb{F}, P)$ avec $\Omega = [0, 1)$, $\mathbb{F}$ les boréliens de Ω et P la mesure de Lebesgue. On pose

↳ à retenir de cet exercice : utilise la loi des grands nbres (loi forte) pour démontrer ce résultat facil, mais surtout LA FORMALISATION : on travaille dans $(J, \mathcal{B}_J, \mu_x)$ → avec la mesure de Lebesgue. On peut avoir des boréliens d'un intervale p. ex.

$$Y(\omega) := k \quad \text{si} \quad \omega \in \left[\frac{k}{10}, \frac{k+1}{10}\right) \quad k = 0, \ldots, 9 .$$

Soit θ la transformation de Ω sur Ω, $\omega \mapsto 10 \cdot \omega \mod 1$. On définit pour tout entier $n \geq 0$ $X_n := Y \circ \theta^n$. $X_n(\omega)$ donne la $(n+1)^{\text{ieme}}$ décimale de ω. Montrer que pour tout $k \in \mathbb{N}$ les variables aléatoires $X_0, \ldots, X_k$ sont indépendantes et identiquement distribuées. Donner la distribution des X_k. On appelle motif m_r de longueur r une suite $a_1 a_2 \cdots a_r$, $a_i \in \{\, 0, 1, \ldots, 9 \,\}$, $i = 1, \ldots, r$. Soit m_r un motif fixé, par exemple $r = 10$ et $m_r = 0123456789$; $N_n(\omega; m_r)$ est le nombre de fois que r décimales consécutives de ω coincident avec le motif m_r parmi les n premières décimales de ω. Montrer que pour presque tout $\omega \in \Omega$

$$\lim_{n \to \infty} \frac{N_n(\omega; m_r)}{n} = \frac{1}{10^r} .$$

$0 \quad \tau \quad 2\tau \quad \longrightarrow \quad M\tau = t$

τ macroscopique

$P_\tau(1) = \lambda \cdot \tau$

$P_\tau(0) = 1 - \lambda\tau$

Soit X_j la v.A. du nb. de part. émises durant l'intervalle $[(j-1)\tau, j\tau]$

N_t "nb. part. émises durant $[0,t]$" $= \sum_{j=1}^{M} X_j$

$P(N_t = n) = \binom{M}{N} \left(\frac{\lambda t}{M}\right)^n \left(1 - \frac{\lambda t}{M}\right)^{M-n} \xrightarrow{M\to\infty}$ c.f. ex. 1 série 9

$\lim_{M\to\infty} P(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$

CORRIGÉ DE LA SÉRIE 10

Exercice 1 : Un calcul immédiat donne

$$\mathbb{E}(X) = \frac{\Gamma(t+1)}{\lambda\Gamma(t)} = \frac{t}{\lambda} \quad , \quad \sigma^2(X) = \frac{t}{\lambda^2}.$$

La densité de la variable $X + Y$ est donnée par la formule

$$\varphi_{X+Y}(u) = \int_{-\infty}^{\infty} \varphi_X(u-y)\varphi_Y(y)\,dy.$$

Concrètement, sans expliciter les constantes nous avons

$$\begin{aligned}
\varphi_{X+Y}(u) &= \frac{1}{\Gamma(s)\Gamma(t)} \int_0^u \lambda e^{-\lambda(u-y)}[\lambda(u-y)]^{s-1}\lambda e^{-\lambda y}[\lambda y]^{t-1}\,dy \\
&= K e^{-\lambda u} \int_0^u [u-y]^{s-1} y^{t-1}\,dy \\
&= K e^{-\lambda u} u^{s+t-1} \int_0^1 (1-x)^{s-1} x^{t-1}\,dx \\
&= C e^{-\lambda u} u^{s+t-1}.
\end{aligned}$$

C est indépendante de u; par conséquent $X + Y$ a une distribution gamma de paramètres $(s+t, \lambda)$. c)+ d): $T_n \leq x$ si et seulement si le nombre d'émissions jusqu'au temps x est au moins n, i.e. $\{T_n \leq x\} = \{N(t) \geq n\}$. Nous avons donc

$$P(\{T_n \leq x\}) = \sum_{j=n}^{\infty} \frac{e^{-\lambda x}(\lambda x)^j}{j!}.$$

Nous obtenons la densité de T_n en dérivant par rapport à x. Un calcul simple montre que T_n a une distribution gamma de paramètres (n, λ).

Exercice 2 : On utilise l'inégalité de Markov. Posons $Y := |\frac{S_n}{n} - p|$.

$$P(\{Y \geq \varepsilon\}) = P(\{Y^4 \geq \varepsilon^4\}) \leq \frac{\mathbb{E}(Y^4)}{\varepsilon^4}.$$

$$\mathbb{E}(Y^4) = \frac{1}{n^4} \sum_{i,j,k,l=1}^{n} \mathbb{E}[(X_i - p)(X_j - p)(X_k - p)(X_l - p)].$$

Les seuls termes non nuls sont du type $\mathbb{E}[(X_i - p)^2(X_j - p)^2] = p^2(1-p)^2$ et $\mathbb{E}[(X_i - p)^4] = p(1-p)(p^3 + (1-p)^3)$. Il y a $3n(n-1)$ termes du premier type et n du second type. Donc

$$\mathbb{E}(Y^4) = \frac{p(1-p)}{n^4}[n(p^3 + (1-p)^3) + 3p(1-p)(n^2 - n)] \leq \frac{1}{4n^2}.$$

Le lemme de Borel-Cantelli donne pour chaque $\varepsilon > 0$

$$P(\limsup_n \{ |\frac{S_n}{n} - p| \geq \varepsilon \}) = 0.$$

Par conséquent (k entier)

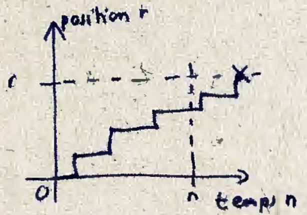
$$N := \bigcup_{k\geq 1} \limsup_n \{ |\frac{S_n}{n} - p| \geq \frac{1}{k} \}$$

est un ensemble négligeable. L'ensemble complémentaire est de probabilité un; si $\omega \notin N$, alors $\forall k$ il existe $n_k(\omega)$ tel que si $n \geq n_k(\omega)$ on a $|\frac{S_n(\omega)}{n} - p| \leq \frac{1}{k}$.

Série 11, Exercice 1 : en bref

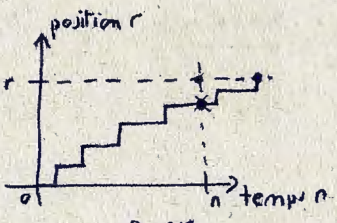
Soit : $S_n(\omega) = \sum_{j=1}^{n} X_j$; $T_r(\omega) = \min\{m : S_m \geq r\}$

A voir : $P(T_r > n) = P(S_n < r)$



$T_r > n$

équivalent $\longleftrightarrow$



$S_n < r$

distribution : $P(S_n(\omega) = r) = \binom{n}{r} p^{r}(1-p)^{n-r}$ (Bernoulli)

$P(T_r(\omega) = n) \iff \begin{cases} \sum_{j=1}^{n-1} X_j = r-1 & \text{"="}\ A \\ X_n = 1 & \text{"="}\ B \end{cases}$; $A \cap B = \emptyset$, indépendants

$\Rightarrow P(T_r(\omega) = n) = P(A \cap B) = P(A)\cdot P(B) = \binom{n-1}{r-1} p^{r-1}(1-p)^{n-r}\cdot p = \binom{n-1}{r-1} p^{r}(1-p)^{n-r}$

(Pascal : proba de : on a r succès avec n épreuves)

$P(T_r(\omega) < \infty) = 1$: $P(T_r(\omega) < \infty) = P(\lim_{k\to\infty} T_r(\omega) < k) = 1 - P(\lim_{k\to\infty} T_r(\omega) \geq k)$

$\geq 1 - \lim_{k\to\infty} \frac{\mathbb{E}(T_r)}{k} = 1 - \lim_{k\to\infty} \frac{1}{k}\cdot\frac{r}{p} = 1$ #

Série 11, exercice 2 : "Axiomatisation"

Soit l'espace $(\Omega, \mathcal{F}, p)$

$\Omega =$ "ensemble de toutes les épreuves", i.e. tous les cas possibles de tirages de boules dans les 2 urnes jusque à ce que 1 urne est vide

$\Omega = \{ \omega_1, \omega_2, \dots \omega_k \}$

$\omega_i = (A, B, A, A, \dots, A)$ — A, B : tirage dans l'urne A, B

$\mathcal{F}$: algèbre de Boole des événements ($\mathrm{card}(\Omega) < \infty \rightarrow$ pas besoin de la σ-additivité)

p : mesure de proba.

Soit les événements $A, B \in \mathcal{F}$, $E \in \mathcal{F}$

$A =$ "l'urne A est vide et il reste k boules ds l'urne B"

$B =$ "l'urne B est vide et il reste k boules dans l'urne A"

$E = A \cup B \in \mathcal{F}$ (union disjointe)

$\mathrm{Card}(\Omega) < \infty \rightarrow$ esp. prob. discret :

$$P(A) = \frac{\sum_{\omega \in A} p(\omega)}{\sum_{\omega \in \Omega} p(\omega)} \quad ; \quad P(E) = 2 \cdot P(A) = P(A) + P(B)$$

1) Pas formalisé

Chaque tirage indépendant du précédent, $p(\text{tiré ds } A) = p(\text{tiré ds } B) = 1/2$.

Calcul de $P(A)$: pour vider l'urne A, on doit tirer N boules de A + 1 boule virtuelle pour se rendre compte que l'urne est effectivement vide. Il doit rester k boules dans l'urne B, donc on en a tiré $(N-k)$. Par l'indépendance : pour 1 tirage donné, la proba :

$$\left(\frac{1}{2}\right)^{N+1} \left(\frac{1}{2}\right)^{N-k} = \frac{1}{2^{2N+1-k}}$$

$\exists$ plusieurs façons de tirer N boules de l'urne A, parmi un total de $(2N-k)$ boules tirées : $\binom{2N-k}{N}$

La proba est donc :

$$P(A) = \binom{2N-k}{N} \frac{1}{2^{2N+1-k}}$$

$$\boxed{P(E) = \binom{2N-k}{N} \frac{1}{2^{2N-k}}}$$

2) Avec l'exercice 1

Soit X_j la V.A. qui décrit le j-ème tirage : $\begin{cases} X_j = 1 & \text{si boule tirée dans A} \\ X_j = 0 & \text{si boule tirée dans B} \end{cases}$

$P(X_j = 1) = P(X_j = 0) = 1/2$ (équiproba., indépendance VS)

Soit la V.A. qui compte le nb. de boules tirées dans A :

$$S_n(\omega) = \sum_{j=1}^{n} X_j \quad ; \quad T_r(\omega) = \min\{ m \text{ t.q. } S_m(\omega) \geq r \}$$

Pour obtenir l'événement A, on veut que :

$$S_{n-1}(\omega) = N \quad ; \quad X_n = 1 \quad \longrightarrow \quad S_n = N+1$$

En tout on doit avoir tiré $n = (2N-k+1)$ boules, autrement dit :

$$T_{N+1} = 2N - k + 1$$

i.e. le plus petit indice m pour lequel on a $S_m = N+1$ (vidé l'urne A) est $2N-k+1$, d'où :

$$P(E) = 2 \cdot P(A) = 2 \cdot P(T_{N+1} = 2N-k+1) = \binom{2N-k}{N} \cdot \left(\frac{1}{2}\right)^{N+1} \cdot \left(\frac{1}{2}\right)^{2N-k-N} \cdot 2$$

$$\boxed{P(E) = \binom{2N-k}{N} \left(\frac{1}{2}\right)^{2N-k}}$$

PROBABILITÉ ET STATISTIQUE

SÉRIE 12

à retenir sur l'ex: – déf. de l'entropie relative
– ce serait bien : tout ce qui est mis en évidence

Exercice 1 : Soit $(\Omega, \mathcal{F}, P)$. $\underline{A} = \{A_1, \ldots, A_n\}$ est une partition finie de Ω, $\mathcal{A}$ l'algèbre engendrée par $\underline{A}$; on note $p_i := P(A_i)$. Soit Q une autre mesure de probabilité sur $(\Omega, \mathcal{F})$. On note $q_i := Q(A_i)$. On compare les deux mesures de probabilités restreintes à $\mathcal{A}$. (Une autre façon équivalente de considérer l'exercice est la suivante: soit $E = \{a_1, \ldots, a_n\}$ et $X : \Omega \to E$ une variable aléatoire $\mathcal{A}$-mesurable à valeur dans E. $p(a_i) = p_i$, resp. $q(a_i) = q_i$ donne la distribution de X par rapport à P resp. Q.) On définit l'**entropie relative** des mesures de probabilité **restreintes à l'algèbre** $\mathcal{A}$ par la formule

$$D_{\mathcal{A}}(P\|Q) := \sum_{A \in \underline{A}} P(A) \log \frac{P(A)}{Q(A)} = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}.$$

Dans cette formule on utilise les conventions $0 \log \frac{0}{q} = 0$ et $p \log \frac{p}{0} = \infty$.

a) Montrer que $D_{\mathcal{A}}(P\|Q) \geq 0$; $D_{\mathcal{A}}(P\|Q) = 0$ ssi $P = Q$.

b) Soit $\mathcal{B}$ une algèbre finie telle que $\mathcal{A} \subset \mathcal{B}$. Montrer $D_{\mathcal{A}}(P\|Q) \leq D_{\mathcal{B}}(P\|Q)$.

"assez intuitif" : si $\mathcal{B}$ contient $\mathcal{A}$, alors $\mathcal{B}$ a plus de parties que $\mathcal{A}$, alors son désordre est plus grand, donc son entropie relative sera plus grande.

Indication: (vérifier que $t \log t$ est convexe.)

c) Soit $T : \Omega \to \mathbb{R}$ une variable aléatoire, qui est $\mathcal{A}$-mesurable. On définit $\mathcal{P}_Q(T) := \log \mathbb{E}_Q(\exp(T))$. Montrer que pour toute mesure de probabilité S on a l'inégalité

$$\mathbb{E}_S(T) - D_{\mathcal{A}}(S\|Q) \leq \mathcal{P}_Q(T),$$

et que l'on a égalité si et seulement si $S = Q_T$ où

$$Q_T(A_i) := \frac{\exp(T(\omega))}{\exp \mathcal{P}_Q(T)} Q(A_i) \quad \forall \omega \in A_i, \ \forall i = 1, \ldots n.$$

Indication: exprimer $\mathbb{E}_S(T) - D_{\mathcal{A}}(S\|Q)$ à l'aide de $D_{\mathcal{A}}(Q_T\|Q)$ et $\mathcal{P}_Q(T)$.

d) Montrer que $\max_{A \in \mathcal{A}} |Q(A) - P(A)| = 1/2 \sum_{i=1}^{n} |Q(A_i) - P(A_i)|$. par clair

e) Montrer que $2(\max_{A \in \mathcal{A}} |Q(A) - P(A)|)^2 \leq \cancel{2} D_{\mathcal{A}}(P\|Q)$. Indication: vérifier ce résultat lorsque $\underline{A}$ ne contient que deux éléments, puis utiliser b).

Exercice 2 : On considère une suite de variables aléatoires définies sur $(\Omega; \mathcal{F}, P)$, indépendantes et identiquement distribuées. Chaque variable X_n, $n \geq 1$, prend ses valeurs dans $E = \{a_1, \ldots, a_r\}$; la distribution de X_n est aussi noté P, $p_j := P(X_n = a_j)$, $\forall j$. Soit $N_n(\omega; a_j)$ le nombre d'indices k, $1 \leq k \leq n$, tels que $X_k(\omega) = a_j$. La distribution empirique des variables $X_1, \ldots, X_n$ est la mesure de probabilité définie sur E par la formule

$$T_n(\omega)(a_j) := \frac{N_n(\omega; \overset{q_j}{\cancel{A_j}})}{n}.$$

Estimer le nombre de valeurs différentes pour T_n; soit Q une ces valeurs; montrer

$$C_n^- \exp(-nD(Q\|P)) \leq P(T_n = Q) \leq C_n^+ \exp(-nD(Q\|P)), \quad C_n^\pm = O(n^{\pm(r-1)/2}).$$

$\|P - Q\| := \sum_{i=1}^{r} |q_j - p_j|$ est une distance sur l'ensemble des mesures de probabilité définies sur E. Montrer que $\lim_{n \to \infty} \|T_n - P\| = 0$ avec probabilité 1.

CORRIGÉ DE LA SÉRIE 11

Exercice 1 : Si $\omega \in \{T_r > n\}$ alors $S_n(\omega) < r$; par conséquent $\{T_r > n\} \subset \{S_n < r\}$. Si $\omega \in \{S_n < r\}$ alors $T_r(\omega) > n$, ce qui implique $\{T_r > n\} \supset \{S_n < r\}$. La distribution de S_n est une distribution binomiale Bi(n,p). Calculons $P(T_r = k)$; on a

$$\{ T_r = k \} = \{ S_{k-1} = r-1 \} \cap \{ X_k = 1 \} .$$

Ces deux derniers événements sont indépendants; par conséquent $(k \geq r)$

$$P(T_r = k) = P(S_{k-1} = r-1) P(X_k = 1) = \binom{k-1}{r-1} p^r (1-p)^{k-r} .$$

$P(\{ T_r < \infty \}) = 1$ si et seulement si $P(\{ T_r < \infty \}^c) = 0$. Ceci suit de

$$\{ T_r < \infty \}^c = \bigcap_n \{ S_n \leq r-1 \}$$

et $\lim_{n \to \infty} P(S_n \leq r-1) = 0$.

Exercice 2 : Les deux urnes sont numérotées I et II. Supposons que c'est l'urne I qui est vide en premier et que lorsque l'on découvre quelle est vide, il y a k boules dans l'urne II. Notons cet événement par $E_I(k)$. Cela signifie que l'événement $E_I(k)$ a lieu lors du tirage $n = N + 1 + (N - k)$. En utilisant la distribution de T_r de l'exercice 1 avec $p = 1/2$, $r = N+1$, $P(E_I(k)) = \binom{n-1}{N} 2^{-n}$. La probabilité de l'événement cherché est $P(E_I(k)) + P(E_{II}(k)) = \binom{2N-k}{N} 2^{-(2N-k)}$.

Exercice 3 : On introduit de manière récursive les intervalles ci-dessous. $I_k := [\frac{k}{10}, \frac{k+1}{10})$, $k = 0, 1, \ldots, 9$. On divise ensuite chaque I_k en 10 intervalles de longueur égale, ce qui donne 10^2 intervalles I_{kj}, $k, j = 0, 1, \ldots, 9$ etc. De manière générale on note ω_n la n^{ieme} décimale de ω, $n \geq 1$.

$$I_{j_1, \cdots, j_r} = \{ \omega \in \Omega : \omega_k = j_k \, , k = 1, \ldots, r \} .$$

Nous avons $\{ X_n(\omega) = k \} = \cup_{j_1, \ldots, j_n} I_{j_1, \cdots, j_n, k}$; comme tous ces intervalles sont disjoints $P(X_n(\omega) = k) = \frac{1}{10}$. L'identité $\cap_{j=0}^n \{ X_j = k_j \} = I_{k_0, \cdots, k_n}$, donne $P(X_0 = k_0; \cdots; X_n = k_n) = 10^{-(n+1)} = P(X_0 = k_0) \cdots P(X_n = k_n)$. Posons $Y_{m_r;k}(\omega) := 1$ si les décimales $\omega_k \cdots \omega_{k+r}$ coïncident avec m_r et $Y_{m_r;k}(\omega) := 0$ sinon. Nous avons

$$N_n(\omega; m_r) = \sum_{j=0}^{r-1} \sum_{\substack{s=0,1,\ldots \\ j+sr \leq n-r}} Y_{m_r; j+sr}(\omega) .$$

Les variables aléatoires des r sommes de droite sont indépendantes; par la loi forte des grands nombres, presque certainement

$$\lim_{n \to \infty} \frac{1}{n} \sum_{\substack{s=0,1,\ldots \\ j+sr \leq n-r}} Y_{m_r; j+sr}(\omega) = \frac{1}{r} \mathbb{E}(Y_{m_r;k}) = \frac{1}{r} \frac{1}{10^r} .$$

!

Remarque: On ne sait pas si ce résultat est valable pour le nombre π; on a trouvé le motif 0123456789 pour la première fois en juillet 1997; le motif commence à la 17,387,594,880 décimale! Par contre le nombre de Champernowne 0.123456789101112131415161718192021.... a la propriété ci-dessus.

Questions de mathématiciens : démontrer que le nb. π a cette propriété ... intuitivement oui, mais on ne l'a pas démontré. clairement

PROBABILITÉ ET STATISTIQUE

SÉRIE 13

Exercice 1 : Soit $\Omega_n := \{-1,1\}^n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_k = \pm 1 \forall k\}$. On introduit les variables aléatoires $\sigma_k(\omega) := \omega_k$, $\forall k$. En mécanique classique ces variables aléatoires sont appelées variables de spin. L'énergie des variables de spin est la variable aléatoire

$$H_n := -\frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j - h\sum_{i=1}^{n}\sigma_i\,.$$

h est un paramètre réel décrivant un champ magnétique externe. La mesure de probabilité est celle de Gibbs, P_n, qui est définie par la formule

$$P_n(\omega) := \frac{\mathrm{e}^{-\beta H_n(\omega)}}{Z_n}\,,$$

où Z_n est la constante de normalisation, $Z_n = \sum_\omega \mathrm{e}^{-\beta H_n(\omega)}$. Le paramètre réel positif β est l'inverse de la température. Le but de l'exercice est d'étudier la distribution de la variable aléatoire $m_n := \frac{1}{n}\sum_{i=1}^{n}\sigma_i$, plus précisément le comportement asymptotique lorsque $n \to \infty$ des expressions $\frac{1}{n}\ln P_n(m_n \in B)$, $B \subset [-1,1]$. Le principe de base pour comprendre ce comportement asymptotique est le *principe du terme dominant*: soit $B = \cup_{k=1}^{r} B_k \subset [-1,1]$; montrer que

?

i) $$\limsup_{n\to\infty}\left(\frac{1}{n}\ln P_n(m_n \in B)\right) = \max_k \limsup_{n\to\infty}\left(\frac{1}{n}\ln P_n(m_n \in B_k)\right).$$

énergie libre

Soit $B := [a,b]$; posons $s(x) := -\frac{1+x}{2}\ln\frac{1+x}{2} - \frac{1-x}{2}\ln\frac{1-x}{2}$, $x \in [-1,1]$ et $f(x) := s(x) + \frac{\beta x^2}{2} + \beta h x$. Étudier soigneusement $P_n(m_n \in B)$ pour n grand et montrer que

ii) $$\lim_{n\to\infty}\frac{1}{n}\ln P_n(m_n \in B) = \max_{x\in[a,b]}(f(x) - \max_{x'\in[-1,1]} f(x'))\,.$$

Soit $M(\beta,h)$ les maximas de la fonction f. Montrer que la distribution de la variable aléatoire m_n est concentrée, lorsque $n \to \infty$, sur l'ensemble $M(\beta,h)$, dans le sens suivant: $\forall\varepsilon > 0$, soit $M_\varepsilon(\beta,h) := \{x : \min_{y\in M(\beta,h)}|x-y| \le \varepsilon\}$; $P_n(m_n \notin M_\varepsilon(\beta,h))$ tend vers zéro exponentiellement vite,

$$\lim_{n\to\infty}\frac{1}{n}\ln P_n(m_n \notin M_\varepsilon(\beta,h)) < 0\,.$$

Supposons que $h = 0$. Montrer que la loi (faible) des grands nombres **n'est pas vérifiée** pour $\frac{1}{n}\sum_{i=1}^{n}\sigma_i$ si $|M(\beta,0)| \ge 2$, i.e. si l'on a une **transition de phase**. Montrer qu'une telle situation a lieu pour $h = 0$ et β convenable.

la loi faible ne marche pas dans le cas du modèle de Curie-Weiss

A retenir de l'exercice : – modèle physique (Ising) de particules avec un spin

– $\beta \sim \frac{1}{T}$: si $\beta \uparrow$, alors il n'y a plus indépendance et la loi faible des grd nb. n'est plus vérifiée (intéraction entre particules)

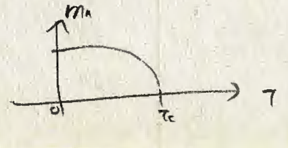
en fait : $\longrightarrow$ pt. critique : température de Curie T_c

$$m_n = \frac{1}{n}\sum_{i=1}^{n}\sigma_i$$: moment résultant

si $T < T_c$, β suffisament grd $\longrightarrow$ observe aimantation m_n résultante non nulle

$\longrightarrow$ modèle sans intéraction plus valable

$\longrightarrow$ loi faible plus valable



CORRIGÉ DE LA SÉRIE 12

Exercice 1 : $t \mapsto -\log t$ est strictement convexe. Par l'inégalité de Jensen

$$-D_{\mathcal{A}}(P\|Q) = \sum_{i=1}^{n} p_i \log \frac{q_i}{p_i} \leq \log \sum_{i=1}^{n} p_i \frac{q_i}{p_i} = 0\,.$$

On $=$ à la place de $\leq$ ssi $q_i = p_i$. b) La condition $\mathcal{A} \subset \mathcal{B}$ signifie que $\forall i$, $A_i = \cup_{i_j} B_{i_j}$, $B_{i_j} \in \underline{B}$. Nous avons avec $f(t) = t \log t$, $t_{i_j} = \frac{P(B_{i_j})}{Q(B_{i_j})}$ et $\alpha_{i_j} = \frac{Q(B_{i_j})}{\sum_j Q(B_{i_j})}$,

$$\begin{aligned} D_{\mathcal{A}}(P\|Q) &= \sum_{i=1}^{n} P(A_i) \log \frac{P(A_i)}{Q(A_i)} = \sum_{i=1} (\sum_j P(B_{i_j})) \log \frac{(\sum_j P(B_{i_j}))}{(\sum_j Q(B_{i_j}))} \\ &= \sum_{i=1} (\sum_j Q(B_{i_j})) f(\sum_j \alpha_{i_j} t_{i_j}) \leq \sum_{i=1} (\sum_j Q(B_{i_j})) \sum_j \alpha_{i_j} f(t_{i_j}) \\ &= D_{\mathcal{B}}(P\|Q)\,. \end{aligned}$$

c) On peut écrire $\mathbb{E}_S(T) - D_{\mathcal{A}}(S\|Q)$ ainsi, en choisissant $\omega_i \in A_i$,

$$\sum_i \Big(T(\omega_i) - \log \frac{S(A_i)}{Q(A_i)}\Big) S(A_i) = -D_{\mathcal{A}}(S\|Q_T) + \mathcal{P}_Q(T)\,.$$

Comme $D_{\mathcal{A}}(S\|Q_T) \geq 0$ et $D_{\mathcal{A}}(S\|Q_T) = 0$ ssi $S = Q_T$, c) est démontré.
d) Soit $I := \{i : P(A_i) > Q(A_i)\}$ et posons $A := \cup_{i\in I} A_i$; calculons

$$\sum_j |P(A_j) - Q(A_j)| = \sum_{i\in I} (P(A_i) - Q(A_i)) + \sum_{j \in I^c} (Q(A_j) - P(A_j)) =$$

$$(P(A) - Q(A)) + (Q(A^c) - P(A^c)) = 2(P(A) - Q(A)) = 2 \max_{A' \in \mathcal{A}} |Q(A') - P(A')|\,.$$

e) Supposons que $\underline{A} = \{A, A^c\}$. Posons $p := P(A)$ et $q := Q(A)$. À démontrer

$$G(q) := p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q} - 2(p-q)^2 \geq 0\,.$$

Nous avons $G(p) = 0$ et la dérivée de $G(q)$ s'écrit $\frac{(p-q)[4(1-q)q-1]}{q(1-q)}$. $G(q) \geq 0$ puisque $(1-q)q \leq 1/4$. Cas général, utiliser $D_{\mathcal{A}}(P\|Q) \leq D_{\mathcal{B}}(P\|Q)$ si $\mathcal{A} \subset \mathcal{B}$.

Exercice 2 : Les valeurs de la variable aléatoire T_n sont des mesures de probabilités sur E. Chaque mesure Q sur E est spécifiée par $q_1 \geq 0, \ldots, q_r \geq 0$, $\sum_j q_j = 1$. Pour n fixé $q_j \in \{\, k/n : k = 0, 1, \ldots, n\}$. Le nombre de valeurs que peut prendre T_n est au plus $(n+1)^r$. Soit Q une de ces valeurs.

$$P(T_n = Q) = P(T_n(a_j) = q_j\,, j = 1, \ldots, r) = \binom{n}{nq_1, \ldots, nq_r} p_1^{nq_1} \cdots p_r^{nq_r}\,.$$

En utilisant la formule de Stirling sous la forme $W^-(n) n^n e^{-n} < n! < W^+(n) n^n e^{-n}$, avec $W^-(n) = \sqrt{2\pi n} \exp(\frac{1}{12n+1})$ et $W^+(n) = \sqrt{2\pi n} \exp(\frac{1}{12n})$,

$$\frac{W^-(n)}{\prod_j W^+(nq_j)} \exp(-nD(Q\|P)) \leq P(T_n = Q) \leq \frac{W^+(n)}{\prod_j W^-(nq_j)} \exp(-nD(Q\|P))\,.$$

De ces deux estimations nous obtenons

$$P(\{\, \|T_n - P\| \geq \varepsilon\}) \leq (n+1)^r C_n^+ \exp(-nD(Q\|P)) \leq (n+1)^r C_n^+ \exp(-n \frac{\varepsilon^2}{8})\,.$$

On utilise ensuite le lemme de Borel Cantelli comme dans l'exercice 2 série 10.

EPFL
Département de mathématiques

Physique, 2ème année

EXAMEN: PROBABILITÉ ET STATISTIQUE

- L'examen est oral. Il dure en règle générale 15 minutes.
- Il y a 15 minutes de préparation, sans document.
- L'examen porte sur **tout** le cours **et** sur **toutes** les séries.

La ou les questions posées seront soit d'ordre pratique, soit d'ordre théorique, ou les deux. Si la question posée est d'ordre pratique, cette question sera directement liée à un des exercices ou à un exemple traité dans le cours.

L'un des buts de l'examen est de revoir la matière du cours globalement. Essayer de dégager les points principaux des différents chapitres et de faire le lien avec les exercices. Il faut savoir de façon précise les définitions du cours ainsi que les théorèmes, en particulier il faut pouvoir illustrer concrètement les notions du cours. (Les exercices peuvent être utiles à cet égard!) Les détails techniques des démonstrations des théorèmes ne sont pas exigés. L'accent est mis sur la compréhension des différentes notions du cours ainsi que des théorèmes.

Exercices cibles de l'examen: S4 : 1,2,3
S5 : 1,2
S11 : {1+2}
+ calcul de : – espérances
– densités de $|X|$, X^2 à partir de celle de X . . . etc

CORRIGÉ DE LA SÉRIE 13

Exercice 1 : Principe du terme dominant.

$$\max_k P_n(m_n \in B_k) \leq P_n(m_n \in \cup_k B_k) \leq r \max_k P_n(m_n \in B_k).$$

ln étant une fonction croissante, il s'ensuit que $\limsup_n \frac{1}{n} \ln P_n(m_n \in \cup_k B_k) = \limsup_n \max_k \frac{1}{n} \ln P_n(m_n \in B_k)$. Pour conclure on utilise le résultat élémentaire: soient r suites de nombres réels $\{a_k(n)\}_n$, $k = 1, \ldots, r$; alors $\limsup_n \max_k a_k(n) = \max_k \limsup_n a_k(n)$. D'une part $\limsup_n \max_k a_k(n) \geq \limsup_n a_k(n)$ $\forall k$, d'autre part $\forall \varepsilon > 0$ $\exists n_\varepsilon$ tel que $\forall n \geq n_\varepsilon$, $\max_k a_k(n) \leq \max_k \limsup_n a_k(n) + \varepsilon$.

m_n prend les valeurs $x_k := -1 + \frac{2k}{n}$, $k = 0, 1 \ldots, n$. La distribution de m_n s'écrit

$$P_n(m_n \in B) = \frac{\sum_{x_k \in B} \sum_{\omega:\, m_n = x_k} e^{\frac{n\beta}{2} m_n(\omega)^2 + n\beta h m_n(\omega)}}{\sum_{x_k \in [-1,1]} \sum_{\omega:\, m_n = x_k} e^{\frac{n\beta}{2} m_n(\omega)^2 + n\beta h m_n(\omega)}}.$$

Par la formule de Stirling $\exists$ c_1 et c_2 t.q.

$$c_1 n^{-1/2} e^{ns(x_k)} \leq \binom{n}{n-k} \leq c_2 n^{1/2} e^{ns(x_k)}.$$

Remarque: Pour k fixé les estimations ne sont pas optimales; cependant elles sont valables **uniformément** en $x_k \in [-1, 1]$, ce qui est important plus bas.
On obtient ainsi immédiatement, en posant $f(x) := \frac{\beta x^2}{2} + \beta h x + s(x)$,

$$c_1 n^{-1/2} e^{nf(x_k)} \leq \sum_{\omega:\, m_n = x_k} e^{\frac{n\beta}{2} m_n(\omega)^2 + n\beta h m_n(\omega)} \leq c_2 n^{1/2} e^{nf(x_k)}.$$

On estime $P_n(m_n \in [a, b])$ en estimant de la même façon le numérateur et le dénominateur. Considérons par exemple le dénominateur.

$$c_1 n^{-1/2} e^{n \max_{x' \in [-1,1]} f(x')} \leq Z_n \leq (n+1) c_2 n^{1/2} e^{n \max_{x' \in [-1,1]} f(x')}.$$

Il s'ensuit que

$$\left| \frac{1}{n} \ln P_n(m_n \in [a, b]) - \Big(\max_{x \in [a,b]} f(x) - \max_{x' \in [-1,1]} f(x') \Big) \right| = O(\frac{\ln n}{n}).$$

Nous avons: $f'(x) = -\frac{1}{2} \ln \frac{1+x}{1-x} + \beta x + \beta h$ et $f''(x) = \beta - \frac{1}{1-x^2}$. $f'(x) = 0$ ssi $\tanh(\beta x + \beta h) = x$. $f'(x) = 0$ possède une, deux ou trois solutions. f est concave si $\beta \leq 1$; si $\beta > 1$ elle est convexe sur $(-\frac{\sqrt{\beta-1}}{\sqrt{\beta}}, \frac{\sqrt{\beta-1}}{\sqrt{\beta}})$, sinon elle est concave. En particulier f est strictement concave aux extrémités de $[-1, 1]$. L'ensemble $M(\beta, h) \subset \{x : \tanh(\beta x + \beta h) = x\}$. x est un maximum de f ssi $f(x) - \max_{x' \in [-1,1]} f(x') = 0$, ce qui donne immédiatement le résultat de concentration,

$$\lim_{n\to\infty} \frac{1}{n} \ln P_n(m_n \notin M_\varepsilon(\beta, h)) = \max_{x \in M_\varepsilon(\beta,h)} f(x) - \max_{x' \in [-1,1]} f(x') < 0.$$

√ Soit $h = 0$; par symétrie $\mathbb{E}(\sigma_k) = 0$ $\forall k$. Si $\beta > 1$, alors $|M(\beta, 0)| = 2$; la distribution de m_n est concentrée sur les deux maximas non nuls de f. Si $\varepsilon > 0$ est assez petit,

$$\lim_{n\to\infty} \frac{1}{n} \ln P_n(m_n \in [-\varepsilon, \varepsilon]) = \max_{x \in [-\varepsilon,\varepsilon]} f(x) - \max_{x' \in [-1,1]} f(x') < 0,$$

$$\boxed{\lim_{n\to\infty} P_n(|\frac{1}{n} \sum_{k=1}^n \sigma_k| \geq \varepsilon) = 1 - \lim_{n\to\infty} P_n(|\frac{1}{n} \sum_{k=1}^n \sigma_k| \leq \varepsilon) = 1.}$$

La loi des grands nombres n'est pas vérifiée dans ce cas.

i.e. $m_n \neq 0$ → la résultante du moment magnétique est > 0.