Cours de Probabilités, L3S5, UE 14A

(Licence de Physique de l'Université Louis Pasteur)

J. FRANCHI

Premier semestre 2005-2006. 56 heures, dont la moitié en T.D..

Contenu

I. Fondements de la théorie des probabilités

1 Probabilité

Une probabilité est d'abord une *fonction* qui à un *événement* associe un nombre réel compris entre 0 et 1. Cela implique de préciser ce qu'est un événement. Or cela n'a de sens que dans le cadre d'un ensemble d'*épreuves aléatoires* ou *tirages*, qu'on note généralement Ω. Il peut s'agir par exemple de lancers de dés ou de pièces de monnaie, de tirages d'urne, de durées de vie (d'atomes ou d'individus), de tirs sur une cible, etc... Ces premiers exemples familiers montrent déjà que l'ensemble Ω peut être fini, dénombrable, ou continu. Ce sera donc a priori un ensemble non vide quelconque.

Lorsque Ω est fini ou dénombrable, toutes ses parties seront des événements. Tandis qu'en général il est nécessaire de se restreindre à un sous-ensemble de parties de Ω : $\mathcal{T} \subset \mathcal{P}(\Omega)$, qu'on nomme *$\sigma$-algèbre* ou *tribu*. On a naturellement besoin de pouvoir considérer la réunion et la conjonction (intersection) de 2 événements, de même que le complémentaire d'un événement ; en outre il faut aussi pouvoir considérer une réunion dénombrable d'événements. Enfin il est naturel de considérer l'événement impossible (vide). D'où la définition suivante.

Définition 1 *Une <u>tribu</u> (ou σ-algèbre) est une partie de $\mathcal{P}(\Omega)$ stable par réunion dénombrable et par passage au complémentaire, et contenant $\emptyset$. Le couple $(\Omega, \mathcal{T})$ est appelé <u>espace probabilisable</u>.*

On vérifie aussitôt les propriétés suivantes :

Proposition 1 *Ω est un événement (certain). La différence $A \setminus B := A \cap B^c$ et la différence symétrique $A \Delta B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$ de deux événements A et B sont des événements. Toute intersection dénombrable d'événements est un événement.*

Nous pouvons maintenant définir rigoureusement ce qu'est une probabilité : *une mesure positive de masse 1.*

Définition 2 *Une <u>probabilité</u> sur l'espace probabilisable $(\Omega, \mathcal{T})$ est une fonction $I\!P$ de $\mathcal{T}$ dans $[0, 1]$ qui vérifie : $I\!P(\Omega) = 1$, et la propriété d'additivité dénombrable :*
$I\!P\Big(\bigcup_{n \in I\!N} A_n\Big) = \sum_{n \in I\!N} I\!P(A_n)$ pour tout suite $\{A_n \,|n \in I\!N\}$ d'événements deux à deux disjoints. Le triplet $(\Omega, \mathcal{T}, I\!P)$ est appelé <u>espace probabilisé</u> ou <u>espace de probabilité</u>. Les événements de probabilité nulle sont dits <u>négligeables</u>. Les événements de probabilité 1 sont dits <u>presque sûrs</u>.

C'est toujours dans le cadre d'un espace de probabilité, plus ou moins bien précisé, que peut avoir lieu un calcul de probabilité. Il est généralement préférable de bien le préciser. En effet c'est la non-précision de l'espace considéré qui est à l'origine de paradoxes ou d'erreurs courantes sur les probabilités.

On vérifie aisément les propriétés suivantes :

Proposition 2 *(i) L'événement impossible est négligeable :* $I\!\!P(\emptyset) = 0$.

(ii) $I\!\!P$ *est croissante :* $A \subset B \Rightarrow I\!\!P(A) \leq I\!\!P(B)$, *pour tous événements* A *et* B.

(iii) $I\!\!P\Big(\bigcup\limits_{n\in I\!\!N} A_n\Big) \leq \sum\limits_{n\in I\!\!N} I\!\!P(A_n)$ *pour tout suite* $\{A_n \,|n \in I\!\!N\}$ *d'événements.*

(iv) Toute intersection dénombrable d'événements presque sûrs est presque sûre.

(v) $I\!\!P(A \setminus B) = I\!\!P(A) - I\!\!P(B)$, *lorsque* A *et* B *sont deux événements tels que* $B \subset A$. *En particulier,* $I\!\!P(A^c) = 1 - I\!\!P(A)$ *pour tout événement* A.

(v) $I\!\!P(A \cup B) = I\!\!P(A) + I\!\!P(B) - I\!\!P(A \cap B)$, *pour tous événements* A *et* B.

(vi) $I\!\!P$ *est continue :* $I\!\!P\Big(\bigcup\limits_{n\in I\!\!N} A_n\Big) = \lim\limits_{n\to\infty} \uparrow I\!\!P(A_n)$, *pour toute suite croissante d'événements* $\{A_n \,|\, n \in I\!\!N\}$ *(id est* $A_n \subset A_{n+1}$ $(\forall\, n)$*) ; et de même* $I\!\!P\Big(\bigcap\limits_{n\in I\!\!N} B_n\Big) = \lim\limits_{n\to\infty} \downarrow I\!\!P(B_n)$, *pour toute suite décroissante d'événements* $\{B_n \,|\, n \in I\!\!N\}$ *(id est* $B_n \supset B_{n+1}$ $(\forall\, n)$*).*

1.1 Exemples

1. Probabilité discrète sur un ensemble $\Omega = \{\omega_1, .., \omega_N\}$ fini :
elle est clairement définie par la liste des probabilités des singletons : $p_j := I\!\!P(\{\omega_j\})$.
Nous avons en effet $I\!\!P(A) = \sum\limits_{\omega_j\in A} p_j$ pour toute partie $A \subset \Omega$.

Réciproquement, toute liste $\{p_1, .., p_N\}$ de réels $p_j \geq 0$ tels que $\sum\limits_{j=1}^{N} p_j = 1$ définit bien (par la même formule) une probabilité unique sur Ω.

2. Probabilité discrète sur $I\!\!N$ (ou sur n'importe quel autre ensemble dénombrable) :
elle est encore définie par la liste des probabilités des singletons : $p_j := I\!\!P(\{\omega_j\})$.
Nous avons en effet $I\!\!P(A) = \sum\limits_{\omega_j\in A} p_j$ pour toute partie $A \subset I\!\!N$.

Réciproquement, toute suite $\{p_j \,|\, j \in I\!\!N\}$ de réels $p_j \geq 0$ tels que $\sum\limits_{j\geq 1} p_j = 1$ définit bien (par la même formule) une probabilité unique sur $I\!\!N$.

3. *Cordes.* On tire une corde au hasard dans un disque de rayon R. Quelle est la probabilité que la longueur ℓ de la corde soit $\geq R$?

a. ℓ varie continûment dans $[0, 2R]$, de sorte que la probabilité cherchée vaut $1/2$.

b. ℓ est déterminée par la distance d de la corde au centre du disque ; d varie continûment dans $[0, R]$, et $\ell = 2\sqrt{R^2 - d^2} \geq R \Leftrightarrow d \leq R\sqrt{3}/2$, de sorte que la probabilité cherchée vaut $\sqrt{3}/2$.

c. ℓ est déterminée par le milieu M de la corde, qui varie continûment dans le disque ; et $\ell \geq R$ a lieu ssi M est dans le disque concentrique de rayon $\sqrt{3}/2$, de sorte que la probabilité cherchée vaut $3/4$.

Explication : la probabilité choisie est très insuffisamment précisée par l'expression "tirage au hasard". Ici on a considéré successivement la probabilité uniforme sur l'ensemble : des longueurs, des distances au centre, des milieux. Ce sont trois probabilités différentes !

<u>Exercice n^o 1.1</u> Est-il plus probable d'obtenir au moins une fois 6 en lançant 4 dés usuels, ou bien d'obtenir au moins une fois un double 6 en lançant 24 fois 2 dés usuels ?

<u>Exercice n^o 1.2</u> On lance n fois de suite 3 dés normaux. Pour quelles valeurs de n la probabilité d'avoir obtenu au moins un 421 dépasse-t-elle $\frac{1}{2}$?

<u>Exercice n^o 1.3</u> On lance 5 pièces de monnaie. Calculer les probabilités des événements suivant : "la 1ère pièce donne face" ; "face sort exactement 2 fois" ; "face sort au plus 3 fois".

<u>Exercice n^o 1.4</u> On lance 10 dés usuels. Calculer les probabilités des événements suivant : "6 ne sort pas" ; "6 sort 1 fois exactement" ; "6 sort 3 fois exactement" ; "6 sort 2 fois au moins" ; "6 sort 3 fois au moins".

<u>Exercice n^o 1.5</u> Une armoire contient 10 paires de chaussures, parmi lesquelles on prélève au hasard 8 chaussures. Quelle est la probabilité d'avoir ainsi k paires de chaussures exactement ?

<u>Exercice n^o 1.6</u> Une urne contient n boules noires et b boules blanches. Deux joueurs X et Y tirent avec remise une boule dans l'urne, tour à tour, X tirant le premier. Quelle est la probabilité que X soit le premier à tirer une boule noire ? Même question sans remise.

<u>Exercice n^o 1.7</u> Une loterie comporte 100 billets, dont les seuls billets gagnants suivant : 1 billet gagne 50 euros, 5 billets gagnent chacun 30 euros, 10 billets gagnent chacun 10 euros. Quelle est la probabilité qu'un acheteur de 3 billets gagne 30 euros (au moins, puis exactement) ?

<u>Exercice n^o 1.8</u> Un joueur X lance 2 dés usuels, et obtient ainsi la somme S.

a) Calculer la probabilité que $S > n$, en fonction des différentes valeurs de l'entier n.

b) Un joueur Y relance les 2 dés et obtient une somme T. Quelles sont les probabilités que $S = T$, que $S > T$, que $S \geq T$?

<u>Exercice n^o 1.9</u> Un sac contient n jetons numérotés de 1 à n, qu'on tire tous 1 à 1, sans remise. $i)$ Calculer $p_n := \mathbb{P}$(au moins un jeton sorte au rang indiqué par son numéro), sa limite p_∞, et majorer $|p_n - p_\infty|$.

$ii)$ Soit $p_n(k) := \mathbb{P}$(exactement k jetons sortent au rang indiqué par leur numéros), pour $k \in \{1, .., n\}$. Déduire de (i) une formule pour $p_n(k)$, puis $\lim_{n\to\infty} p_n(k)$ (pour tout $k \in \mathbb{N}$).

1.2 Probabilités conditionnelles

Définition 3 (Probabilité conditionnelle) *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$, et un événement $C \in \mathcal{T}$, non négligeable. La probabilité conditionnelle relative à C (ou* *"sachant C") est définie par :* $\mathbb{P}(A/C) := \mathbb{P}(A \cap C)/\mathbb{P}(C)$.

On vérifie immédiatement qu'il s'agit encore d'une probabilité sur $(\Omega, \mathcal{T})$.

Exercice n^o 1.2.1 Lancer de 2 dés usuels : $\Omega = \{1, .., 6\}^2$. $\mathbb{P}$ uniforme. Soient X_1 le chiffre indiqué par le premier dé, S la somme des chiffres indiqués par les 2 dés, et $C = \{S = 5\}$. Dresser le tableau des valeurs de $\mathbb{P}(\cdot/C)$, puis de $\mathbb{P}(X_1 = \cdot/C)$.

Exercice n^o 1.2.2 Vous allez chez des gens dont vous savez qu'ils ont 2 enfants, dont au moins une fille. a) Quelle est la probabilité que l'autre enfant soit aussi une fille ?

b) En l'absence de l'information qu'ils ont au moins une fille (pour cette question seulement), mais en voyant une fille ouvrir la porte, quelle est la probabilité que l'autre enfant soit aussi une fille ? c) Une fille vous ouvre la porte ; quelle est la probabilité que l'autre enfant soit aussi une fille ? d) La fille qui vous a ouvert vous dit qu'elle est l'aînée des 2 enfants ; quelle est la probabilité que l'autre enfant soit aussi une fille ?

Exercice n^o 1.2.3 Trois condamnés X, Y, Z sont informés que l'un d'eux, choisi au hasard, va être exécuté, et que les 2 autres vont être libérés. Mais ils ne doivent pas encore savoir qui le hasard a désigné. X demande au geôlier de lui nommer l'un de ses 2 codétenus devant être libéré, arguant que cette information serait innocente, puisqu'il sait que l'un des 2 au moins doit l'être. Le geôlier refuse, arguant que cette information modifierait réellement l'estimation que X peut faire de ses chances. Qui a raison ?

Exercice n^o 1.2.4 Une urne contient b boules blanches et n boules noires. Quand une boule est tirée, on le remet dans l'urne, avec ℓ boules de la même couleur. On effectue ainsi 3 tirages au hasard. a) Quelle est la probabilité que la 1ère boule tirée soit noire sachant que la seconde est blanche ? b) Quelle est la probabilité que la 3ème boule soit noire ?

Exercice n^o 1.2.5 Vous attendez un ami de Vancouver, qui voyage jusqu'à Strasbourg avec changement d'avion à New York, Londres et Francfort. La probabilité d'attentat est estimée à p pour chacun des 4 vols, avec indépendance entre les 4 vols. Votre ami n'arrivant pas, quelle est la probabilité que l'attentat ait eu lieu : a) dans le 1er avion ? b) dans le 2ème avion ? c) dans le 3ème avion ? c) dans le 4ème avion ?

Pour effectuer un calcul, il est très souvent indispensable de pouvoir "distinguer des cas". Cela s'exprime par la formule suivante, très élémentaire et très utile à la fois.

Proposition 3 (Formule des probabilités totales) *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$, et une partition de Ω en événements non négligeables :* $\Omega = \bigsqcup_{j=1}^{N} C_j$. *Alors nous avons* $\mathbb{P}(A) = \sum_{j=1}^{N} \mathbb{P}(A/C_j)\mathbb{P}(C_j)$, *pour tout événement A.*

On a souvent à inverser un conditionnement. Cela se fait simplement, au moyen de la formule élémentaire suivante, très utile aussi.

Proposition 4 (Formule de Bayes) *Fixons un espace de probabilité* $(\Omega, \mathcal{T}, I\!\!P)$*, et une partition de* Ω *en événements non négligeables :* $\Omega = \bigsqcup_{j=1}^{N} C_j$*. Alors nous avons pour tout événement non négligeable* A *et tout* $k \in \{1, .., N\}$ *:*

$$I\!\!P(C_k/A) = I\!\!P(A/C_k)I\!\!P(C_k) \Big/ \sum_{j=1}^{N} I\!\!P(A/C_j)I\!\!P(C_j)\,.$$

<u>Exercice n^o 1.2.6</u> Trois machines U, V, W produisent une même pièce dans une usine. U assure 40% de la production, V 35%, et W le reste. U produit 20% de pièces défectueuses, V 15%, et W 10%.

a) Quelle est la probabilité qu'une pièce prise au hasard soit défectueuse ?

b) Quelle est la probabilité qu'une pièce défectueuse prise au hasard provienne de U ?

<u>Exercice n^o 1.2.7</u> 12% des individus d'une population sont porteurs d'une certaine maladie. Un test relatif à cette maladie est fiable à 95%, dans le cas d'un malade comme dans le cas d'un sujet sain. a) Quelle est la probabilité qu'un individu présentant un test positif soit effectivement malade ? b) Quelle est la probabilité qu'un individu présentant un test négatif soit effectivement sain ?

<u>Exercice n^o 1.2.8</u> Émile possède 5 pièces de monnaie, dont 2 sont normales, 2 ont 2 côtés "face", et une a 2 côtés "pile".

a) Il prend une pièce au hasard et la lance ; quelle est la probabilité qu'il voie "face" ?

b) Il voit "face" ; quelle est la probabilité que l'autre côté de la pièce soit aussi "face" ?

Il relance la même pièce.

c) Quelle est la probabilité que le côté caché de la pièce soit "face" ?

d) Il voit de nouveau "face" ; quelle est la probabilité que l'autre côté de la pièce soit aussi "face" ? Il choisit ensuite au hasard une des autres pièces et la lance.

e) Quelle est la probabilité de voir de nouveau "face" (pour la troisième fois) ?

<u>Exercice n^o 1.2.9</u> Un livre a une probabilité $p > 0$ de se trouver dans une commode comportant k tiroirs, et des chances égales de se trouver dans chacun des tiroirs.

i) On ouvre les $(k-1)$ premiers tiroirs, sans le trouver ; quelle est la probabilité de le trouver dans le dernier tiroir ?

ii) Soit $j \in \{2, .., k-1\}$. On ouvre les $(k-j)$ premiers tiroirs, sans le trouver ; quelle est la probabilité de le trouver dans le dernier tiroir ? dans l'un des j derniers tiroirs ?

<u>Exercice n^o 1.2.10</u> Le quart d'une population est vacciné contre le choléra. Au cours d'une épidémie, on constate qu'il y a parmi les malades un vacciné pour 4 non-vaccinés, et qu'il y a un malade sur 12 parmi les vaccinés. Quelle est la probabilité qu'un non-vacciné tombe malade ? Le vaccin est-il efficace ?

1.3 Événements indépendants

Définition 4 *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$. Deux événements A et B sont dits indépendants lorsque $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

Exemples : 1) "tirer un roi" et "tirer un trèfle", dans un jeu de bridge ou de belote. 2) "obtenir un chiffre pair avec le 1er dé" et "obtenir un 6 avec le 2ème dé", lors du lancer de 2 dés.

Exercice n^o 1.3.1 Montrer que 2 événements A et B sont indépendants ssi A et B^c le sont, et ssi A^c et B^c le sont, ou bien encore (lorsqu'ils sont non négligeables) ssi $\mathbb{P}(A/B) = \mathbb{P}(A)$, et ssi $\mathbb{P}(B/A) = \mathbb{P}(B)$.

Exercice n^o 1.3.2 Un urne contient des jetons numérotés, rouges ou noirs. Lors du tirage d'un jeton la probabilité d'en tirer un rouge est $3/5$; d'en tirer un de numéro impair est $2/3$; d'en tirer un rouge et pair est p. Que vaut la probabilité d'en tirer un noir impair ? Pour quelles valeurs de p les événements "noir" et "impair" sont-ils indépendants ?

Définition 5 *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$. Des événements $A_1, .., A_n$ sont dits indépendants lorsque $\mathbb{P}(A_{i_1} \cap A_{i_2} \cap .. \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2})..\mathbb{P}(A_{i_k})$ pour tous $1 \leq i_1 < .. < i_k \leq n$. Les événements d'une suite sont dits indépendants lorsque toute sous-suite finie est constituée d'événements indépendants.*

Proposition 5 *Les événements $A_1, .., A_n$ sont indépendants ssi $\mathbb{P}(A_1^{\varepsilon_1} \cap A_2^{\varepsilon_2} \cap .. \cap A_n^{\varepsilon_n}) = \mathbb{P}(A_1^{\varepsilon_1})\mathbb{P}(A_2^{\varepsilon_2})..\mathbb{P}(A_n^{\varepsilon_n})$ pour tout choix de $\varepsilon_1, .., \varepsilon_n$, avec $A_j^{\varepsilon_j} = A_j$ ou $A_j^{\varepsilon_j} = A_j^{\mathsf{c}}$.*

Exercice n^o 1.3.3 On jette un dé normal n (≥ 3) fois de suite. Montrer que les événements $\{$les lancers j et k donnent le même résultat$\}$ sont deux à deux indépendants, mais non indépendants.

Exercice n^o 1.3.4 Sur $\Omega := \{a, b, c, d\}$ on définit $\mathbb{P}$ par : $\mathbb{P}(\{a\}) = \alpha$, $\mathbb{P}(\{b\}) = \beta$, $\mathbb{P}(\{c\}) = \gamma$, $\mathbb{P}(\{d\}) = \delta$. Trouver les valeurs de α, β, γ, δ telles que les événements $A := \{b, c\}$, $B := \{c, a\}$, $C := \{a, b\}$ soient 2 à 2 indépendants mais non indépendants.

Exercice n^o 1.3.5 Pour ridiculiser un astrologue (ou un "voyant", ou n'importe quelle autre sorte de charlatan), on le défie de prédire le résultat de 11 lancers successifs d'une pièce de monnaie usuelle. Quelle est la probabilité (en fonction de $n \in \{0, 1, .., 5\}$) qu'il se trompe au plus n fois ?

2 Lemmes de Borel-Cantelli

Ces fameux lemmes, quoique simples, illustrent bien l'hypothèse d'additivité dénombrable faite sur les probabilités. Ils sont très utiles.

Pour toute suite d'ensembles $\{A_n \,|\, n \in \mathbb{N}\}$, $\limsup_n A_n := \cap_n \cup_{\{m \geq n\}} A_m$ est l'ensemble des éléments qui appartiennent à une infinité d'ensembles A_n. (De même, $\liminf_n A_n := \cup_n \cap_{\{m \geq n\}} A_m$ est l'ensemble des éléments qui appartiennent à tous les ensembles A_n sauf au plus un nombre fini.)

Proposition 6 *(lemmes de Borel-Cantelli) Soit $\{A_n \,|\, n \in \mathbb{N}\}$ une suite d'événements.*
1) Si $\sum_n \mathbb{P}(A_n) < \infty$, alors $\mathbb{P}(\limsup_n A_n) = 0$.
2) Si $\sum_n \mathbb{P}(A_n) = \infty$ et si les A_n sont indépendants, alors $\mathbb{P}(\limsup_n A_n) = 1$.

Preuve 1) $\mathbb{P}(\cap_n \cup_{\{m \geq n\}} A_m) \leq \sum_{\{m \geq n\}} \mathbb{P}(A_m) \to 0$.
2) $1 - \mathbb{P}(\cap_n \cup_{\{m \geq n\}} A_m) = \mathbb{P}(\cup_n \cap_{\{m \geq n\}} A_m^c) = \lim \uparrow_n \mathbb{P}(\cap_{\{m \geq n\}} A_m^c)$
$= \lim_n \prod_{\{m \geq n\}} \mathbb{P}(A_m^c) = \lim_n \prod_{\{m \geq n\}} (1 - \mathbb{P}(A_m)) \leq \lim_n \exp\Big(-\sum_{\{m \geq n\}} \mathbb{P}(A_m)\Big) = 0$. $\diamond$

Exemples 1) Un singe tapant sur un clavier au hasard et indéfiniment tapera presque sûrement à un certain moment les œuvres complètes de Victor Hugo sans erreur (et même une infinité de fois). Evidemment, ce resultat ne fournit aucun majorant du temps qu'il faudra attendre pour voir ceci se réaliser une première fois !

2) Si l'univers est infini, il y a une infinité de planètes habitées par des êtres vivant.

Exercice n^o 2.1 Montrer que si une suite de v.a. $\{Y_n \,|\, n \in \mathbb{N}\}$ est telle que $\sum_n \|Y_n\|_2^2$ converge, alors cette suite converge presque sûrement vers 0.

Exercice n^o 2.2 a) Montrer que si Y est une v.a.r. ≥ 0 , alors $\mathbb{E}(Y)^2 \leq \mathbb{E}(Y^2) \times \mathbb{P}(Y > 0)$.
b) Déduire que si $\{A_1, .., A_n\}$ sont des événements non tous négligeables, alors

$$\mathbb{P}\Big(\bigcup_{j=1}^{n} A_j\Big) \geq \Big(\sum_{j=1}^{n} \mathbb{P}(A_j)\Big)^2 \Big/ \Big[\sum_{j=1}^{n} \mathbb{P}(A_j) + 2 \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j)\Big] .$$

c) Déduire que pour le deuxième lemme de Borel-Cantelli il suffit de supposer l'indépendance des événements 2 à 2 (en plus de la divergence de la série).
d) Montrer que s'il existe un $c > 0$ tel que $\mathbb{P}(A_i \cap A_j) \leq c\,\mathbb{P}(A_i)\mathbb{P}(A_j)$ pour tous $i \neq j$ et si $\sum_j \mathbb{P}(A_j) = \infty$, alors $\mathbb{P}(\limsup_j A_j) > 0$.
e) Considérons un jeu illimité de pile ou face, avec $p := \mathbb{P}(\text{“}pile\text{”}) \in]0, 1[$.
Notons X_j la variable de Bernoulli valant 1 ssi le j-ième lancer donne pile, et posons $A_j^k := \{X_j = .. = X_{j+k-1} = 1\}$, pour $j, k \in \mathbb{N}^*$. Montrer que $\mathbb{P}\Big(\cap_k \cup_j A_j^k\Big) = 1$. Interpréter.

3 Variables aléatoires et leur lois

Définition 6 *Une variable aléatoire (“v.a.”) est une fonction V définie sur un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d$, telle que $\{V \in E\} = V^{-1}(E) \in \mathcal{T}$, pour tout E pavé ou ouvert (ou fermé ou borélien) de $\mathbb{R}^d$. Sa loi est la probabilité $\mathbb{P}_V := \mathbb{P} \circ V^{-1}$*

sur $I\!R^d$, définie par : $I\!P_V(E) := I\!P(V^{-1}(E)) = I\!P(V \in E)$. Lorsque $d = 1$, on parle de variable aléatoire réelle, "v.a.r.".

On vérifie immédiatement que la loi d'une variable aléatoire est bien une probabilité (sur $V(\Omega) \subset I\!R^d$, ou directement sur $I\!R^d$, la tribu étant celle des "boréliens", engendrée par les pavés ou par les ouverts de $I\!R^d$). Notons que dans le cas où Ω est discret (fini ou dénombrable), dans la définition ci-dessus la condition de mesurabilité sur V est vide.

On vérifie les propriétés suivantes, qui autorisent toutes les opérations usuelles sur les variables aléatoires.

Proposition 7 *Une fonction de Ω dans $I\!R^d$ est une v.a. ssi ses coordonnées (dans n'importe quelle base de $I\!R^d$) sont des v.a.r.. Une combinaison linéaire de v.a. est encore une v.a.. Un produit de v.a.r. est une v.a.r.. La composée d'une v.a. par une fonction continue de $I\!R^d$ dans $I\!R^{d'}$ est encore une v.a..*

Exemples : La somme et le produit des chiffres indiqués par 2 dés ; les durées de vie de particules fissiles ; le nombre de fois qu'une suite de N lancers d'une pièce de monnaie donne pile ; les temps qu'il faut attendre pour tirer dans une urne, lors de tirages successifs, une boule rouge, une boule verte,... Etc...

La notion d'*espérance* est au cœur de la théorie des probabilités. Il s'agit en fait simplement d'une intégrale, qui est aussi une moyenne.

Définition 7 *L'espérance d'une variable aléatoire V définie sur un espace de probabilité $(\Omega, \mathcal{T}, I\!P)$ est son intégrale par rapport à la mesure $I\!P$:*

$$I\!E(V) := \int_\Omega V \, dI\!P \, .$$

Elle existe lorsque (la norme de) V est intégrable, c'est-à-dire lorsque $I\!E(\|V\|) < \infty$.

Corollaire 1 *Soient V une v.a. et f une fonction continue de $I\!R^d$ dans $I\!R^{d'}$. Alors*

$$I\!E(f \circ V) = \int_\Omega f \circ V \, dI\!P = \int_{I\!R^d} f \, dI\!P_V \, .$$

Lorsque la loi de V est discrète, c'est-à-dire lorsque presque sûrement V ne prend qu'un ensemble fini ou dénombrable de valeurs, alors on a

$$I\!E(f \circ V) = \sum_v v \, I\!P(V = v) \, ,$$

cette série convergeant dès que V est intégrable (ce qui a lieu par exemple dès qu'elle ne prend qu'un nombre fini de valeurs).

Proposition 8 *L'espérance est linéaire :* $\mathbb{E}(\sum_{j=1}^{N} \lambda_j V_j) = \sum_{j=1}^{N} \lambda_j \mathbb{E}(V_j)$ *, pour toutes v.a. intégrables* $V_1, .., V_N$ *et tous réels* $\lambda_1, .., \lambda_N$ *. En particulier, l'espérance d'une v.a. vectorielle intégrable, dans n'importe quelle base, a pour coordonnées les espérances des coordonnées.*

<u>Exercice n^o 3.1</u> On tire au hasard et sans remise toutes les boules d'une urne remplie de r boules rouges et n boules noires.

1) Montrer que que la probabilité de tirer une boule rouge au k-ième tirage ne dépend pas de k. (Il y a au moins 2 solutions bien différentes)

2) Quelle est l'espérance du rang de la k-ième boule rouge ? (On pourra considérer le vecteur $(X_1, X_2 - X_1, .., X_r - X_{r-1}, r+n+1-X_r)$, X_j désignant le rang de la j-ième boule rouge).

<u>Exercice n^o 3.2</u> Une urne U contient initialement 2 boules blanches, et une urne U' 2 boules noires. A chaque instant, on tire au hasard une boule par urne, et on les interchange.

Notons X_k le nombre de boules blanches dans U au k-ième instant, et V_k le vecteur-colonne donnant la loi de X_k.

a) Quelle est la relation entre V_{k+1} et V_k ?
b) Calculer $\lim_{\{k\to\infty\}} \mathbb{P}(X_k = 2)$.
Soient T le premier instant où U contient deux boules noires, $p_k := \mathbb{P}(T \geq k, X_k = 1)$, et $q_k := \mathbb{P}(T \geq k, X_k = 2)$.
c) Exprimer (p_{k+1}, q_{k+1}) en fonction de (p_k, q_k), puis p_{k+1} en fonction de (p_k, p_{k-1}).
d) Déduire la valeur de p_k, puis la loi de T. Que vaut $\mathbb{P}(T = \infty)$?

<u>Exercice n^o 3.3</u> Un marchand de journaux a X clients par jour, X étant une v.a. entière intégrable, de loi supposée connue. Il gagne a par journal vendu, perd b par journal invendu, et perd c par client insatisfait. Quel est le nombre n de journaux qu'il doit commander par jour pour optimiser son gain moyen ?

<u>Exercice n^o 3.4</u> Montrer que pour toute v.a.r. $Z \geq 0$, on a $\mathbb{E}(Z) = \int_0^\infty \mathbb{P}(Z > s)\, ds$. Particulariser au cas où Z prend ses valeurs dans $\mathbb{N}$.

<u>Exercice n^o 3.5</u> Au cours d'un jeu illimité de pile ou face avec $\mathbb{P}$(pile)$= p$, on note X_k le rang de la k-ième apparition de "pile". Calculer la loi de X_k, son espérance et sa variance. Pour $n \in \mathbb{N}^*$, calculer $\mathbb{P}\Big((\exists\, k \in \mathbb{N}^*)\ X_k = n\Big)$.

Définition 8 *Une variable aléatoire* V *(définie sur un espace de probabilité* $(\Omega, \mathcal{T}, \mathbb{P})$ *et à valeurs dans* $\mathbb{R}^d$*) admet la* <u>*densité*</u> h *lorsque sa loi est donnée par*

$$\mathbb{E}(f \circ V) := \int_{\mathbb{R}^d} f(v)\, h(v)\, dv \,, \quad \textit{pour toute fonction continue bornée } f : \mathbb{R}^d \to \mathbb{R} .$$

Notons que la densité h est nécessairement une fonction mesurable de $\mathbb{R}^d$ dans $\mathbb{R}_+$, d'intégrale égale à 1.

Attention, la plupart des variables aléatoires ne sont ni discrètes ni à densité ! Il suffit de songer à un v.a. V valant une v.a. discrète X avec probabilité $1/2$ et une v.a. à densité Y avec probabilité $1/2$: $\mathbb{P}_V = (\mathbb{P}_X + \mathbb{P}_Y)/2$. Imaginer par exemple un tir sur cible, avec probabilité $1/2$ de rater la cible, auquel cas la v.a. prend la valeur disons -1, et probabilité $1/2$ d'atteindre la cible, auquel cas la v.a. prend ses valeurs dans le disque formé par la cible, avec par exemple la loi uniforme.

Cela dit, les lois usuelles, celles qu'on rencontre le plus souvent, sont discrète ou à densité (on dit aussi : absolument continues).

Définition 9 *Une v.a. V (définie sur un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$) est dite de carré intégrable ou dans L^2 lorsque $\mathbb{E}(\|V\|^2) < \infty$. Elle est alors nécessairement intégrable.*

La variance d'une v.a.r. V de carré intégrable est :

$$Var(V) := \mathbb{E}\Big[|V - \mathbb{E}(V)|^2\Big] = \mathbb{E}(V^2) - \mathbb{E}(V)^2 .$$

La covariance de deux v.a.r. V, V' de carré intégrable est :

$$Cov(V, V') := \mathbb{E}\Big[(V - \mathbb{E}(V) \times (V' - \mathbb{E}(V')\Big] = \mathbb{E}(VV') - \mathbb{E}(V)\mathbb{E}(V') .$$

La matrice de covariance (ou de dispersion) d'une v.a. $V = (V_1, .., V_d) \in \mathbb{R}^d$ de carré intégrable est :

$$K_V := \mathbb{E}\Big[(V - \mathbb{E}(V) \times^t (V - \mathbb{E}(V)\Big] = \mathbb{E}(V\,{}^tV) - \mathbb{E}(V)\mathbb{E}({}^tV) = \Big(\Big(Cov(V_i, V_j)\Big)\Big)_{1\le i,j\le d} .$$

Corollaire 2 *La covariance est bilinéaire, symétrique, positive, et $Var(V)$ est nulle ssi V est p.s. constante. En outre $Var(\lambda V + V') = \lambda^2 Var(V) + 2\,\lambda\, Cov(V, V') + Var(V')$.*

Proposition 9 *(Inégalité de Schwarz) La covariance est majorée par le produit des écart-types : $|Cov(V, V')| \le \sigma(V)\sigma(V')$, où l'écart-type de la v.a.r. V est $\sigma(V) := \sqrt{Var(V)}$. De sorte que le coefficient de corrélation linéaire $\varrho(V, V') := \dfrac{Cov(V, V')}{\sigma(V)\sigma(V')}$ (défini pour V et V' v.a.r. de carré intégrable et non p.s. constantes) appartient à $[-1, 1]$. Il vaut ± 1 ssi $V = aV' + b$ (p.s. pour a, b réels fixes).*

Exercice n^o 3.6 a) Vérifier que $K_V = \mathbb{E}(V\,{}^tV) - \mathbb{E}(V)^t\mathbb{E}(V) = \Big(\Big(Cov(V_i, V_j)\Big)\Big)_{1\le i,j\le d}$.
b) Vérifier que pour tout $u \in \mathbb{R}^d$ on a $Var({}^tuV) = {}^tuK_V u = \sum_{1\le i,j\le d} u_i u_j Cov(V_i, V_j)$.
c) Montrer que K_V est une matrice symétrique positive.
d) Montrer que K_V est inversible ssi il n'existe pas d'hyperplan de $\mathbb{R}^d$ contenant p.s. V.

e) Soit M une matrice de $\mathcal{M}_{n,d}(\mathbb{R})$. Calculer $\mathbb{E}(MV)$ et K_{MV}. Même question pour AV, si A est une application affine de $\mathbb{R}^d$ dans $\mathbb{R}^n$.

Exercice n^o 3.7 Inégalité de Markov, ou de Bienaymé-Tchebitchev :

Vérifier que $\mathbb{P}[|V| \geq v] \leq \mathbb{E}[|V|^k]/v^k$, pour tous $v > 0$ et $k \geq 0$.

Exercice n^o 3.8 Vous écrivez chaque jour avec probabilité 1 si vous n'avez pas écrit la veille, et avec probabilité 1/2 sinon. Montrez que vous écrivez ainsi en moyenne 243 lettres par an. (Considérer pour chaque jour la variable indicatrice de l'événement "écrire".)

Proposition 10 *(Inégalité de Jensen) Pour toute v.a.r. intégrable V et toute fonction convexe ψ de $\mathbb{R}$ dans $\mathbb{R}$, telle que $\psi \circ V$ soit positive ou intégrable, on a*

$$\mathbb{E}[\psi \circ V] \geq \psi(\mathbb{E}[V]) .$$

Exemples : on a $|x_1 + .. + x_n|^p \leq n^{p-1} (|x_1|^p + .. + |x_n|^p)$ pour tous $n \in \mathbb{N}^*$, $p \in [1, \infty[$, et $x_1, .., x_n \in \mathbb{R}$; et $(\mathbb{E}[|V|^p])^{1/p} \geq (\mathbb{E}[|V|^q])^{1/q}$ si $p \geq q > 0$.

4 Lois usuelles

4.1 Lois usuelles discrètes

Définition 10 *La loi uniforme sur un espace de probabilité fini est celle qui attribue la même valeur à tous les singletons.*

Corollaire 3 *Si $\mathbb{P}$ est uniforme sur Ω fini, alors $\mathbb{P}(A) = \mathrm{Card}(A)/\mathrm{Card}(\Omega)$ pour toute partie A de Ω.*

Exercice n^o 4.1.1 Calculer l'espérance et la variance d'une v.a. uniforme sur un intervalle de $\mathbb{Z}$.

Définition 11 *La loi de Bernoulli de paramètre p, notée $\mathcal{B}(p)$, est sur l'espace $\{0, 1\}$ la loi (de probabilité) qui attribue la valeur p au singleton $\{1\}$. Ici $0 \leq p \leq 1$.*

Définition 12 *La loi binômiale de paramètres n et p, notée $\mathcal{B}(n, p)$, est sur l'espace $\{0, .., n\}$ la loi (de probabilité) qui attribue la valeur $C_n^k\, p^k (1-p)^{n-k}$ au singleton $\{k\}$. Ici $n \in \mathbb{N}^*$ et $0 \leq p \leq 1$.*

Remarque 1 C'est la loi de la somme de n v.a. indépendantes de même loi $\mathcal{B}(p)$.

Exercice n^o 4.1.2 Montrer que l'espérance et la variance d'une v.a. de loi $\mathcal{B}(n,p)$ valent np et $np(1-p)$.

Définition 13 *La loi hypergéométrique de paramètre N,n,p, notée $\mathcal{H}(N,n,p)$, est sur l'espace $\{0,..,n\}$ la loi (de probabilité) qui attribue la valeur $\dfrac{C_{Np}^k\,C_{N(1-p)}^{n-k}}{C_N^n}$ au singleton $\{k\}$. Ici $N,Np,n\in\mathbb{N}$, $n\le N$, et $0\le p\le 1$.*

Remarque 2 C'est la loi du nombre d'éléments présentant un caractère donné K dans un échantillon de taille n, prélevé au hasard (uniformément parmi les parties de cardinal n) dans un ensemble de cardinal N, dont une proportion p présente le caractère K.

Exercice n^o 4.1.3 a) Montrer que l'espérance d'une v.a. de loi $\mathcal{H}(N,n,p)$ vaut np.
b) Calculer sa variance. c) Vérifier que $\lim\limits_{N\to\infty}\mathcal{H}(N,n,p)(k)=\mathcal{B}(n,p)(k)$, pour n,p,k fixés.

Définition 14 *La loi géométrique de paramètre p, notée $\mathcal{G}(p)$, est sur l'espace $\mathbb{N}^*$ la loi (de probabilité) qui attribue la valeur $(1-p)^{n-1}p$ au singleton $\{n\}$. Ici $0<p<1$.*

Remarque 3 C'est la loi du nombre N de tentatives nécessaires pour obtenir un résultat positif (succès), lors d'une suite de tentatives identiques et indépendantes, p étant la probabilité de succès à chaque tentative. On a $\mathbb{P}(N>n)=(1-p)^n$ pour tout $n\in\mathbb{N}$.

Exercice n^o 4.1.4 a) Les lois géométriques vérifient la propriété de non-vieillissement :
$\mathbb{P}(N>n+m/N>m)=\mathbb{P}(N>n)$ pour tous $n,m\in\mathbb{N}$.
b) Y a-t-il d'autres lois sur $\mathbb{N}^*$ qui vérifient cette propriété ?

Exercice n^o 4.1.5 Montrer que l'esp. et la variance d'une v.a. de loi $\mathcal{G}(p)$ valent $\frac{1}{p}$ et $\frac{1-p}{p^2}$.

Définition 15 *La loi de Poisson de paramètre λ, notée $\mathcal{P}(\lambda)$, est sur l'espace $\mathbb{N}$ la loi (probabilité) qui attribue la valeur $e^{-\lambda}\dfrac{\lambda^n}{n!}$ au singleton $\{n\}$. Ici $\lambda>0$.*

Exercice n^o 4.1.6 a) Montrer que l'esp. et la variance d'une v.a. de loi $\mathcal{P}(\lambda)$ valent λ.
b) Vérifier que $\lim\limits_{np\to\lambda}\mathcal{B}(n,p)(k)=\mathcal{P}(\lambda)(k)$, pour $\lambda>0$ fixé et $n\to\infty$.

Exercice n^o 4.1.7 Quelle est la valeur la plus probable pour une variable aléatoire poissonnienne de paramètre λ ?

Exercice n^o 4.1.8 Un trousseau de n clefs contient une seule clef ouvrant une serrure donnée. On les essaie l'une après l'autre au hasard. Calculer la loi, l'espérance et la variance du du nombre d'essais nécessaires. Même question si on réessaie à chaque fois une clef au hasard sans avoir écarté la précédente.

4.2 Lois usuelles à densité

Définition 16 *La loi uniforme sur un ouvert (ou un fermé) O de $I\!R^d$, notée $\mathcal{U}(O)$, est la loi admettant une densité constante (par rapport à la mesure de Lebesgue) sur O.*

Exercice n^o 4.2.1 Calculer l'esp. et la variance d'une v.a. uniforme sur un intervalle de $I\!R$.

Définition 17 *La loi exponentielle de paramètre λ, notée $\mathcal{E}(\lambda)$, est la loi admettant sur $I\!R_+$ la densité $t \mapsto \lambda\, e^{-\lambda t}$. Ici $\lambda > 0$.*

Exercice n^o 4.2.2 a) Les lois exponentielles vérifient la propriété de non-vieillissement : si la loi de Y est $\mathcal{E}(\lambda)$, alors $I\!P(Y > s+t/Y > s) = I\!P(Y > t)$ pour tous $s, t > 0$.

b) Y a-t-il d'autres lois à densité sur $I\!R_+$ qui vérifient cette propriété ?

c) En déduire la loi de la durée de vie d'un atome fissile, en fonction de sa demi-vie.

Exercice n^o 4.2.3 Montrer que l'espérance et l'écart-type d'une v.a. de loi $\mathcal{E}(\lambda)$ valent $\frac{1}{\lambda}$.

Définition 18 *Une variable aléatoire réelle est dite gaussienne centrée réduite ou normale centrée réduite ou gaussienne standard lorsqu'elle admet (sur $I\!R$) la densité : $t \longmapsto \exp(-t^2/2)/\sqrt{2\pi}$. Une variable aléatoire réelle X est dite gaussienne ou normale lorsqu'il existe $m \in I\!R$ et $\sigma > 0$ tels que $(X-m)/\sigma$ soit normale centrée réduite. On dit alors que la loi de X est $\mathcal{N}(m, \sigma^2)$.*

Exercice n^o 4.2.4 a) Vérifier que $m = I\!E(X)$ et $\sigma^2 = Var(X)$, et que $X \in \cap_{p<\infty} L^p$.

b) Vérifier que $\frac{A^2\, e^{-x^2/2}}{(A^2+1)\, x} \leq \int_x^\infty e^{-t^2/2} dt \leq \frac{e^{-x^2/2}}{x}$ sur $[A, \infty[$, pour tout $A > 0$. Donner un équivalent de $I\!P(X > x)$ quand $x \to +\infty$.

c) Montrer que la densité de X est $t \longmapsto \exp(-(t-m)^2/2\sigma^2)/\sigma\sqrt{2\pi}$.

Exercice n^o 4.2.5 Pour être en cours à 8h, un étudiant en voiture a le choix entre un trajet sur petite route, dont la durée (en minutes) X suit la loi normale de moyenne 35,2 et de variance 25, et un trajet sur autoroute, dont la durée Y suit la loi normale de moyenne 40 et de variance 4. Il désire arriver à l'heure. Quel trajet doit-il préférer s'il part à 7h15 ? Et s'il part à 7h30 ?

Définition 19 *Une variable aléatoire V à valeurs dans $I\!R^d$ est dite gaussienne ou normale lorsque $u \cdot V = {}^t\!u\, V$ est une v.a.r. normale ou p.s. constante pour tout $u \in I\!R^d$.*

On note $\mathcal{N}(m, K)$ la loi d'un vecteur gaussien d'espérance m et de matrice de covariance K. Une probabilité (resp. une densité) est dite gaussienne lorsqu'elle est la loi (resp. la densité) d'un vecteur gaussien.

Exercice n^o 4.2.6 a) Vérifier que si V est un vecteur gaussien de $\mathbb{R}^d$ et si A est une application affine de $\mathbb{R}^d$ dans $\mathbb{R}^n$, alors AV est un vecteur gaussien.

b) Montrer que si V est un vecteur gaussien, alors ses coordonnées dans n'importe quelle base sont gaussiennes (ou p.s. constantes).

Proposition 11 *Soit K une matrice réelle symétrique positive, de format $d \times d$ et de rang r, et soit M une matrice réelle de format $d \times r$ telle que $K = M \times {}^t M$. Soit $m \in \mathbb{R}^d$. Soit V un vecteur de $\mathbb{R}^r$ formé de coordonnées gaussiennes standard indépendantes. Alors $m + MV$ est un vecteur gaussien de loi $\mathcal{N}(m, K)$.*

Corollaire 4 *Si la matrice de covariance d'un vecteur gaussien d-dimensionnel de loi $\mathcal{N}(m, K)$ est inversible, alors ce vecteur admet la densité sur $\mathbb{R}^d$:*

$$v \longmapsto (2\pi)^{-d/2} \, (det\, K)^{-1/2} \, \exp\Big(-\tfrac{1}{2}\, {}^t(v-m) K^{-1} (v-m) \Big) \, .$$

Preuve Soit V un vecteur de $\mathbb{R}^d$ formé de coordonnées gaussiennes standard indépendantes. Soit M une matrice réelle de format $d \times d$ telle que $K = M \times {}^t M$. Alors M est inversible, et nous avons pour toute fonction-test f, par changement de variable :

$$\mathbb{E}(f(m+MV)) = \int_{\mathbb{R}^d} f(m+Mx) \prod_{j=1}^{d} (2\pi)^{-1/2} \, e^{-x_j^2/2} dx = \int_{\mathbb{R}^d} f(m+Mx) \, (2\pi)^{-d/2} e^{-\, {}^t x x/2} dx$$

$$= \int_{\mathbb{R}^d} f(v) \, (2\pi)^{-d/2} \, e^{-\, {}^t(M^{-1}(v-m)) \times (M^{-1}(v-m))/2} \, |det\,(M^{-1})| \, dv$$

$$= \int_{\mathbb{R}^d} f(v) \, (2\pi)^{-d/2} \, (det\, K)^{-1/2} \, \exp\Big(-\tfrac{1}{2}\, {}^t(v-m) K^{-1} (v-m) \Big) \, dv \, . \ \diamond$$

Exercice n^o 4.2.7 : Simulation Soit F une fonction de répartition sur $\mathbb{R}$. Pour tout $p \in [0,1]$, posons $G(p) := \inf\{x \in \mathbb{R} \,|\, F(x) \geq p\}$. (Nota Bene : $\inf \mathbb{R} = -\infty$ et $\inf \emptyset = +\infty$.)

a) Justifier l'existence dans $\mathbb{R}$ de $G(p)$ si $p \in]0,1[$. b) Montrer que $(\forall x \in \mathbb{R})$ $G(F(x)) \leq x$.

c) Montrer que si $G(p) \in \mathbb{R}$, alors $F(G(p)) \geq p$. d) Montrer que $G(p) \leq x \Leftrightarrow F(x) \geq p$.

e) Montrer que si U est une variable aléatoire de loi uniforme sur $[0,1]$, alors $G \circ U$ admet F pour fonction de répartition. Nota Bene : Ceci est utilisé pour simuler des variables aléatoires.

f) Que vaut G lorsque F est bijective de $\mathbb{R}$ sur $]0,1[$? Comment simuler la loi $\mathcal{E}(\lambda)$?

4.3 Quelques lois de la mécanique statistique

La mécanique statistique classique s'appuie sur le modèle statistique de Maxwell, dans lequel les N particules considérées sont réparties parmi n états, dont n_j ont un niveau d'énergie e_j, de sorte que $n = n_1 + .. + n_d$. Soit N_j le nombre de particules ayant le niveau d'énergie e_j, de sorte que $N = N_1 + .. + N_d$. Quelle est la loi du vecteur $(N_1, .., N_d)$?

La réponse est la loi multinômiale (généralisant la loi binômiale) :

$$\mathbb{P}\Big((N_1, .., N_d) = (k_1, .., k_d)\Big) = \frac{N!}{k_1! \times .. \times k_d!} \times \frac{n_1^{k_1} \times .. \times n_d^{k_d}}{n^N} \ ,$$

pour tout $(k_1, .., k_d) \in \mathbb{N}^d$ tel que $k_1 + .. + k_d = N$. En effet, le coefficient multinômial $\frac{N!}{k_1! \times .. \times k_d!}$ (c'est un coefficient binômial lorsque $d = 2$) est le nombre de partitions de l'ensemble des N particules en d classes de cardinaux respectifs $k_1, .., k_d$; et d'autre part $n_1^{k_1} \times .. \times n_d^{k_d}$ est le nombre de façons de répartir k_1 particule d'énergie e_1 parmi n_1 états,.., k_d particule d'énergie e_d parmi n_d états ; tandis que n^N est le nombre total de répartitions des N particules parmi les n états.

Notons que cela montre du même coup la généralisation de la formule du binôme :

$$\sum_{\{(k_1,..,k_d)\in \mathbb{N}^d \,|\, k_1+..+k_d=N\}} \frac{N!}{k_1! .. k_d!} \times z_1^{k_1} .. z_d^{k_d} = (z_1 + .. + z_d)^N \ , \quad \text{pour tous } \ z_1, .., z_d \in \mathbb{C} \,,$$

et en particulier : $\sum_{\{(k_1,..,k_d)\in \mathbb{N}^d \,|\, k_1+..+k_d=N\}} \frac{N!}{k_1! .. k_d!} = d^N$ pour tous $N, d \in \mathbb{N}^*$.

Ce modèle de Maxwell n'est en fait pas vraiment réaliste, du fait qu'il suppose les particules distinguables, ce qui n'est pas le cas en mécanique quantique. Supposant les particules indistinguables, on a le choix entre deux options : on peut les répartir entre les différents états soit avec répétitions, soit sans répétition. L'option avec répétitions est la statistique de Bose-Einstein, relative aux particules de spin entier (photons, mésons,..), et l'option sans répétition est la statistique de Fermi-Dirac, relative aux particules de spin demi-entier (électrons, neutrons, protons,..), qui obéissent au principe d'exclusion de Pauli.

Pour la statistique de Fermi-Dirac, la loi généralise la loi hypergéométrique :

$$\mathbb{P}\Big((N_1, .., N_d) = (k_1, .., k_d)\Big) = \frac{C_{n_1}^{k_1} \times .. \times C_{n_d}^{k_d}}{C_n^N} \ ,$$

pour tout $(k_1, .., k_d) \in \mathbb{N}^d$ tel que $k_1 + .. + k_d = N$. En effet, il faut simplement répartir k_1 particules parmi n_1 états d'énergie e_1, ..., k_d particules parmi n_d états d'énergie e_d, cependant qu'il y a en tout N particules à répartir parmi n états d'énergie ; avec à chaque fois une seule particule au plus par état, ce qui revient exactement à choisir le sous-ensemble des états occupés parmi les états disponibles, d'où les coefficients binômiaux.

Pour la statistique de Bose-Einstein, la loi est

$$\mathbb{P}\Big((N_1, .., N_d) = (k_1, .., k_d)\Big) = \frac{C_{n_1+k_1-1}^{k_1} \times .. \times C_{n_d+k_d-1}^{k_d}}{C_{n+N-1}^N} \ ,$$

pour tout $(k_1,..,k_d)\in \mathbb{N}^d$ tel que $k_1+..+k_d=N$. Le décompte est en effet le même que pour la statistique de Fermi-Dirac, les combinaisons devant être remplacées par des combinaisons avec répétitions. Or C^N_{n+N-1} est le nombre des combinaisons avec répétitions de N particules placées dans n états ; ce qui se voit en juxtaposant sur un segment les N particules, et en considérant la bijection qui à chaque ensemble de $(n-1)$ signes délimitant n sous-segments (figurant les états) associe la combinaison avec répétitions ainsi constituée.

Remarquons une structure de produit commune à ces trois “statistiques” : elles sont toutes trois de la forme $\mathbb{P}\Big((N_1,..,N_d)=(k_1,..,k_d)\Big)=c(N,n)\prod\limits_{j=1}^{d}\lambda(n_j,k_j)$.

Exercice n^o 4.3.1 : Montrer que lorsque le nombre n des états tend vers l'infini de façon que les proportions $n_1/n,..,n_d/n$ convergent vers $p_1,..,p_d$, alors les statistiques de Fermi-Dirac et de Bose-Einstein convergent vers la loi multinômiale (analogue à la statistique de Maxwell).

5 Variables aléatoires indépendantes

Définition 20 *Fixons un espace de probabilité $(\Omega,\mathcal{T},\mathbb{P})$. Des v.a. (définies sur $(\Omega,\mathcal{T})$) $V_1,..,V_n,..$ sont dites indépendantes lorsque pour tous $B_1,..,B_n,..$, les événements $\{V_1\in B_1\},..,\{V_n\in B_n\},..$ sont indépendants (revoir la définition 5, section 1.3).*

Proposition 12 *Les v.a. $V_1,..,V_n,..$ sont indépendantes ssi*

$$\mathbb{E}\Big(f_1\circ V_1\times..\times f_n\circ V_n\Big)=\mathbb{E}\Big(f_1\circ V_1\Big)\times..\times\mathbb{E}\Big(f_n\circ V_n\Big)$$

pour tous $n\in\mathbb{N}^*$ et toutes fonctions mesurables positives $f_1,..,f_n$.

Cela se dit aussi sous la forme : la loi de $(V_1,..,V_n,..)$ est le produit des lois des V_n.

En particulier, lorsque les v.a.r. indépendantes V_n admettent des densités h_n sur $\mathbb{R}$, alors la v.a. $(V_1,..,V_d)$ admet la densité $h_1\otimes..\otimes h_d$ sur $\mathbb{R}^d$.

Remarque 4 1) Les événements $A_1,..,A_n,..$ sont indépendants ssi les v.a. $1_{A_1},..,1_{A_n},..$ sont indépendantes (revoir la proposition 5, section 1.3).

2) Si des v.a.r. $V_1,..,V_n,..$ sont indépendantes, alors leurs covariances sont nulles. Mais la réciproque est fausse.

3) Les v.a. discrètes $V_1,..,V_n,..$ sont indépendantes ssi pour tous $v_1,..,v_n,..$ les événements $\{V_1=v_1\},..,\{V_n=v_n\},..$ sont indépendants.

Exemples : les résultats de différents dés, de lancers successifs d'une pièce, de tirages successifs (roulette,..), les durées de vie de différents atomes fissiles (en l'absence de toute réaction en chaîne), etc...

Exercice n^o 5.1 a) Soient U_1, U_2, U_3 trois v.a. indépendantes, uniformes sur $[0,1]$. Montrer qu'elles sont p.s. 2 à 2 distinctes. On les réordonne en $V_1 < V_2 < V_3$. Montrer que (V_1, V_2, V_3) admet la densité $6 \times 1_{\{0<v_1<v_2<v_3<1\}}$, et en déduire les densités de V_1, V_2, V_3.

b) Soient $U_1, .., U_n$ des v.a. indépendantes, uniformes sur $[0,1]$, et $J_n := \min\{U_1, .., U_n\}$, $M_n := \max\{U_1, .., U_n\}$. Calculer $\mathbb{P}(x < J_n, M_n < y)$, et en déduire la loi (conjointe) de (J_n, M_n). c) Que vaut $\lim\limits_{n\to\infty} \mathbb{P}(J_n < x < y < M_n)$?

Exercice n^o 5.2 Soient $Y_1, .., Y_n$ des v.a. indépendantes, exponentielles. Quelle est la loi de $\min\{Y_1, .., Y_n\}$?

Exercice n^o 5.3 Soient $X_1, .., X_n$ n v.a. indépendantes, X_j étant poissonnienne de paramètre λ_j. a) Quelle est la loi de $X_1 + X_2$, puis de $X_1 + .. + X_n$?

b) Calculer la loi, l'espérance et la variance de X_j sachant $X_1 + .. + X_n$.

c) Soient Y_j, $j \in \mathbb{N}$, des v.a. indépendantes à valeurs dans $\{1, .., n\}$, de même loi donnée par $\mathbb{P}(Y_j = k) = \alpha_k$, pour $1 \le k \le n$. Soient N une variable de Poisson de paramètre λ, indépendante des Y_j, et $N_k := \sum\limits_{j=1}^{N} 1_{\{Y_j = k\}}$. Montrer que les variables $N_1, .., N_k$ sont poissonniennes indépendantes.

Exercice n^o 5.4 Soient X_1, X_2 deux v.a.r. indépendantes de lois exponentielles de paramètres λ_1, λ_2. a) Calculer les lois de $J := \min\{X_1, X_2\}$, $M := \max\{X_1, X_2\}$.

b) Supposant que $\lambda_1 = \lambda_2$, montrer que J et $M - J$ sont des variables indépendantes.

Exercice n^o 5.5 Soient X, Y deux variables aléatoires indépendantes, de densité respectivement f, g, avec $n \in \mathbb{N}^*$, $f(x) = 1_{\{x>0\}} \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}$, et $g(y) = 1_{\{y>0\}} \lambda\, e^{-\lambda y}$.

Calculer la densité de $X + Y$, $\mathbb{E}(X + Y)$, et $Var(X + Y)$.

Exercice n^o 5.6 Calculer la densité du carré d'une v.a.r. normale standard, puis de la somme de 2 tels carrés, puis du quotient de deux v.a.r. normales standards indépendantes.

Exercice n^o 5.7 Soient X et Y 2 v.a.r. indépendantes, admettant toutes 2 la densité $t \mapsto t^{-2}\, 1_{[1,\infty[}(t)$. On pose $U := XY$ et $V := X/Y$.

a) Calculer les lois de (U, V), de U, et de V. U et V sont-elles indépendantes ?

b) Calculer $\mathbb{E}(U^{-1/2} V^{-1})$.

Exercice n^o 5.8 Trois clients A, B, C arrivent au même temps 0 à la poste, où 2 guichets sont ouverts, qu'occupent A et B tout de suite. C remplace le premier des 2 qui a terminé. On admet que les temps de service X, Y, Z requis par ces 3 clients sont des v.a.r.i.i.d. de même loi $\mathcal{E}(\lambda)$.

a) Quelle est la loi du temps d'attente T de C ? b) Calculer la probabilité que C termine (et parte) le dernier. c) Calculer la loi du temps du dernier départ.

Exercice n^o 5.9 Quand la somme de 2 variables aléatoires binômiales indépendantes est-elle binômiale ?

Exercice n^o 5.10 Au cours d'un jeu illimité de pile ou face avec $\mathbb{P}(\text{pile})= p$, on note X_k le rang de la k-ième apparition de "pile". Calculer la loi de X_k, son espérance et sa variance. Pour $n \in \mathbb{N}^*$, calculer $\mathbb{P}\Big((\exists\, k \in \mathbb{N}^*)\ X_k = n\Big)$.

Exercice n^o 5.11 On effectue n tirages indépendants avec remise dans une urne contenant une proportion p_j de boules marquées j, pour $1 \le j \le r$, $r > 1$ étant fixé. On note N_j le nombre de boules marquées j qu'on tire ainsi. Préciser la loi du vecteur $N := (N_1, .., N_r)$, et calculer l'espérance et la variance de N_j, la covariance de N_j et N_k, et le nombre moyen de j tels que $N_j = 0$.

6 Transformées de Laplace et de Fourier

Définition 21 *Soient $V_1, .., V_n$ des v.a. (définies sur un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$) à valeurs dans $\mathbb{R}_+$. La transformée de Laplace de la variable V_1 est la fonction de $\mathbb{R}_+$ dans $[0,1[$ définie par : $\lambda \mapsto \mathbb{E}(e^{-\lambda V_1})$. La transformée de Laplace de la variable $(V_1, .., V_n)$ est la fonction de $(\mathbb{R}_+)^n$ dans $[0,1[$ définie par : $(\lambda_1, .., \lambda_n) \mapsto \mathbb{E}(e^{-\lambda_1 V_1 - .. - \lambda_n V_n})$.*

Remarque 5 Dans le cas d'une v.a. positive discrète V_1, on écrit la transformée de Laplace sous une forme légèrement différente. C'est (modulo un changement de variable trivial) ce qu'on nomme la fonction génératrice, définie sur $]0,1]$ par : $s \mapsto \mathbb{E}(s^{V_1})$.

Définition 22 *Soient $V_1, .., V_n$ des v.a.r. (sur un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$). La transformée de Fourier de la variable V_1 est la fonction de $\mathbb{R}$ dans (le disque unité de) $\mathbb{C}$ définie par : $t \mapsto \mathbb{E}(e^{\sqrt{-1}\, t V_1})$. La transformée de Fourier de la variable $(V_1, .., V_n)$ est la fonction de $\mathbb{R}^n$ dans (le disque unité de) $\mathbb{C}$ définie par : $(t_1, .., t_n) \mapsto \mathbb{E}(e^{\sqrt{-1}\,(t_1 V_1 + .. + t_n V_n)})$.*

Théorème 1 *Les transformées de Laplace et de Fourier sont injectives : la loi d'un vecteur aléatoire $(V_1, .., V_n)$ est déterminée par sa transformée de Fourier (et aussi par sa transformée de Laplace s'il est à coordonnées positives).*

Remarque 6 Noter que la transformée de Fourier (ou de Laplace) d'une v.a. est en fait celle de sa loi, et que l'injectivité du théorème 1 ne vaut que pour les lois, et non pour les variables elles-mêmes. Ainsi deux v.a. ayant la même loi ont la même transformée de Fourier, sans pour autant devoir être égales. Il suffit de considérer par exemple une variable normale centrée et la variable opposée.

Corollaire 5 *Des v.a. $V_1, .., V_n$ sont indépendantes ssi la transformée de Fourier du vecteur aléatoire $(V_1, .., V_n)$ est le produit (tensoriel) des transformées de Fourier des V_j :*

$$\mathbb{E}(e^{\sqrt{-1}\,(t_1 \cdot V_1 + .. + t_n \cdot V_n)}) = \mathbb{E}(e^{\sqrt{-1}\ t_1 \cdot V_1}) .. \mathbb{E}(e^{\sqrt{-1}\ t_n \cdot V_n}), \quad \forall\, t_1 \in \mathbb{R}^{d_1}, .., t_n \in \mathbb{R}^{d_n}.$$

Même énoncé avac la transformation de Laplace si les V_j sont à coordonnées positives.

Preuve Dans le sens direct, c'est simplement la proposition 12, avec $f_j(v_j) := e^{\sqrt{-1}\ t_j \cdot v_j}$. Pour la réciproque, il suffit de considérer des v.a. $V'_1, .., V'_n$ indépendantes et telle que chaque V'_j ait la même loi que V_j. Le sens direct montre que le membre de droite de la formule de l'énoncé est la transformée de Fourier de $(V'_1, .., V'_n)$. Donc le théorème 1 assure que $(V_1, .., V_n)$ a la même loi que $(V'_1, .., V'_n)$, et donc que $V_1, .., V_n$ sont indépendantes. $\diamond$

Proposition 13 *La transformée de Fourier d'une v.a.r. normale N de loi $\mathcal{N}(m, \sigma^2)$ est donnée par*

$$\mathbb{E}(e^{\sqrt{-1}\ t N}) = e^{\sqrt{-1}\ tm - \sigma^2 t^2/2}, \quad \forall\, t \in \mathbb{R}.$$

Corollaire 6 *La transformée de Fourier d'un vecteur gaussien V (à valeurs dans $\mathbb{R}^d$, identifié à une colonne, et de matrice de covariance K_V) est donnée par*

$$\mathbb{E}(e^{\sqrt{-1}\ {}^t u V}) = \exp\left(\sqrt{-1}\ {}^t u \mathbb{E}(V) - \tfrac{1}{2}\, {}^t u K_V u\right), \quad \forall\, u \in \mathbb{R}^d.$$

Exercice n^o 6.1 a) Déduire le corollaire 6 de la définition 19 et de la proposition 13.

b) Soient N gaussienne standard, et ε indépendante de N et uniforme sur ± 1. Montrer que εN est gaussiennne standard, mais que le vecteur $(N, \varepsilon N)$ n'est pas gaussien.

c) Montrer que toute combinaison linéaire (non triviale) de v.a. normales indépendantes est normale. Donner un contrexemple simple s'il n'y a pas indépendance.

d) Montrer que si le vecteur aléatoire V a ses coordonnées gaussiennes et indépendantes, alors il est gaussien (revenir à la définition 19).

e) Montrer que les coordonnées (dans la base canonique de $\mathbb{R}^d$) d'un vecteur gaussien sont indépendantes ssi la matrice de covariance est diagonale, et donc ssi elles sont non corrélées. Vérifier que c'est faux si le vecteur n'est pas gaussien (même si ses coordonnées sont gaussiennes).

Exercice n^o 6.2 Soit V un vecteur gaussien d-dimensionnel, et pour $1 \le j \le k$ soient M_j une matrice réelle de format $d_j \times d$ et $V_j := M_j V$. Montrer que V_i et V_j sont indépendants ssi $M_i K_V\, {}^t M_j = 0$. Généraliser à l'indépendance de $V_1, .., V_k$.

Exercice n^o 6.3 Soit $\{X_j \,|\, j \in \mathbb{N}^*\}$ une suite de v.a.i.i.d. $\mathcal{N}(0,1)$. Soit $\{a_j \,|\, j \in \mathbb{N}\}$ une suite de réels tels que $a_j a_{j+1} = 0$ pour tout j et tels que $\sum_j a_j^2 < \infty$. Soit $Y_n := \sum_{j=1}^n a_{n-j} X_j$, pour tout $n \in \mathbb{N}^*$.

a) Etudier la convergence en loi de la suite Y_n. b) Les v.a. Y_n et Y_{n+1} sont-elles indépendantes ? c) Etudier la convergence dans L^2 de la suite Y_n.

Exercice n^o 6.4 Soit $V := (V_0, ..V_d)$ un vecteur gaussien $(d+1)$-dimensionnel, dont les coordonnées sont $\mathcal{N}(0,1)$ et vérifient : $Cov(V_0, V_j) = p$ pour $1 \le j \le d$ et $Cov(V_i, V_j) = p^2$ pour $1 \le i \ne j \le d$, p étant un paramètre. Posons $W_j := (1-p^2)^{-1/2}(V_j - pV_0)$ pour $1 \le j \le d$. Déterminer successivement les lois de : $(V_0, W_1, .., W_d)$; $S := \sum_{j=1}^d V_j$; S/V_0.

Exercice n^o 6.5 a) Calculer les fonctions génératrices des lois usuelles : $\mathcal{B}(n,p)$, $\mathcal{G}(p)$, $\mathcal{P}(\lambda)$.

b) Calculer les transformées de Fourier des lois usuelles : $\mathcal{U}([m,n])$, $\mathcal{B}(n,p)$, $\mathcal{G}(p)$, $\mathcal{P}(\lambda)$, $\mathcal{U}([a,b])$, $\mathcal{E}(\lambda)$.

c) Retrouver à partir de là les valeurs des espérances et variances de ces différentes lois.

Exercice n^o 6.6 Étude de la transmission du nom "Chenin", porté à la génération 0 par un unique homme (= humain mâle). Notons Z_n le nombre d'hommes s'appelant "Chenin" à la génération $n \in \mathbb{N}$, F la fonction génératrice de Z_1 ($F(s) := \mathbb{E}(s^{Z_1})$), p_k la probabilité (supposée fixe) qu'un homme ait k enfants mâles, et posons $m := \sum_{\{k\in\mathbb{N}\}} k\, p_k$.
Supposons l'indépendance entre les descendances des différents hommes.
a) Calculer par récurrence la fonction G_n, génératrice de Z_n, en fonction de F.
b) Vérifier que F est monotone et convexe sur $[0,1]$.
c) En déduire en fonction de m le comportement asymptotique de $\alpha_n := \mathbb{P}(Z_n = 0)$. Interprétation ?

Exercice n^o 6.7 Un joueur va au casino avec une fortune initiale $a \in \mathbb{N}^*$. À chaque partie, il gagne 1 avec probabilité p et perd 1 avec probabilité $q = 1-p$. Les parties sont supposées indépendantes.

1) Fixons un entier $b > a$, et notons $P_b(a)$ la probabilité qu'a le joueur d'atteindre la fortune b avant d'être ruiné.

a) Montrer que $P_b(a) = p\,P_b(a+1) + q\,P_b(a-1)$. b) Déduire la valeur de $P_b(a)$.

2) Autorisons le joueur à s'endetter, notons T le premier instant où sa fortune vaut $a+1$, puis, pour $n \in \mathbb{N}$, $g_n := \mathbb{P}(T = n)$, et enfin g la fonction génératrice de T.

a) Montrer que $g_{n+2} = q\,(g_1 g_n + .. + g_n g_1)$. b) Déduire que $g(s) - ps = qs\, g^2(s)$.
c) Calculer g, $\mathbb{P}(T < \infty)$, et $\mathbb{E}(T)$.

7 Convergences des variables aléatoires

Soient X et X_n, $n \in \mathbb{N}$, des v.a. à valeurs dans $\mathbb{R}^d$, définies sur un même espace de probabilité. Notons $\|X\| := [\,\omega \mapsto \|X(\omega)\|\,]$ une norme de X dans $\mathbb{R}^d$.

Définition 23 *On dit que la suite $\{X_n\,,\, n \in \mathbb{N}\}$ converge vers X*

i) dans L^p (où $p \in [1,\infty]$) lorsque $\mathbb{E}(\|X_n - X\|^p) \to 0$; la convergence dans L^1 est aussi appelée convergence en moyenne, et la convergence dans L^2 convergence en moyenne quadratique ;

ii) presque sûrement lorsque $\mathbb{P}(X_n$ converge vers $X) = 1$;

iii) en probabilité lorsque pour tout $\varepsilon > 0$ $\mathbb{P}(\|X_n - X\| > \varepsilon)$ tend vers 0.

Théorème 2 *1) Convergence monotone : Si (V_n) est une suite p.s. croissante de v.a.r. ≥ 0, alors $\lim_n \uparrow \mathbb{E}(V_n) = \mathbb{E}(\lim_n \uparrow V_n)$.*

2) Convergence dominée (de Lebesgue) : *Si (X_n) est une suite de v.a. qui converge p.s. vers une v.a. X, et si elle est dominée par une v.a.r. fixe intégrable Y : $\|X_n\| \leq Y \in L^1$, alors $X \in L^1$ et X_n converge vers X dans L^1.*

Exercice n^o 7.1 a) Montrer que la convergence presque sûre et la convergence dans L^p entraînent (chacune) la convergence en probabilité.
b) Montrer que la convergence dans L^p entraîne la convergence dans L^q si $p > q$.
c) Supposons que X_n converge en probabilité vers X et que Y_n converge en probabilité vers Y. Soient a et b réels. Montrer que $aX_n + bY_n$ et $X_n \cdot Y_n$ convergent en probabilité.

Exercice n^o 7.2 Trouver

a) Une suite de v.a.r. qui converge dans L^q mais pas dans L^p (pour $p > q$ fixés).

b) Une suite de v.a.r. qui converge dans L^p mais pas presque sûrement (pour $p < \infty$ fixé).

c) Une suite de v.a.r. qui converge presque sûrement mais dans aucun L^p.

Exercice n^o 7.3 Soit $\{X_n \,|\, n \in \mathbb{N}^*\}$ une suite de v.a.r. indépendantes, telles que $\mathbb{P}(X_n = n) = 1/n$ et $\mathbb{P}(X_n = 0) = 1 - 1/n$ pour chaque $n \in \mathbb{N}^*$. Montrer que cette suite converge en probabilité, mais ni presque sûrement ni dans aucun L^p.

Exercice n^o 7.4 a) Montrer qu'une suite de v.a. qui converge en probabilité admet une sous-suite qui converge presque sûrement. (Utiliser le premier lemme de Borel-Cantelli, proposition 6 section 2.)
b) Montrer qu'une suite convergente en probabilité et dominée par une variable de L^p converge dans L^p.

Exercice n^o 7.5 a) Montrer que $\{X_n\,,\, n \in \mathbb{N}\}$ converge en probabilité vers X ssi $\mathbb{E}(\|X_n - X\| \wedge 1)$ tend vers 0.

b) Vérifier que $(X, Y) \longmapsto \mathbb{E}(\|X - Y\| \wedge 1)$ est une distance sur $L^0(\Omega, \mathcal{T}, \mathbb{P}; \mathbb{R}^d)$.

Définition 24 *On dit que la suite $\{X_n\,,\, n \in \mathbb{N}\}$ de variables aléatoires à valeurs dans $\mathbb{R}^d$ converge en loi vers X (à valeurs dans $\mathbb{R}^d$) lorsque la loi de X_n converge vaguement (ou étroitement) vers celle de X. Id est : lorsque $\mathbb{E}(f \circ X_n)$ tend vers $\mathbb{E}(f \circ X)$, pour toute fonction f continue à support compact (ou continue bornée) de $\mathbb{R}^d$ dans $\mathbb{R}$.*

Proposition 14 *La convergence en probabilité entraîne la convergence en loi.*

Exercice n^o 7.6 a) Prouver que lorsque X est p.s. constante, les convergences vers X en probabilité et en loi sont équivalentes.
b) Montrer que ces deux convergences sont généralement non équivalentes.

Exercice n^o 7.7 Montrer que si X_n converge en loi vers X et si Y_n converge en probabilité vers 0, alors $X_n \cdot Y_n$ converge en probabilité vers 0. (Distinguer entre $\{|Y_n| > A\}$ et $\{|Y_n| \leq A\}$).

Remarque 7 *a) Si les lois de X_n et de X ont un même support dénombrable discret S (en général $S \subset \mathbb{Z}^d$), alors la convergence en loi de X_n vers X équivaut à :*

$$\mathbb{P}(X_n = s) \text{ tend vers } \mathbb{P}(X = s) \quad \text{pour tout } s \in S\,.$$

b) Si X_n converge en loi vers X et si g est continue, alors $g \circ X_n$ converge en loi vers $g \circ X$.

Proposition 15 *La convergence en loi équivaut à la convergence simple des transformées de Fourier.*

Proposition 16 *La convergence en loi d'une suite de v.a.r. $\{X_n \,|\, n \in \mathbb{N}^*\}$ vers une v.a.r. X équivaut à la convergence simple de F_n vers F en chaque point de continuité de F ; F_n désignant la fonction de répartition de X_n, et F celle de X.*

Exemples a) La loi $B(n, p_n)$ converge vers la loi $\mathcal{P}(\lambda)$ si $np_n \to \lambda > 0$.

b) La loi $H(N, n, p_N)$ converge vers la loi $B(n, p)$ si $N \to \infty$ et $p_N \to p \in]0,1[$.

c) Soient B une variable de Bernoulli ($\mathbb{P}(B = 1) = \mathbb{P}(B = -1) = 1/2$), et pour tout $n \in \mathbb{N}$: $X_n = B = X'_{2n}$ et $X'_{2n+1} = -B$. Alors les 2 suites X et X' convergent en loi, mais pas la suite $X + X'$.

Exercice n^o 7.8 Notons $\{X_n \,|\, n \in \mathbb{N}^*\}$ une suite de v.a.r. indépendantes gaussiennes standard. Etudier la convergence en loi de $\frac{nX_1+(n-1)X_2+..+X_n}{n\sqrt{n}}$.

Exercice n^o 7.9 Soit $\{X_n \,|\, n \in \mathbb{N}\}$ une suite de v.a.r. de fonction de répartition commune F, telle que $\lim\limits_{x\to\infty} x\,(1 - F(x)) = \lim\limits_{x\to\infty} x\,F(-x) = 0$. Posons $M_n := \max\{X_1, .., X_n\}$, et $m_n := \min\{X_1, .., X_n\}$. Montrer que $\frac{M_n}{n}$ et $\frac{m_n}{n}$ convergent en probabilité vers 0.

8 Loi des Grands Nombres

Théorème 3 *Soit $\{X_n \,|\, n \in \mathbb{N}^*\}$ une suite de v.a. indépendantes et de même loi intégrable. Alors $\frac{X_1+..+X_n}{n}$ converge p.s. vers $\mathbb{E}(X_1)$.*

Remarque 8 La réciproque suivante de la loi des grands nombres est vraie :

Si $\{X_n \,|\, n \in \mathbb{N}^*\}$ est une suite de v.a.r indépendantes et de même loi telle que la suite des moyennes de Césaro $\frac{X_1+..+X_n}{n}$ converge p.s., alors X_1 est intégrable.

Preuve Remarquons qu'alors $\frac{X_n}{n} = \frac{X_1+..+X_n}{n} - \frac{X_1+..+X_{n-1}}{n-1} \times \frac{n-1}{n}$ doit converger p.s. vers 0. Donc $\mathbb{E}(|X_1|) \leq \sum_n \mathbb{P}(|X_n| > n)$ est finie d'après le deuxième lemme de Borel-Cantelli. $\diamond$

Remarque 9 1) Le théorème 3 ci-dessus constitue la loi forte des grands nombres. On appelle loi faible des grands nombres le résultat (strictement plus faible) qui énonce la convergence en probabilité des moyennes, au lieu de leur convergence presque sûre.

2) On montre que la convergence dans la loi forte des grands nombres a lieu également dans L^1.

Exercice n^o 8.1 Prouver la formulation suivante de la loi faible des grands nombres :

Si les v.a.r. X_n sont non-corrélées 2 à 2 et ont la même loi admettant un second moment, alors leurs moyennes de Césaro convergent dans L^2 (et donc en probabilité) vers $I\!\!E(X_1)$.

Exercice n^o 8.2 Soit $\{X_n \,|\, n \in I\!\!N^*\}$ une suite de v.a.r.i.i.d. telles que $I\!\!E(X_1^+) = \infty > I\!\!E(X_1^-)$. Montrer que p.s. $\frac{X_1+..+X_n}{n} \longrightarrow +\infty$. (Utiliser $X_n^k := X_n \wedge k$).

Exercice n^o 8.3 a) Soient $p \in [0,1]$ et f une fonction continue de $[0,1]$ dans $I\!\!R$. Etudier la convergence de $n \longmapsto \sum_{k=0}^n C_n^k \, p^k (1-p)^{n-k} f(k/n)$.

b) Soient $\lambda > 0$ et g une fonction continue bornée de $I\!\!R_+$ dans $I\!\!R$. Etudier la convergence de $n \longmapsto e^{-\lambda n} \sum_{k\geq 0} \frac{(\lambda n)^k}{k!} g(k/n)$; puis de $n \longmapsto e^{-\lambda n} \sum_{k=0}^{[\lambda n]} \frac{(\lambda n)^k}{k!}$ (utiliser le théorème suivant ; interpréter en terme de médiane).

9 Théorème Central Limite

Théorème 4 *Soit $\{X_n \,|\, n \in I\!\!N^*\}$ une suite de v.a. à valeurs dans $I\!\!R^d$, indépendantes, et de même loi admettant un second moment. Alors $\frac{X_1+..+X_n - n I\!\!E(X_1)}{\sqrt{n}}$ converge en loi, vers la loi gaussienne $\mathcal{N}(0, K_{X_1})$, où $K_{X_1} := I\!\!E(X_1 \, {}^t X_1) - I\!\!E(X_1)\, {}^t I\!\!E(X_1)$ désigne la matrice de covariance de X_1.*

Preuve L'énoncé étant clairement invariant par une translation sur la suite X, il suffit de considérer le cas où X_1 est centrée. Notons $\phi(v) := I\!\!E(e^{i\, v \cdot X_1})$ la transformée de Fourier de X_1, qui est de classe C^2, puisque X_1 est supposée de carré intégrable. La formule de Taylor-Young à l'origine s'écrit alors : $\phi(v) = 1 - \frac{1}{2}\, {}^t v K(X_1) v + o(\|v\|^2)$.
Par indépendance, la transformée de Fourier de $\frac{X_1+..+X_n}{\sqrt{n}}$ est égale à $\phi(v/\sqrt{n}\,)^n$, qui vaut $\left[1 - \frac{{}^t v K(X_1) v + n\, o(\|v\|^2/n)}{2n}\right]^n$, et donc qui converge vers $\exp\left[-\frac{1}{2}\, {}^t v K(X_1) v\right]$, du fait de la convergence de $(1+\frac{z}{n})^n$ vers e^z uniformément sur les compacts de $\mathbb{C}$.
En effet, si $|z| \leq A$, nous avons :

$$\left| e^z - (1+\tfrac{z}{n})^n \right| \leq \sum_{k=0}^n \left(\tfrac{1}{k!} - \tfrac{C_n^k}{n^k}\right) |z|^k + \sum_{k=n+1}^\infty \tfrac{|z|^k}{k!} \leq \sum_{k=0}^n \left(\tfrac{1}{k!} - \tfrac{C_n^k}{n^k}\right) A^k + \sum_{k=n+1}^\infty \tfrac{A^k}{k!}$$
$$= e^A - (1+\tfrac{A}{n})^n = e^A - e^{n \log(1+\frac{A}{n})} \longrightarrow 0\,.$$

Finalement, nous venons de montrer la convergence simple de la transformée de Fourier de $\frac{X_1+..+X_n}{\sqrt{n}}$ vers celle de $\mathcal{N}(0, K_{X_1})$, ce qui suffit d'après la proposition 15. $\diamond$

Exercice n^o 9.1 (Surlocation) Une agence de voyage dispose de 160 places à louer pour une destination donnée. Elle sait que les locations sont honorées par ses clients avec une probabilité fixe p. Elle vend $N = 160\alpha > 160$ places. Pour quelles valeurs de α la probabilité de ne pas louer trop de places vaut-elle $0,95$, $0,975$?

Exercice n^o 9.2 Le prix S_n d'une action au jour n est modélisé ainsi : $S_0 = s > 0$ est fixe, et $S_{n+1} = (1 + r + \sigma\varepsilon_{n+1})\, S_n$, où $r > 0$ est un taux fixe, $\sigma \in]0, 1 + r[$ est une volatilité fixe, et $\{\varepsilon_n \,|\, n \in I\!N\}$ est une suite i.i.d. de loi de Bernoulli $\mathcal{B}(\pm 1, 1/2)$.

a) Etudier le comportement des suites $(\log S_n)/n$ et S_n .
b) Etudier le comportement de la suite $(\log S_n)/\sqrt{n}$.
c) Etudier le comportement de la suite $[(1 + r)^2 - \sigma^2]^{(-\sqrt{n}/2)} \times S_n^{[1/\sqrt{n}]}$.

Exercice n^o 9.3 Notons $\{X_n, Y_n, Z_n \,|\, n \in I\!N^*\}$ une famille de v.a.i.i.d. de loi commune $\mathcal{B}(\pm 1, 1/2)$. Posons pour tout $n \in I\!N$:
$R_n := \sum_{k=1}^n X_k Y_k$, $S_n := \sum_{k=1}^n Y_k Z_k$, $T_n := \sum_{k=1}^n Z_k X_k$, puis $V_n := (R_n, S_n, T_n)$.

a) R_n, S_n, T_n sont-elles indépendantes 2 à 2 ? indépendantes ?
b) Quelle est la loi de R_n ? c) Calculer la transformée de Fourier de V_n .
d) Etudier la convergence de V_n/n , puis de $V_n/\sqrt{n}$.

Exercice n^o 9.4 3016 mathématiciens sont invités à un colloque ; en moyenne un sur 4 répondra favorablement, et les réponses sont supposées indépendantes les unes des autres. Combien les organisateurs doivent-ils prévoir de places, afin que la probabilité de ne pas en manquer soit $\geq 0,99$?

Exercice n^o 9.5 Une compagnie assure 10000 clients sur la vie, qui payent chacun une prime annuelle de A euros. On estime que chaque client a une probabilité de décès au cours d'une année égale à 6/1000, indépendamment les uns des autres. La prime de décès est de B euros.

a) Quelle est la loi du nombre annuel des décès ? Comment peut-on l'approcher ?

b) Si $B = 1000$, pour quels A la compagnie a-t-elle une probabilité $< 1/100$ d'être en déficit ? c) Si $A = 15$, pour quels B la compagnie a-t-elle une probabilité $> 0,7$ de faire un bénéfice annuel > 50000 euros ?

Exercice n^o 9.6 Soit $\{X_n \,|\, n \in I\!N\}$ une suite de v.a.i.i.d. de loi commune $\mathcal{B}(p)$. Posons $Y_n := X_n X_{n+1}$ et $S_n := (Y_1 + .. + Y_n)/n$, pour tout $n \in I\!N$.
Calculer la loi et l'espérance de Y_n , puis la covariance de Y_n et de Y_{n+k} , l'espérance de S_n^2 , et enfin montrer que S_n converge en probabilité vers p^2.

Exercice n^o 9.7 Soit $\{X_n \,|\, n \in I\!N^*\}$ une suite de v.a. indépendantes de loi de Poisson de paramètre 1.
a) Montrer que $I\!E\Big[\Big(\frac{X_1+..+X_n-n}{\sqrt{n}}\Big)^-\Big] = \frac{n^{(n+\frac{1}{2})}e^{-n}}{n!}$.
b) Montrer que $\Big(\frac{X_1+..+X_n-n}{\sqrt{n}}\Big)^-$ converge en loi vers N^-, partie négative d'une gaussienne centrée réduite.
c) Montrer que $\frac{X_1+..+X_n-n}{\sqrt{n}}$ est bornée dans L^2, puis que $I\!E\Big[\Big(\frac{X_1+..+X_n-n}{\sqrt{n}}\Big)^-\Big] \longrightarrow I\!E(N^-)$.

d) Retrouver la formule de Stirling.

10 Exemple des marches aléatoires

Définition 25 *Soit $\{X_n \,|\, n \in I\!N^*\}$ une suite de v.a. à valeurs dans $\mathbb{Z}^d$, indépendantes, et de même loi μ. On appelle marche aléatoire (sur $\mathbb{Z}^d$) le processus $\{S_n \,|\, n \in I\!N\}$ à valeurs dans $\mathbb{Z}^d$ défini par : $S_0 := 0$, et pour tout $n \in I\!N^*$: $S_n := X_1 + .. + X_n$.*
La marche est dite simple lorsque $\mu(\{e_j\}) = \mu(\{-e_j\}) = 1/(2d)$, où $(e_1, .., e_d)$ désigne la base canonique de $\mathbb{Z}^d$.

Définition 26 *Soit pour tout $n \in I\!N^*$: $\mathcal{F}_n := \sigma(X_1, .., X_n) = \sigma(S_1, .., S_n)$ la tribu représentant l'information connue au temps n. On note aussi $\mathcal{F}_0$ la tribu triviale, et $\mathcal{F}_\infty$ la tribu engendrée par $\cup_{\{n \in I\!N\}} \mathcal{F}_n$.*
On appelle temps d'arrêt toute v.a. N à valeurs dans $\overline{I\!N}$ telle que $\{N \le n\} \in \mathcal{F}_n$ pour tout $n \in I\!N$. On pose alors : $\mathcal{F}_N := \{A \in \mathcal{F}_\infty \,|\, (\forall n \in I\!N)\ A \cap \{N \le n\} \in \mathcal{F}_n\}$.

L'exemple fondamental de temps d'arrêt est le temps d'atteinte par le processus d'une certaine partie E de Z^d : $\min\{n \in I\!N \,|\, S_n \in E\}$ est bien un temps d'arrêt.
Un autre exemple, trivial, est fourni par les temps constants.

Exercice n^o 10.1 Montrer que N est un temps d'arrêt ssi $\{N = n\} \in \mathcal{F}_n$ pour tout $n \in I\!N$, que $\mathcal{F}_N := \{A \in \mathcal{F}_\infty \,|\, (\forall n \in I\!N)\ A \cap \{N = n\} \in \mathcal{F}_n\}$, et que $\mathcal{F}_N$ est une tribu.

Exercice n^o 10.2 Soient N et N' deux temps d'arrêt.
a) Vérifier que $\min\{N, N'\}$ et $\max\{N, N'\}$ sont aussi deux temps d'arrêt.

b) Montrer que les événements $\{N < N'\}, \{N \le N'\}, \{N = N'\}$ appartiennent à $\mathcal{F}_{\min\{N,N'\}}$, puis que $\mathcal{F}_{\min\{N,N'\}} = \mathcal{F}_N \cap \mathcal{F}_{N'}$.

Exercice n^o 10.3 Soit N un temps d'arrêt. Supposons que $I\!E(\|X_1\|)$ et $I\!E(N)$ sont finies. a) Montrer que $I\!E(S_N) = I\!E(X_1) \times I\!E(N)$. (Ecrire $S_N = \sum_j X_j 1_{\{N<j\}^c}$).

b) Déduire (pour la marche simple, $d = 1$) que $\min\{n | S_n = 1\}$ et T ne sont pas intégrables.

Lemme 1 *Soit N un temps d'arrêt. Conditionnellement à $\{N < \infty\}$,*
la suite $\{X_{N+n} \,|\, n \in I\!N^*\}$ est indépendante de $\mathcal{F}_N$, et a la même loi que $\{X_n \,|\, n \in I\!N^*\}$.

Preuve Fixons $A \in \mathcal{F}_N$, $k \in I\!N^*$, et $v_1, .., v_k \in \mathbb{Z}^d$. Nous avons

$$I\!P(A, N < \infty, X_{N+1} = v_1, .., X_{N+k} = v_k) = \sum_{n \in I\!N} I\!P(A, N = n, X_{n+1} = v_1, .., X_{n+k} = v_k)$$

$$= \sum_{n \in I\!N} I\!P(A\,,\, N = n) I\!P(X_{n+1} = v_1\,, .., X_{n+k} = v_k) = I\!P(A\,,\, N < \infty) \prod_{j=1}^{k} \mu(v_j)\,. \ \diamond$$

Remarque 10 Quitte à prendre l'image de Ω par X, on peut toujours se placer sur l'espace canonique, i.e. considérer que $\Omega = (\mathbb{Z}^d)^{I\!N^*}$, $\mathcal{F} = \mathcal{F}_\infty$, de sorte que X devient la suite des coordonnées : $X_n(\omega) = \omega_n$, $\forall n \in I\!N$. $I\!P$ demeure la loi de la marche S, déduite de $(X, \mu^{\otimes I\!N^*})$.

10.1 Temps de retour en 0

Soient $T_0 := 0$, $T_1 = T := \min\{n \in \mathbb{N}^* \,|\, S_n = 0\}$ le premier temps de retour en 0 (qui vaut ∞ s'il n'y en a pas), puis pour $k \in \mathbb{N}^*$: $T_k := \min\{n > T_{k-1} \,|\, S_n = 0\}$ le k-ième temps de retour en 0.

Exercice n^o 10.4 a) Montrer que $\mathbb{P}(T_k < \infty) = \mathbb{P}(T < \infty)^k$. (Utiliser le lemme 1).

b) Pour tout $v \in \mathbb{Z}^d, v \neq 0$, notons $T_k(v)$ le k-ième temps de passage en v. Montrer que

$$\mathbb{P}(T_k(v) < \infty) = \mathbb{P}(T_1(v) < \infty) \times \mathbb{P}(T < \infty)^{k-1}.$$

Proposition 17 *a) Les énoncés suivant sont équivalents :*

(i) $\mathbb{P}(T < \infty) = 1$. *(ii)* $\mathbb{P}(\limsup_n \{S_n = 0\}) = 1$. *(iii)* $\sum_n \mathbb{P}(S_n = 0) = \infty$.

b) Si $\mathbb{P}(T < \infty) < 1$, *alors* $\mathbb{P}(\limsup_n \{S_n = v\}) = 0$ *pour tout* $v \in \mathbb{Z}^d$.

Preuve a) L'exercice 10.4.b montre que (i) entraîne $\mathbb{P}(T_n < \infty\ (\forall n)) = 1$, c'est-à-dire (ii). Le premier lemme de Borel-Cantelli montre que $(ii) \Rightarrow (iii)$. Ensuite nous avons $\sum_n 1_{\{S_n=0\}} = \sum_n 1_{\{T_n<\infty\}}$ presque sûrement, et donc en intégrant : $\sum_{n\in\mathbb{N}} \mathbb{P}(\{S_n = 0\}) = \sum_{n\in\mathbb{N}} \mathbb{P}(\{T_n < \infty\}) = \sum_n \mathbb{P}(\{T < \infty\})^n = [1 - \mathbb{P}(\{T < \infty\})]^{-1}$ d'après l'exercice 10.4.b à nouveau, d'où $(iii) \Rightarrow (i)$.

b) $\mathbb{E}[\sum_n 1_{\{S_n=v\}}] = \sum_n \mathbb{P}(\{T_n(v) < \infty\}) = \mathbb{P}(T_1(v) < \infty)/\mathbb{P}(T = \infty) < \infty$ d'après l'exercice 10.4.c, et donc $\mathbb{P}(\limsup_n \{S_n = v\}) = \mathbb{P}[\sum_n 1_{\{S_n=v\}} = \infty] = 0$. $\diamond$

Proposition 18 *Soient* $\mathcal{R} := \{v \in \mathbb{Z}^d \,|\, \limsup_n \{S_n = v\} = \Omega \ p.s.\}$ *l'ensemble des points presque sûrement infiniment visités par la marche aléatoire, et*
$\mathcal{E} := \{v \in \mathbb{Z}^d \,|\, (\exists n \in \mathbb{N})\ \mathbb{P}(S_n = v) > 0\}$ *l'ensemble des points pouvant être visités par la marche aléatoire.*
Alors soit $\mathcal{R}$ *est vide, soit c'est un sous-groupe de* $\mathbb{Z}^d$ *et il est égal à* $\mathcal{E}$.

Preuve Soient $v \in \mathcal{E}$, et $w \notin v+\mathcal{R}$. Fixons $k \in \mathbb{N}$ tel que $\mathbb{P}(S_k = v) > 0$, et $m \in \mathbb{N}^*$ tel que $\mathbb{P}((\forall n \geq m)\ S_n \neq w - v) > 0$. Alors

$$\mathbb{P}((\forall n \geq m + k)\ S_n \neq w) \geq \mathbb{P}((\forall n \geq m + k)\ S_n - S_k \neq w - v\,,\ S_k = v)$$

$$= \mathbb{P}((\forall n \geq m+k)\ S_n-S_k \neq w-v)\,\mathbb{P}(S_k = v) = \mathbb{P}((\forall n \geq m)\ S_n \neq w-v)\,\mathbb{P}(S_k = v) > 0.$$

Donc nous venons de prouver que si $w - v \notin \mathcal{R}$ et $v \in \mathcal{E}$, alors $w \notin \mathcal{R}$. Autrement dit : $v \in \mathcal{E}$ et $w \in \mathcal{R}$ impliquent $w - v \in \mathcal{R}$.

Remarquons que $\mathcal{R} \subset \mathcal{E}$. Donc prenant v et w dans $\mathcal{R}$, nous voyons que $\mathcal{R}$ est bien un sous-groupe, s'il est non vide. Enfin, dans ce cas toujours, prenant $w = 0$, nous obtenons $\mathcal{E} \subset \mathcal{R}$. $\diamond$

10.2 Récurrence des marches aléatoires simples

Dans le cas de la marche simple sur $\mathbb{Z}^d$ (voir la définition 25), il est clair que $\mathcal{E} = \mathbb{Z}^d$. Les propositions 17 et 18 nous assurent donc que soit la marche simple repasse indéfiniment par chaque point de $\mathbb{Z}^d$, soit chaque point n'est visité qu'un nombre fini de fois par elle, et donc elle tend vers l'infini, tout ceci presque sûrement. Dans le premier cas, on dit que la marche est <u>récurrente</u>, et dans le second, qu'elle est <u>transitoire</u> (ou <u>transiente</u>).

Le célèbre résultat suivant peut se traduire en langage courant par la phrase : un homme ivre finira par retrouver sa demeure, alors qu'un oiseau ivre peut se perdre définitivement.

Théorème 5 *La marche aléatoire simple sur $\mathbb{Z}^d$ est récurrente pour $d=1$ ou $d=2$, mais est transitoire pour $d \geq 3$.*

<u>Preuve</u> La marche est récurrente ssi 0 est récurrent, i.e. ssi la marche repasse presque sûrement indéfiniment par 0, et donc, d'après la proposition 17, ssi $\sum_n \mathbb{P}(S_n = 0) = \infty$. Remarquons que $P(S_{2n+1} = 0) = 0$ pour tout n. Nous avons donc récurrence ssi $\sum_n \mathbb{P}(S_{2n} = 0) = \infty$.

Pour $d = 1$, nous avons aussitôt $\mathbb{P}(S_{2n} = 0) = 2^{-2n} C_{2n}^n \sim (\pi n)^{-1/2}$, puisqu'il faut autant de pas vers la gauche que vers la droite, et en utilisant la formule de Stirling.

Pour $d = 2$, distinguant selon le nombre m de déplacements de la marche dans le sens horizontal, un dénombrement simple donne : $\mathbb{P}(S_{2n} = 0) = 4^{-2n} \sum_{m=0}^n \frac{(2n)!}{m!m!(n-m)!(n-m)!}$, d'où $\mathbb{P}(S_{2n} = 0) = 2^{-4n} C_{2n}^n \sum_{m=0}^n C_n^m C_n^{n-m} = \left(2^{-2n} C_{2n}^n\right)^2 \sim (\pi n)^{-1}$.

Pour $d = 3$, le même genre de considérations donne :

$$\mathbb{P}(S_{2n} = 0) = 6^{-2n} \sum_{0 \leq l+m \leq n} \frac{(2n)!}{[l!m!(n-l-m)!]^2} = 2^{-2n} C_{2n}^n \sum_{0 \leq l+m \leq n} \Big(3^{-n} \frac{n!}{l!m!(n-l-m)!}\Big)^2$$

$$\leq 2^{-2n} C_{2n}^n \max_{0 \leq l+m \leq n} \Big\{3^{-n} \frac{n!}{[l!m!(n-l-m)!]^2}\Big\} \sum_{0 \leq l+m \leq n} \Big(3^{-n} \frac{n!}{l!m!(n-l-m)!}\Big)$$

$$= 2^{-2n} C_{2n}^n \max_{n_1+n_2+n_3=n} \Big\{3^{-n} \frac{n!}{n_1!n_2!n_3!}\Big\} = \frac{2^{-2n} C_{2n}^n \times 3^{-n} \times n!}{\min\Big\{n_1!n_2!n_3! \,\Big|\, n_1+n_2+n_3 = n, |n_i - n_j| \leq 1\Big\}} ,$$

puisque si par exemple $n_1 - n_2 \geq 2$ alors $n_1!n_2!n_3! > (n_1 - 1)!(n_2 + 1)!n_3!$; cela impose $n_1, n_2, n_3 \geq (n-2)/3$, d'où par la formule de Stirling la conclusion pour $d = 3$:

$$\mathbb{P}(S_{2n} = 0) = \mathcal{O}\Big(\frac{n^{-1/2} \times 3^{-n} \times (n/e)^n \sqrt{n}}{(n_1/e)^{n_1} (n_2/e)^{n_2} (n_3/e)^{n_3} \sqrt{n_1 n_2 n_3}}\Big) = \mathcal{O}\Big(\frac{3^{-n} \times n^n}{(\frac{n-2}{3})^{n+3/2}}\Big) = \mathcal{O}(n^{-3/2}) .$$

Finalement, si $d > 3$, considérons la projection S'_n de la marche S_n sur les trois premières coordonnées de $\mathbb{Z}^d$, puis $T_0 := 0$, et pour $n \in \mathbb{N}^*$:
$T_n := \min\{m > T_{n-1} \,|\, S'_m \neq S'_{T_{n-1}}\}$. Alors T_n est p.s. fini, S'_{T_n} est la marche simple sur $\mathbb{Z}^3$, et donc ne passe plus par 0 après un certain temps aléatoire presque sûrement fini N .

Comme S'_n est constante entre deux instants T_{n-1} et T_n, nous avons a fortiori $S'_n \neq 0$ (et donc $S_n \neq 0$) pour $n > T_N$, d'où la transience annoncée. $\diamond$

Exercice n^o 10.5 Marche aléatoire simple sur un arbre homogène.

Soit $\mathcal{A}$ un arbre : c'est un graphe (non vide) non orienté sans boucle, de sorte que 2 points (ou sommets) quelconques de l'arbre sont reliés par un unique chemin injectif ; le nombre d'arêtes que comporte ce chemin définit la distance entre ces deux points ; elle fait de l'arbre un espace métrique. Supposons $\mathcal{A}$ homogène : chaque point admet exactement d voisins (i.e. points situés à distance 1), et connexe.

À quoi $\mathcal{A}$ est-il isométrique lorsque $d = 1$ ou 2 ? Fixons $d \geq 3$, et un point $O \in \mathcal{A}$. La marche simple S sur $\mathcal{A}$ part de $S_0 := O$, et pour tout $n \in \mathbb{N}^*$, la loi de S_n sachant S_{n-1} est uniforme sur l'ensemble des d voisins de S_{n-1}. Notons δ_n la distance entre S_0 et S_n.

a) Montrer qu'il existe une suite b de v.a.i.i.d., de loi à déterminer, telle que pour tout $n \in \mathbb{N}$: $\delta_{n+1} - \delta_n = b_n \, 1_{\{\delta_n > 0\}} + 1_{\{\delta_n = 0\}}$. (Définir b_n sur $\{\delta_n = 0\}$ par b'_n indépendante ayant la bonne loi, puis vérifier que b convient en calculant $\mathbb{P}(b_n = 1 | b_{n-1}, .., b_0)$.)

b) Montrer que la marche S est transitoire, et qu'elle admet une vitesse de fuite, limite presque sûre de δ_n/n, à déterminer. (Commencer par minorer δ_n.)

11 Processus de Poisson

Définition 27 *On appelle processus de comptage tout processus* $(N_t \,|\, t \in \mathbb{R}_+)$ *donné par* $N_t := Card\,(\mathcal{E} \cap [0,t])$ *, où la donnée* $\mathcal{E}$ *est une partie discrète aléatoire de* $\mathbb{R}_+$ *.*

Le processus de comptage N *est dit à accroissements indépendants lorsque pour tout* n *dans* $\mathbb{N}^*$ *et tous* $s_1 < t_1 \leq .. \leq s_n < t_n$ *les variables* $(N_{t_1} - N_{s_1})$, .., $(N_{t_n} - N_{s_n})$ *sont indépendantes. Cela signifie l'indépendance des nombres de points de* $\mathcal{E}$ *situés dans des intervalles disjoints deux à deux.*

Remarques 1) Un processus de comptage est nécessairement croissant, à valeurs dans $\overline{\mathbb{N}}$, et presque sûrement continu à droite en tout temps (y compris aléatoire) :

$$\mathbb{P}(\inf N_{S_n} \neq N_S) \leq \mathbb{P}(N_{S_n} - N_S \geq 1) = \mathbb{P}(\mathcal{E} \cap]S, S_n] \neq \emptyset) \searrow 0 \text{ si } S_n \searrow S \text{ p.s..}$$

2) Les points de $\mathcal{E}$ sont généralement des instants successifs de survenue d'événements (d'un type donné a priori : arrivées dans une file, passages de bus, désintégrations d'atomes, etc..), qu'on observe au cours du temps t. Pour tous $0 \leq s < t$, la variable aléatoire $N_t - N_s$ décompte le nombre de ces événements survenus dans l'intervalle de temps $]s,t]$.

Définition 28 *On appelle processus de Poisson (homogène) d'intensité* λ *(quelconque dans* $\mathbb{R}_+^*$*) tout processus de comptage* N *à accroissements indépendants nul en 0 et tel que pour tous* $s,t \in \mathbb{R}_+$ *la loi de* $(N_{s+t} - N_s)$ *est* $\mathcal{P}(\lambda t)$ *, la loi de Poisson de paramètre* λt *.*

Fixons dorénavant un processus de Poisson N, de filtration $(\mathcal{F}_t)$ et d'intensité λ.

Exercice n^o 11.1 a) Un processus de Poisson est à accroissements stationnaires : la loi de tout vecteur $(N_{a+t_1} - N_{a+s_1}, .., N_{a+t_n} - N_{a+s_n})$ est indépendante de $a \in I\!R_+$.

b) Un processus de Poisson est p.s. fini à tout instant : $I\!P\Big(\bigcap_t\{N_t < \infty\}\Big) = 1$: les événements qu'il décompte ne peuvent pas s'accumuler en un instant fini.

c) En temps t petit, on a : $I\!P(N_t = 1) = \lambda t + o(t)$ et $I\!P(N_t \geq 1) = \lambda t + o(t)$.

d) On a presque sûrement : $\lim\limits_{t\to\infty} N_t/t = \lambda = I\!E(N_t/t)$.

Lemme 2 *On a presque sûrement* $N_{S_1} = 1$ *(où* $S_1 := \inf\{t > 0 \,|\, N_t > 0\}$*).*

Corollaire 7 *Les points de* $\mathcal{E}$ *s'ordonnent en une suite strictement croissante de variables aléatoires tendant vers* $+\infty$ *:* $S_0 := 0 < S_1 < .. < S_n < \ldots$ *Pour tout* $n \in I\!N$ S_n *est le* n*-ième point de* $\mathcal{E}$ *.*

Preuve Définissons par récurrence $S_0 := 0$ et $S_{n+1} := \inf\{t > 0 \,|\, N_t > N_{S_n}\}$. Le lemme 2 permet de voir que ce sont des temps p.s. finis, et que $I\!P(N_{S_{n+1}} - N_{S_n} = 1) = 1$. Ceci montre que les points de $\mathcal{E}$ surviennent bien lors d'une suite strictement croissante, et que S_n est le n-ième point de $\mathcal{E}$. L'exercice (11.1.b) montre que cette suite tend vers l'infini presque sûrement. $\diamond$

Théorème 6 *Pour tout* $n \in I\!N^*$ *, notons* $T_n := S_n - S_{n-1}$ *l'intervalle de temps entre les* $(n-1)$*-ième et* n*-ième points de* $\mathcal{E}$*.*

La suite $(T_n\,,\, n \in I\!N^*)$ *est i.i.d., de loi commune la loi exponentielle de paramètre* λ*.*

Exercice 11.2 a) Montrer que pour tout $n \in I\!N^*$ la loi de S_n est la loi Gamma de paramètres n et λ, i.e. la loi de densité sur $I\!R_+$: $s \longrightarrow \frac{(\lambda s)^{n-1}}{(n-1)!} \times \lambda\, e^{-\lambda s}$.

b) Pour $t > 0$ et $f \in L^1([0,t])$, calculer $I\!E\Big(\sum_{n\in I\!N^*} f(S_n)\, 1_{\{S_n \leq t\}}\Big)$.

c) Fixons $t > 0$ et $n \in I\!N^*$. Montrer que conditionnellement par rapport à $\{N_t = n\}$, le vecteur $(S_1, .., S_n)$ admet la densité de Dirichlet : $(s_1, .., s_n) \longrightarrow \dfrac{n!}{t^n}\, 1_{\{0<s_1<..<s_n<t\}}$.

d) Montrer que (pour $t > 0$ et $n \in I\!N^*$) si $U_1, .., U_n$ sont n variables i.i.d. uniformes sur $[0,t]$, et si $(V_1, .., V_n)$ est le vecteur obtenu en ordonnant $U_1, .., U_n$ dans le sens croissant, alors la "statistique d'ordre" $(V_1, .., V_n)$ admet aussi la densité de Dirichlet ci-dessus.

e) Déduire que pour $0 < s < t$, conditionnellement par rapport à $\{N_t = n\}$, la loi de N_s est $\mathcal{B}(n, s/t)$.

Il est frappant que la réciproque du théorème 6 soit exacte, permettant par exemple de simuler un processus de Poisson.

Théorème 7 *Soit* $\{T_n\,;\, n \in \mathbb{N}^*\}$ *une suite i.i.d. de variables exponentielles de paramètre* $\lambda > 0$ *. Soit (pour tout* $t \in \mathbb{R}_+$*)* $N_t := \sum\limits_{n \in \mathbb{N}^*} 1_{\{T_1+..+T_n \le t\}}$ *. Alors* N *est un processus de Poisson d'intensité* λ *.*

Exercice 11.3 a) Soient N et N' deux processus de Poisson indépendants, de paramètres λ et λ'. Montrer que $N + N'$ définit un processus de Poisson de paramètre $\lambda + \lambda'$.

b) Soit $\{X_n \,|\, n \in \mathbb{N}\}$ une suite i.i.d. indépendante de N. Le processus défini par $\Sigma_t := \sum_{n=1}^{N_t} X_n$ est appelé "processus de Poisson composé". (Exemple : N peut décrire le nombre de clients d'un magasin et X_n la dépense du n-ième client.) Calculer la transformée de Fourier de Σ_t en fonction de celle des X_n et de la fonction génératrice de N_t , puis l'espérance et la variance de Σ_t .

Exercice 11.4 a) Soit N un processus de Poisson d'intensité λ comptant les passages successifs de voitures sur une route. Un piéton a besoin d'une durée $d > 0$ entre deux voitures pour traverser la route. Calculer la transformée de Laplace du temps Y qu'il doit attendre avant de pouvoir traverser. (On trouve $a \longmapsto \frac{a+\lambda}{ae^{\lambda d}+\lambda e^{-ad}}$.) Calculer l'espérance de Y. (On trouve $\frac{e^{\lambda d}-1-\lambda d}{\lambda}$.)

b) Considérons une route à deux voies, les voitures étant comptées par deux processus de Poisson indépendants, d'intensité λ dans un sens et λ' dans l'autre, et nécessitant des durées respective d et d' pour être traversées. Soient Z le temps que le piéton doit attendre pour traverser les deux voies d'un coup, et Z' le temps qu'il doit attendre pour traverser les deux voies sachant qu'il peut les traverser séparément du fait de la présence d'un refuge au milieu de la route. Donner les transformées de Laplace de Z et de Z', leur espérances, et discuter l'intérêt du refuge.

Proposition 19 *(Processus de Poisson marqué) Fixons un entier* $r \ge 2$ *et une probabilité* $p = (p_1, .., p_r)$ *sur l'ensemble* $\{1, .., r\}$ *. Soit* $\{Y_n \,|\, n \in \mathbb{N}^*\}$ *une suite i.i.d. indépendante de* N*, de loi commune* p *. Accolons à chaque instant* S_n *décompté par* N *le type* Y_n *, et notons* N^j *le processus de comptage des instants de type* j *. Alors chaque* N^j *est un processus de Poisson d'intensité* $\lambda\, p_j$ *, et les processus* N^j *sont indépendants.*

Preuve Conditionnellement à $\{N_t = k\}$, les instants $S_1, .., S_k$ se répartissent dans les différents types suivant la loi multinômiale de paramètres k et $p_1, .., p_r$, par définition de la procédure. Nous avons donc pour tous $k_1, .., k_r \in \mathbb{N}$, notant $k := k_1 + .. + k_r$:

$$\mathbb{P}(N_t^1 = k_1, .., N_t^r = k_r) = \frac{k!}{k_1!..k_r!}\,(p_1)^{k_1}..(p_r)^{k_r} \times e^{-\lambda\, t}\,\frac{(\lambda\, t)^k}{k!} = \prod_{j=1}^{r} e^{-\lambda\, p_j\, t}\,\frac{(\lambda\, p_j\, t)^{k_j}}{k_j!} \;.$$

Ceci prouve que pour chaque t les variables $N_t^1, .., N_t^r$ sont indépendantes et ont chacune la loi de Poisson annoncée. Enfin la nullité en 0, la stationnarité, et l'indépendance des accroissements sont immédiats. ◇

Exercice 11.5 Le flux (poissonnien) des trams passant aux halles est en moyenne de 12 trams par heure et comporte une proportion $p = 40\%$ de trams de la ligne D (et de $1-p$ pour la ligne A).

a) Quelle est la probabilité qu'au moins 2 trams D passent en un quart d'heure donné ?

b) Sachant que 3 trams D sont passés en 20 minutes, quel est le nombre moyen de trams passés dans le même temps ?

c) Sachant que 10 trams sont passés en 45 minutes, quelle est la probabilité que la moitié soient des trams D ?

Exercice 11.6 Pour $t > 0$ fixé, vérifier que $S_{N_t} \le t < S_{1+N_t}$. S_{N_t} et S_{1+N_t} sont les points de $\mathcal{E}$ précédant et suivant directement l'instant t. Autrement dit, on peut les voir comme les instants de naissance et de décès de l'intervalle "vivant" à l'instant t. Notons $T'_t := t - S_{N_t}$ et $T''_t := S_{1+N_t} - t$ respectivement l'âge de cet intervalle vivant et le temps qu'il lui reste à vivre.

a) Montrer que pour $a, b > 0$ on a $I\!P(T'_t > a, T_t" > b) = e^{-\lambda(a+b)}\, 1_{\{t>a\}}$.

b) Déduire que T'_t et T''_t sont indépendants et donner leur lois.

c) Comparer la durée de vie moyenne de l'intervalle vivant à l'instant t avec celle du n-ième intervalle de $\mathcal{E}$; explication ?

Exercice 11.7 On veut estimer λ par le maximum de vraisemblance $\hat{\lambda}$ dans les deux cas suivant. a) On observe jusqu'à l'instant $t > 0$ fixé, et on connaît donc $N_t, S_1, .., S_{N_t}$. Calculer $\hat{\lambda}$, et vérifier qu'il est sans biais. (Utiliser la densité sachant N_t.)

b) On observe jusqu'à l'instant S_n, pour n fixé, et on connaît donc $S_1, .., S_n$. Calculer $\hat{\lambda}$, et vérifier qu'il est biaisé.

II. Eléments de statistique mathématique

12 Régression linéaire

Etant donné une statistique double (de données réelles) $X = (x_1, .., x_n)$, $Y = (y_1, .., y_n)$, on cherche une relation linéaire : $Y = aX + b$. Bien entendu elle ne peut être exacte sauf cas exceptionnel. Aussi cherche-t-on les coefficients (a, b) de façon que cette relation soit le plus près possible d'être vérifiée. La notion de proximité qu'on retient est celle qui conduit au calcul le plus simple : on cherche à minimiser la somme des carrés. (a, b) sera donc le couple de réels minimisant $\sum_{j=1}^{n}(y_j - ax_j - b)^2$. La droite d'équation $Y = aX + b$ correspondante est la droite de régression au sens des moindres carrés.

Faisons intervenir les moyennes empiriques $\bar{x} := \frac{1}{n}\sum_{j=1}^{n} x_j$, $\bar{y} := \frac{1}{n}\sum_{j=1}^{n} y_j$, puis la variance empirique $V_x := \frac{1}{n}\sum_{j=1}^{n} x_j^2 - \bar{x}^2$, et la covariance empirique $\kappa := \frac{1}{n}\sum_{j=1}^{n} x_j y_j - \bar{x}\bar{y}$.

Nous avons simplement :

$$\frac{1}{n}\sum_{j=1}^{n}(y_j - ax_j - b)^2 = (\bar{y} - a\bar{x} - b)^2 + V_x\, a^2 - 2\kappa\, a + V_y\,,$$

de sorte qu'apparaît aussitôt (pour peu que la statistique X ne soit pas constante) la solution unique suivante :

$$a = \kappa/V_x\;, \quad b = \bar{y} - a\bar{x}\,.$$

L'erreur quadratique totale vaut alors $\frac{1}{n}\sum_{j=1}^{n}(y_j - ax_j - b)^2 = V_y - \kappa^2/V_x$. Elle est faible lorsque qu'existe une relation linéaire entre les statistiques X et Y. Notons que la droite de régression passe par le centre de gravité $(\bar{x}, \bar{y})$ du <u>nuage</u> formé par les n points (x_j, y_j).

Il est naturel d'utiliser cette droite pour prédire une valeur supplémentaire y_{n+1} de la statistique Y, connaissant une valeur supplémentaire x_{n+1} de la statistique X : $y_{n+1} = ax_{n+1} + b$.

La qualité d'une telle prédiction (qui a priori n'a vraiment de sens que pour x_{n+1} proche de l'intervalle $[\min X, \max X]$) dépend de la valeur de $\sigma^2 := V_y - \kappa^2/V_x$, qui est la variance empirique de la statistique $Y - aX - b$. On peut déterminer une bande de confiance (on parle plutôt d'<u>intervalle de confiance</u>, voir la section 13.4 ci-dessous) par sa largeur autour de la droite de régression, de la façon suivante : pour $p \in]0,1[$ fixé, soit r un réel > 0 tel que $p = \frac{1}{n}\sum_{j=1}^{n} 1_{\{|y_j - ax_j - b| \le r\}}$. Pensant au théorème central limite, il est naturel d'approcher la loi empirique des $(y_j - ax_j - b)$ par une gaussienne, nécessairement centrée et de variance σ^2. De sorte qu'on doit avoir $p \approx \mathbb{P}(|G| \le r/\sigma)$, pour G gaussienne centrée réduite.

On peut ainsi estimer que pour x_{n+1} proche de l'intervalle $[\min X, \max X]$) on doit avoir par exemple 95% de chance de trouver y_{n+1} dans l'intervalle (de confiance à 95%)
$\Big[ax_{n+1} + b - (1,96)\sigma\,,\; ax_{n+1} + b + (1,96)\sigma\Big]$.

Bien entendu, on ne cherche pas toujours une relation linéaire entre les statistiques X et Y. Mais on peut bien souvent s'y ramener, par un changement de variable élémentaire, dès qu'on cherche une relation simple comportant 2 paramètres réels ; par exemple si l'on espère que $Y \approx \log(ae^X + b)$, il suffit de considérer le nuage formé par les (e^{x_j}, e^{y_j}) ; ou si l'on soupçonne que $Y \approx \frac{aX}{X^2+b}$, il suffit de considérer le nuage formé (ou la statistique double formée) par les $(x_j^2, x_j/y_j)$.

<u>Exercice n^o 12.1</u> Soient $X = (1, 3, 5, 7, 8, 9, 11, 12, 15, 18)$ et
$Y = (-33, -97, -134, -141, -130, -102, -32, 37, 343, 860)$. Représenter le nuage, et tester s'il y a une relation $Y = \alpha X^3 + \beta X$, ou $Y = \mathrm{ch}(aX - b) - 150$, ou $Y = \alpha(X - \beta)^2 - 150$.
Déterminer un intervalle de confiance à 95% pour la valeur y associée à $x = 14$.

13 Vraisemblance et estimation

13.1 Maximum de vraisemblance

Il s'agit d'estimer un paramètre inconnu θ d'une loi de probabilité $\mu = \mu_\theta$, sur la base d'une suite de réalisations, ou d'observations, effectuées suivant cette loi. Notons X une v.a. dont la loi est μ, et $x_1, .., x_n$ une suite d'observations, ou statistique, représentant des réalisations de X, qu'on suppose indépendantes. On considère les 2 cas les plus usuels : soit μ est discrète, soit elle admet une densité $g = g_\theta$. La fonction de vraisemblance associée à θ et à $x_1, .., x_n$ est :

$f(x_1, .., x_n \,|\, \theta) := \prod_{j=1}^n \mathbb{P}(X = x_j)$ lorsque μ est discrète, et

$f(x_1, .., x_n \,|\, \theta) := \prod_{j=1}^n g_\theta(x_j)$ lorsque μ admet la densité g_θ .

L'estimateur du maximum de vraisemblance est par définition la valeur de θ maximisant la fonction de vraisemblance :

$$\hat{\theta} := \operatorname{argmax} f(x_1, .., x_n \,|\, \theta) = \operatorname{argmax} \log f(x_1, .., x_n \,|\, \theta).$$

Exemples 1) μ est exponentielle de moyenne θ, id est $\mathcal{E}(1/\theta)$: alors $\log f(x_1, .., x_n \,|\, \theta) = -n \log \theta - (x_1 + .. + x_n)/\theta$, de sorte que $\hat{\theta} = (x_1 + .. + x_n)/n$.

2) $\mu = \mathcal{B}(m, p)$, avec m supposé connu et $\theta = p$: alors $\log f(x_1, .., x_n \,|\, p) = \sum_{j=1}^n \log C_m^{x_j} + \sum_{j=1}^n x_j \log p + (mn - \sum_{j=1}^n x_j) \log(1 - p)$, de sorte que $\hat{p} = (x_1 + .. + x_n)/(mn)$.

3) $\mu = \mathcal{P}(\lambda)$: alors $\log f(x_1, .., x_n \,|\, \lambda) = \sum_{j=1}^n x_j \log \lambda - n\lambda - \log(x_1!..x_n!)$, de sorte que $\hat{\lambda} = (x_1 + .. + x_n)/n$.

4) $\mu = \mathcal{N}(m, \sigma^2)$, avec $\theta = (m, \sigma)$: alors $\log f(x_1, .., x_n \,|\, m, \sigma) = -(n/2) \log(2\pi) - n \log \sigma - \sum_{j=1}^n (x_j - m)^2/(2\sigma^2)$, de sorte que $\hat{m} = \sum_{j=1}^n x_j/n$ et $\hat{\sigma} = \Big(\sum_{j=1}^n (x_j - \hat{m})^2/n \Big)^{1/2}$ (ce sont les seules valeurs qui annulent à la fois les deux dérivées partielles de la fonction de vraisemblance).

Exercice n^o 13.1.1 Supposons qu'un processus de Poisson de paramètre inconnu λ est observé : a) jusqu'à un temps fixe T ; b) jusqu'à la survenue du n-ième événement. Dans chacun de ces deux cas, quel est l'estimateur du maximum de vraisemblance pour λ ?

13.2 Biais, risque

Un estimateur $\hat{\theta} = \hat{\theta}(x_1, .., x_n)$ est dit sans biais lorsque sa moyenne est θ, pour la loi μ_θ : $\mathbb{E}[\hat{\theta}(X_1, .., X_n) |\, \theta] = \int \hat{\theta}(x_1, .., x_n)\, d\mu_\theta(x_1)..d\mu_\theta(x_n) = \theta$, et cela pour tout θ .

La qualité d'un estimateur est évaluée par son risque (quadratique) :

$R(\hat{\theta}, \theta) := \mathbb{E}\Big[[\hat{\theta}(X_1, .., X_n) - \theta]^2 \Big| \, \theta \Big]$, qui est égal à sa variance s'il n'est pas biaisé. Il n'y a généralement pas d'estimateur dont le risque soit minimal uniformément (id est pour tous les θ), même parmi les estimateurs sans biais. Il peut y avoir des estimateurs biaisés meilleurs (id est de risque moindre) que les estimateurs sans biais.

Exemple $\mu_\theta = \mathcal{E}(1/\theta)$, comme dans l'exemple 1 ci-dessus : notons $\hat\theta_1(x) := x_1$, qui est sans biais, de même que l'estimateur $\hat\theta = (x_1 + .. + x_n)/n$ du maximum de vraisemblance. On a aussitôt : $R(\hat\theta_1, \theta) = \theta^2 < R(\hat\theta, \theta) = \theta^2/n$. De même, si on considère une moyenne quelconque : $\hat\theta_\lambda = \lambda_1 x_1 + .. + \lambda_n x_n$, avec $\lambda_1, .., \lambda_n \geq 0$ tels que $\lambda_1 + .. + \lambda_n = 1$, on trouve encore un estimateur sans biais, de risque $R(\hat\theta_\lambda, \theta) = \lambda_1^2 + .. + \lambda_n^2$, qui est minimal pour $\hat\theta_\lambda = \hat\theta$.

On a plus généralement le résultat suivant, valable dès que les v.a.r.i.i.d. X_j sont de carré intégrable.

Proposition 20 *1) Parmi tous les estimateurs sans biais de* $\mathbb{E}(X_1)$ *qui sont des fonctions linéaires de* $(X_1, .., X_n)$*, la moyenne empirique* $\bar X_n := \frac{1}{n}\sum_{j=1}^n X_j$ *est de risque minimal.*

2) Parmi tous les estimateurs sans biais de $Var(X_1)$ *qui sont des fonctions quadratiques de* $(X_1, .., X_n)$*, la déviation standard* $V_n := \frac{1}{n-1}\sum_{j=1}^n (X_j - \bar X_n)^2$ *est de risque minimal.*

Preuve 1) Un estimateur $\hat\theta = h(X_1, .., X_n) = \sum_{j=1}^n \lambda_j X_j$ de $m = \mathbb{E}(X_1)$ est sans biais ssi $\sum_{j=1}^n \lambda_j = 1$. Notons $\bar\theta := \bar h(X_1, .., X_n) := \frac{1}{n!}\sum_{\pi\in\mathcal{S}_n} h(X_{\pi_1}, .., X_{\pi_n})$ sa fonction symétrisée, qui ici vaut clairement $\bar X_n$. La loi de $(X_1, .., X_n)$ étant invariante par les permutations $\pi \in \mathcal{S}_n$, nous avons pour toute fonction symétrique $Y_n = f(X_1, .., X_n)$:

$$\mathbb{E}[\hat\theta\, Y_n] = \int h(X_1, .., X_n) Y_n d\mathbb{P} = \int h(X_{\pi_1}, .., X_{\pi_n}) Y_n\, d\mathbb{P} = \int \bar h(X_1, .., X_n) Y_n\, d\mathbb{P} = \mathbb{E}[\bar\theta\, Y_n].$$

Cela entraîne $\mathbb{E}[\hat\theta] = \mathbb{E}[\bar\theta]$ et $\mathbb{E}[\hat\theta\bar\theta] = \mathbb{E}[\bar\theta^2]$, et donc :

$$\mathbb{E}[(m-\hat\theta)^2] - \mathbb{E}[(m-\bar\theta)^2] = \mathbb{E}[\hat\theta^2] - \mathbb{E}[\bar\theta^2] = \mathbb{E}[(\hat\theta - \bar\theta)^2] + 2\,\mathbb{E}[(\hat\theta - \bar\theta)\,\bar\theta] = \mathbb{E}[(\hat\theta - \bar\theta)^2],$$

d'où en effet $\mathbb{E}\Big[(\mathbb{E}(X_1) - \hat\theta)^2\Big] \geq \mathbb{E}\Big[(\mathbb{E}(X_1) - \bar\theta)^2\Big] = \mathbb{E}\Big[(\mathbb{E}(X_1) - \bar X_n)^2\Big]$.

2) De même, un estimateur $\hat\theta = g(X_1, .., X_n) = \sum_{1\leq j,k\leq n} \lambda_{jk} X_j X_k$ de $v = Var(X_1)$ est sans biais ssi $\sum_j \lambda_{jj} = 1 = -\sum_{j\neq k} \lambda_{jk}$, puisque $\mathbb{E}[\hat\theta] = v\sum_j \lambda_{jj} + m^2\Big(\sum_j \lambda_{jj} + \sum_{j\neq k} \lambda_{jk}\Big)$. Notons de nouveau $\bar\theta := \bar g(X_1, .., X_n)$ sa fonction symétrisée, qui est une forma quadratique dont les coefficients doivent vérifier les 2 équations ci-dessus, et être symétriques, donc égaux sur la diagonale, et égaux hors de la diagonale. Donc $\bar\theta = \frac{1}{n}\sum_j X_j^2 - \frac{2}{n(n-1)}\sum_{j\neq k} X_j X_k$. Or on obtient la même formule en développant V_n, de sorte que $\bar\theta = V_n$. La fin de la preuve est exactement la même qu'en (1) ci-dessus pour l'espérance. ◇

Exercice n° 13.2.1 Soient $X_1, .., X_n, ..$ des v.a.r.i.i.d. de loi $\mathcal{B}(p)$.

a) Si (pour n fixé) $X_1 + .. + X_n = k$, quel est l'estimateur du maximum de vraisemblance pour p ? Est-il sans biais ?

b) Supposons qu'on s'arrête au premier n tel que $X_1 + .. + X_n = k$, pour k fixé ; quel est alors l'estimateur du maximum de vraisemblance pour p ? Est-il sans biais ?

13.3 Statistique exhaustive

Une statistique exhaustive des v.a.i.i.d. $X_1, .., X_n$ est une fonction (déterministe et indépendante de θ) $\tau = \tau(X_1, .., X_n)$ qui permet de factoriser la fonction du maximum de vraisemblance sous la forme : $f(X_1, .., X_n \,|\, \theta) = h(\tau, \theta) \times k(X_1, .., X_n)$, pour une fonction k indépendante de θ.

Dans les exemples 1,2,3 ci-dessus, $\tau(X_1, .., X_n) := X_1 + .. + X_n$ est une statistique exhaustive. Dans l'exemple 4 ci-dessus,
$\tau(X_1, .., X_n) = (\bar{X}, V) := \left(\frac{X_1+..+X_n}{n}, (X_1 - \bar{X})^2 + .. + (X_n - \bar{X})^2\right)$ est une statistique exhaustive ; donc $\left((X_1 + .. + X_n), (X_1^2 + .. + X_n^2)\right)$ en est une autre.

Un théorème (de Rao-Blackwell) énonce que s'il existe une statistique exhaustive $\tau = \tau(X_1, .., X_n)$ et si $\hat{\theta} = \hat{\theta}(X_1, .., X_n)$ est un estimateur sans biais de θ, alors une certaine fonction de τ (égale à $\mathbb{E}[\hat{\theta} \,|\, \tau]$) est un estimateur sans biais de θ, de risque $\leq R(\hat{\theta}, \theta)$. Ce qui signifie qu'en présence d'une statistique exhaustive τ, les estimateurs sans biais ne sont à rechercher que parmi les fonctions de τ. Lorsque de plus il n'existe qu'un seul estimateur sans biais fonction d'une statistique exhaustive, alors il est clairement optimal parmi les estimateur sans biais.

Exemple Reprenons l'exemple 3 vu plus haut : $\mu = \mathcal{P}(\lambda)$. La statistique $\bar{X} := \sum_{j=1}^{n} X_j/n$ est exhaustive, et optimale pour l'estimation de λ.

Cherchons un estimateur optimal pour $\theta = e^{-\lambda}$, en observant que l'estimateur élémentaire $1_{\{X_1=0\}}$ est sans biais. La statistique $\bar{X}$ est encore exhaustive, et donc $\mathbb{P}\left(X_1 = 0 \,\middle|\, \bar{X}\right)$ est un bon estimateur, en fait optimal. Or (voir l'exercice 5.3) la loi de X_1 sachant $n\bar{X}$ est $\mathcal{B}(n\bar{X}, 1/n)$, de sorte que $\mathbb{P}\left(X_1 = 0 \,\middle|\, \bar{X}\right) = (1 - 1/n)^{n\bar{X}}$.

Exercice n^o 13.3.1 a) Soient $X_1, .., X_n$ des v.a.r.i.i.d. uniformes sur un intervalle $[0, \theta]$. Montrer que $\max\{X_1, .., X_n\}$ est l'estimateur du maximum de vraisemblance pour θ, et que c'est aussi une statistique exhaustive pour θ.

b) Soient $X_1, .., X_n$ des v.a.r.i.i.d. de loi $\mathcal{U}([\theta, \theta'])$. Trouver une statistique exhaustive pour (θ, θ'), et les estimateurs du maximum de vraisemblance pour θ et θ'.

Exercice n^o 13.3.2 Soient $X_1, .., X_n$ des v.a.r.i.i.d. de loi $\mathcal{E}(\lambda)$. Montrer que l'estimateur sans biais de risque minimum de $e^{-2\lambda}$ est $\hat{\theta} := 1_{\{S_n > 2\}} \left(\frac{S_n - 2}{S_n}\right)^{n-1}$. (On a noté $S_n := X_1 + .. + X_n$.) (Indication : Considérer $\mathbb{P}(X_1 > 2)$, et utiliser que la loi de X_1 sachant S_n est celle du minimum de $(n-1)$ v.a.i.i.d. uniformes sur $[0, S_n]$ (à justifier).)

Exercice n^o 13.3.3 Soient $X_1, .., X_n$, $Y_1, .., Y_n$ des v.a.r.i.i.d., les X_j de loi $\mathcal{E}(\lambda)$, et les Y_j de loi $\mathcal{E}(\lambda')$. Trouver une statistique exhaustive pour (λ, λ'), et les estimateurs du maximum de vraisemblance pour λ et λ'. Ont-ils un biais ?

13.4 Intervalles de confiance

On veut estimer la moyenne inconnue m d'une loi μ (sur $I\!R$) d'écart-type σ , dont un échantillon est $x_1,..,x_n$. On assimile pour cela ces valeurs à des v.a.r.i.i.d. de loi μ, et on leur applique la loi des grands nombres : $\bar{x} := (x_1 + .. + x_n)/n$ doit être proche de m, au moins pour n assez grand.

Afin de quantifier cette proximité, on cherche un intervalle $J = [a,b] = J(x_1,..,x_n)$, dont on puisse dire qu'il contient m avec une grande probabilité. Pour cela, on applique le théorème central limite : $(m-\bar{x})\sqrt{n}\,/\sigma = (nm - x_1 + .. + x_n)/(\sigma\sqrt{n}\,)$ doit être proche d'une v.a. G de loi $\mathcal{N}(0,1)$; posant $\alpha := (a-\bar{x})\sqrt{n}\,/\sigma$ et $\beta := (b-\bar{x})\sqrt{n}\,/\sigma$, ce qui équivaut à $a = \bar{x} + \alpha\,\sigma/\sqrt{n}$ et $b = \bar{x} + \beta\,\sigma/\sqrt{n}$, on obtient :

$$I\!P(m \in J) = I\!P\Big(\alpha \leq (m-\bar{x})\sqrt{n}\,/\sigma\ \leq \beta\Big) \approx I\!P(G \in [\alpha,\beta]).$$

On choisit ordinairement (mais ce n'est pas une obligation) l'intervalle $I := [\alpha,\beta]$ centré, et de sorte que $I\!P(G \in I)$ vaille le plus souvent $0,9$ ou $0,95$. Les tables de la loi normale contiennent les valeurs classiques :

$$I\!P(\,|G| \leq 1,65) \approx 0,90 \quad \text{ et } \quad I\!P(\,|G| \leq 1,96) \approx 0,95\,.$$

Prenant ainsi $\alpha = -(1,96) = -\beta$, on obtient

$$I\!P\Big(\bar{x} - (1,96)\,\sigma/\sqrt{n}\ \leq m \leq \bar{x} + (1,96)\,\sigma/\sqrt{n}\Big) \approx 0,95\,.$$

On dit que $J = [\bar{x} - (1,96)\,\sigma/\sqrt{n}\,,\ \bar{x} + (1,96)\,\sigma/\sqrt{n}\,]$ est un intervalle de confiance à 95% pour m .

Il est généralement irréaliste de supposer l'écart-type σ connu. On peut alors le remplacer par un estimateur (voir la proposition 20 ci-dessus) :

$$\sigma \approx \hat{\sigma} = \hat{\sigma}(x_1,..,x_n) = \sqrt{\frac{1}{n-1}\,\sum_{j=1}^{n}(x_j-\bar{x})^2}\ = \sqrt{\frac{1}{n-1}\,\Big(\sum_{j=1}^{n}x_j^2 - n\,\bar{x}^2\Big)}\ .$$

L'intervalle de confiance $J = [\bar{x} + \alpha\,\hat{\sigma}/\sqrt{n}\,,\ \bar{x} + \beta\,\hat{\sigma}/\sqrt{n}\,]$ est alors réellement fonction de l'échantillon $(x_1,..,x_n)$ (et de n).

Exercice n^o 13.4.1 Une caractéristique électrique d'un composant électronique varie, du fait des dispersions de fabrication, suivant une loi normale $\mathcal{N}(\mu,\mu^2)$. On se propose d'estimer son espérance μ par un intervalle de confiance. Un prélèvement de 20 composants effectué sur un lot de série a donné les résultats suivant : 10,1 ; 10,5 ; 9,4 ; 10,2 ; 9,5 ; 9,8 ; 10,1 ; 10,3 ; 10,6 ; 9,7 ; 10,2 ; 10,4 ; 10,3 ; 9,6 ; 9,9 ; 9,8 ; 10,1 ; 10,3 ; 10,5 ; 9,8 . ($\overline{x} = 10,055$).

a) Construire un intervalle de confiance pour μ de niveau de confiance 90%. Est-il le plus court parmi tous les intervalles de confiance de niveau de confiance 90% ?

b) Construire un intervalle de confiance pour μ de niveau de confiance 95%.

Exercice n^o 13.4.2 La fiabilité dans le temps d'un composant électrique, c'est-à-dire sa durée de vie X, est supposée de type exponentiel : $\mathbb{P}(X \geq t) = \exp(-t/\lambda), t \geq 0$. On cherche à estimer à l'aide d'un intervalle de confiance le paramètre λ. Si la moyenne observée d'un échantillon de taille n est $\overline{x}$, donner un intervalle de confiance de niveau de confiance 95% du paramètre λ. Est-il le plus court parmi tous les intervalles de confiance de niveau de confiance 95% ? Traiter d'abord le cas de n grand, puis ensuite le cas général (plus délicat).

Exercice n^o 13.4.3 Ayant lu dans le magazine ELLE du 19.8.02 que 55% des Français utilisaient des bains moussants, vous décidez de vérifier vous-même l'affirmation du magazine ELLE par un mini-sondage, en posant autour de vous la question :

"Utilisez-vous des bains moussants ? OUI, NON"

et en construisant un intervalle de confiance sur une proportion.

a) Vous envisagez de ne considérer l'affirmation comme confirmée que si cet intervalle de 90% basé sur le groupe des personnes interrogées est inclus dans $[54\%; 56\%]$. Est-ce raisonnable ?

b) Découragé(e) par la réponse obtenue en a), vous vous contentez d'un échantillon de 50 personnes. Si la proportion de oui est de 53% parmi les 50 réponses, déterminez un intervalle de confiance de 90% .

13.5 Estimation bayésienne

Lorsqu'on se donne une loi a priori sur l'ensemble des valeurs possibles de θ, disons ν , on considère l'estimateur bayésien $\hat{\theta}_B := \mathbb{E}\left[\theta \,\middle|\, X_1, .., X_n\right]$ (où l'espérance est calculée suivant la loi ν a priori donnée), qui a la vertu de minimiser $\mathbb{E}\left[(\hat{\theta} - \theta)^2 \,\middle|\, X_1, .., X_n\right]$, parmi tous les estimateurs $\hat{\theta} = \hat{\theta}(X_1, .., X_n)$, et pour toute valeur de $(X_1, .., X_n)$ (vérification facile : il suffit d'écrire $(\hat{\theta} - \hat{\theta}_B + \hat{\theta}_B - \theta)^2 = ...$ et de constater la nullité du terme rectangle).

La relation avec la formule de Bayes est directe : la densité conditionnelle de θ sachant $(X_1, .., X_n)$ est $f(\theta \,|\, X_1, .., X_n) = f(X_1, .., X_n \,|\, \theta) \Big/ \int f(X_1, .., X_n \,|\, \theta) d\nu(\theta)$.

Ce qui entraîne aussitôt :

$$\hat{\theta}_B = \int f(X_1, .., X_n \,|\, \theta)\, \theta \, d\nu(\theta) \Big/ \int f(X_1, .., X_n \,|\, \theta) d\nu(\theta) \,;$$

et aussi

$$\mathbb{P}[a(X_1, .., X_n) < \theta < b(X_1, .., X_n)] = \frac{\int_{a(X_1,..,X_n)}^{b(X_1,..,X_n)} f(X_1, .., X_n \,|\, \theta) \, d\nu(\theta)}{\int f(X_1, .., X_n \,|\, \theta) d\nu(\theta)} \,,$$

qui permet d'obtenir un intervalle de confiance $[a(x_1, .., x_n) \,,\, b(x_1, .., x_n)]$.

Exemples 4) $\mu_\theta = \mathcal{B}(\theta)$ et $\nu = \mathcal{U}([0,1]$: $f(\theta \,|\, x_1,..,x_n) = \dfrac{\theta^{\sum_{j=1}^n x_j} \times (1-\theta)^{n-\sum_{j=1}^n x_j}}{\int_0^1 \theta^{\sum_{j=1}^n x_j} (1-\theta)^{n-\sum_{j=1}^n x_j}\, d\theta}$,
d'où, en utilisant que $\displaystyle\int_0^1 \theta^k (1-\theta)^\ell d\theta = \frac{k!\,\ell!}{(k+\ell+1)!}$: $I\!\!E\Big[\theta \,\Big|\, X_1,..,X_n\Big] = \frac{\sum_{j=1}^n X_j + 1}{n+2}$.

5) $\mu_\theta = \mathcal{N}(\theta, 1)$ et $\nu = \mathcal{N}(0,1)$: on obtient $I\!\!E\Big[\theta \,\Big|\, X_1,..,X_n\Big] = \frac{\sum_{j=1}^n X_j}{n+1}$, car on trouve $f(\theta \,|\, x_1,..,x_n) = k(x_1,..,x_n) \times \exp\Big[-\frac{1}{2}(n+1)\Big(\theta - \frac{\sum_{j=1}^n x_j}{n+1}\Big)^2\Big]$, de sorte que la loi conditionnelle de θ sachant $(X_1,..,X_n)$ est $\mathcal{N}\Big(\frac{\sum_{j=1}^n X_j}{n+1}, \frac{1}{n+1}\Big)$; un intervalle de confiance à 95% pour θ est $\Big[\frac{\sum_{j=1}^n x_j}{n+1} - \frac{1{,}96}{\sqrt{n+1}}, \frac{\sum_{j=1}^n x_j}{n+1} + \frac{1{,}96}{\sqrt{n+1}}\Big]$.

Exercice n^o 13.5.1 Un processus de Poisson de paramètre inconnu λ est observé durant le temps fixe T. Soit N le nombre d'événements observés durant ce temps. On se donne pour λ la loi a priori $\mathcal{E}(\alpha)$. Calculer l'estimateur bayésien de λ.

14 Tests

Au lieu d'estimer le paramètre inconnu θ, on cherche ici à savoir s'il vérifie une hypothèse statistique notée H_0, nommée aussi hypothèse nulle, et pour cela on emploie un test. Par exemple $H_0 = \{\theta < 7\}$, et le test est une règle devant permettre de trancher ou de parier, en fonction d'un échantillon statistique relatif à μ_θ, si l'hypothèse nulle H_0 doit être acceptée ou bien rejetée. Dans le cas le plus général, on considère une hypothèse alternative H_1, contre laquelle on teste H_0, conditionnellement à $H_0 \cup H_1$. Ici on a choisi simplement $H_1 = (H_0)^{\mathbf{c}}$.

Bien entendu, l'acceptation ou le rejet de l'hypothèse ne peut pas être une décision déterministe : il y a toujours en pratique une probabilité non nulle que la décision arrêtée soit erronée. Il faut donc décider a priori un seuil de confiance, id est un niveau de test, c'est-à-dire une probabilité de rejet à tort. La plupart du temps, on fixe ce niveau à 5%. Autrement dit, on veut que $\sup_{\theta \in H_0} I\!\!P(\text{rejeter } H_0) = 0,05$.

Lorsque θ est la moyenne de μ_θ, il est naturel de prendre le test sous forme d'une fonction de la moyenne empirique $\bar{X} = \bar{X}_n$. Si par exemple $H_0 = \{\theta \le a\}$, on prendra pour test $\{\bar{X} > a'\}$ (c'est la région critique : si la statistique observée tombe dedans, on décide de rejeter l'hypothèse nulle), avec $a' > a$ à déterminer pour que le niveau soit bien 5%, c'est-à-dire ici pour que $\mu_a^{\otimes n}(\bar{X} > a') = 0,05$.

Pour n grand, si σ est supposé connu, recourant au théorème central limite, on assimilera $\sqrt{n}\,(\bar{X}_n - a)/\sigma$ à une variable gaussienne centrée réduite G, de sorte que a' doit vérifier :

$$0,05 = \mu_a^{\otimes n}(\bar{X} > a') = I\!\!P\Big(\sqrt{n}\,(\bar{X}_n - a)/\sigma > \sqrt{n}\,(a' - a)/\sigma\Big) \approx I\!\!P\Big(G > \sqrt{n}\,(a' - a)/\sigma\Big),$$

ce qui donne $\sqrt{n}\,(a' - a)/\sigma \approx 1,96$, id est $a' \approx a + (1,96)\,\sigma/\sqrt{n}$.

Mais comme déjà dit dans la section 13.4, il est généralement irréaliste de supposer l'écart-type σ connu.

Exemples 1) $\mu_\theta = \mathcal{E}(1/\theta)$. Via le changement de variable $s_1 = x_1, .., s_n = x_1 + .. + x_n$:

$$\mu_a^{\otimes n}(\bar{X} < a') = a^{-n} \int_{x_1+..+x_n<na'} e^{-(x_1+..+x_n)/a} dx_1..dx_n = \int_{x_1+..+x_n<na'/a} e^{-(x_1+..+x_n)} dx_1..dx_n$$

$$= \int_{\{0<s_1<..<s_n<na'/a\}} e^{-s_n} ds_1..ds_n = \int_0^{na'/a} \Big(\int_{\{0<s_1<..<s_{n-1}<s_n\}} ds_1..ds_{n-1} \Big) e^{-s_n}\, ds_n$$

$$= \int_0^{na'/a} \frac{s^{n-1}}{(n-1)!}\, e^{-s}\, ds = 1 - e^{-na'/a} \sum_{j=0}^{n-1} \frac{(na'/a)^j}{j!} = e^{-na'/a} \sum_{j\ge n} \frac{(na'/a)^j}{j!} \;.$$

(Comparer avec le théorème 7 de la section 11.) Il est assez légitime d'approcher la loi $\mathcal{P}(na'/a)$ par une loi normale, id est $\mathcal{N}(na'/a, na'/a)$. Ce qui donne (pour une variable gaussienne centrée réduite G) :

$$0,05 = \mu_a^{\otimes n}(\bar{X} > a') = \mathcal{P}(na'/a)([0,n[) \approx I\!P\Big(G > \sqrt{n}\,(\sqrt{a'/a} - \sqrt{a/a'}\,)\Big),$$

et donc $a'/a \approx 1 + (0,98)/\sqrt{n}$.

2) $\mu_\theta = \mathcal{P}(\theta)$. Alors $\mu_a^{\otimes n}(\bar{X} > a') = \mathcal{P}(na)[\{j \,|\, j > na'\}] = e^{-na} \sum_{j>na'} \frac{(na)^j}{j!}$.

Il est légitime d'approcher la loi $\mathcal{P}(na)$ par une loi normale, id est $\mathcal{N}(na, na)$. Ce qui donne (pour une variable gaussienne centrée réduite G) :

$$0,05 = \mu_a^{\otimes n}(\bar{X} > a') \approx I\!P\Big(G > \sqrt{n/a}\,(a' - a)\Big),$$

et donc $a' \approx a + (1,96)\sqrt{a/n}$.

Exercice n^o 14.1 Deux associés A et B se partagent les tâches d'un cabinet commun de conseil. L'associé A pense qu'il traite moins de 30% des dossiers. Pour vérifier cela, il décide de choisir au hasard un échantillon de 100 dossiers (traités soit par lui-même, soit par l'associé B), sur lesquels il prévoit d'effectuer deux tests, avec le même risque de rejet à tort de 5%, en prenant successivement comme hypothèse nulle

i) $H_0 : p \le 30\%$ ii) $H_0' : p \ge 30\%$, p étant la proportion réelle des dossiers traités par A.

a) Pour combien de dossiers traités par A (au plus) l'hypothèse H_0 est-elle acceptée ?

b) Pour combien de dossiers traités par A (au plus) H_0' est-elle rejetée ?

c) Déterminer la probabilité d'accepter l'hypothèse H_0 si la proportion réelle de dossiers traités par A est de 40%.

d) Déterminer la probabilité de rejeter l'hypothèse H_0' si la proportion réelle de dossiers traités par A est de 40%.

BIBLIOGRAPHIE

BREMAUD P. *Introduction aux probabilités.* Springer, Berlin 1984.

CHUNG K.L. *A course in probability theory.* Academic Press 1974.

COTTRELL M., GENON-CATALOT V., DUHAMEL C., MEYRE T. *Exercices de probabilités.* Cassini, Paris 1999.

FOATA D., FUCHS A. *Calcul des probabilités.* Dunod, Paris 1998.

HOGG R., CRAIG A. *Introduction to mathematical statistics.* The Macmillan Company, New York 1965.

LESIGNE E. *Une introduction aux théorèmes limites du calcul des probabilités.* Ellipses, Paris 1997.

LETAC G. *Intégration et probabilités.* Masson, Paris 1997.

METIVIER M., NEVEU J. *Cours de probabilités.* Ecole polytechnique, 1979.

REES D.G. *Essential Statistics.* Chapman & Hall, London New York 1989.

ROSS S. *Introduction to probability models.* Academic press, 1980.