Résumé de cours de probabilités pour l'agrégation interne

Table des matières

1 Espace de probabilités

1.1 Comment définir "la somme" d'une famille dénombrable de réels ?

La bonne façon de faire des probabilités est de regarder l'intégrale de Lebesgue, on peut s'en passer si on se limite au cas dénombrable, dans ce cas la notion de famille sommable suffit (hors programme voir partie 1.2), si on veut y échapper on peut tout ramener à des série mais c'est un peu alambiqué. C'est toutefois ce que je présente rapidement dans cette partie 1.1. Il est entendu que faire la somme d'un nombre fini de réels ne pose pas de problème.

Proposition 1 : Soit σ une bijection de $\mathbb{N}$ dans $\mathbb{N}$, si la série $\sum u_n$ est absolument convergente, alors la série de terme général $u_{\sigma(n)}$ est convergente et a même somme que $\sum u_n$.

Exemple 1 : Soit (u_n) définie par $u_{2n} = \frac{1}{n+1}$ et $u_{2n-1} = \frac{-1}{n}$, et σ la bijection de $\mathbb{N}$ sur $\mathbb{N}$ définie par

$$\begin{cases} \sigma(3n) = 2n \\ \sigma(3n+1) = 4n+1 \\ \sigma(3n+2) = 4n+3 \end{cases}$$

Il est facile de voir $\sum u_n = 0$ et on montre à l'aide de série entière par exemple que $\sum u_{\sigma(n)} = \ln 2$. La convergence absolue est donc une condition importante pour pouvoir sommer sans tenir compte de l'ordre.

Proposition 2 : Soit I un ensemble dénombrable et ϕ une bijection de $\mathbb{N}$ dans I, si la série $\sum u_{\phi(n)}$ est absolument convergente, alors pour toute bijection ψ de $\mathbb{N}$ dans I la série $\sum u_{\psi(n)}$ est absolument convergente et a même somme que la précédente.

Définition 1 : Dans ce cas on dit que la famille $(u_i)_{i\in I}$ est sommable et on note $\sum_{i\in I} u_i$ la somme de cette série.

Proposition 3 : Soit I fini ou dénombrable, non vide, et $(I_x)_{x\in X}$ une partition finie ou dénombrable de I. $(u_i)_{i\in I}$ est sommable ssi pour tout $x \in X$ la famille $(u_i)_{i\in I_x}$ est sommable et si on note S_x sa somme la famille $(S_x)_{x\in X}$ est sommable. On a alors :

$$\sum_{x\in X} S_x = \sum_{x\in X} \sum_{i\in I_x} u_i = \sum_{i\in I} u_i$$

1.2 Familles sommables

Hors programme mais c'est la façon agréable de penser les choses, avec les résultats qui précédent on peut s'en passer.

1.2.1 Familles positives

Définition 2 : Soit $(u_i)_{i\in I}$ une famille de réels positifs, indicée par I fini ou dénombrable, non vide, cette famille est dite sommable si l'ensemble des sommes d'un nombre finie de u_i est majorée, on appelle alors somme de la famille et on note $\sum_{i\in I} u_i$ la borne supérieur de l'ensemble $\{\sum_{i\in J} u_i | J \text{ est une partie finie de } I\}$.

Proposition 4 : Si $(u_i)_{i\in I}$ famille positive, est sommable toute famille $(u_i)_{i\in J}$ ou J est une partie de I est sommable et $\sum_{i\in J} u_i \leq \sum_{i\in I} u_i$.

Proposition 5 : Une famille positive $(u_i)_{i\in I}$ avec I dénombrable est sommable ssi il existe une bijection φ de $\mathbb{N}$ dans I telle que la série de terme général $u_{\varphi(n)}$ converge.

Proposition 6 : Soit I fini ou dénombrable, non vide, et $(I_x)_{x\in X}$ une partition finie ou dénombrable de I. $(u_i)_{i\in I}$ est sommable ssi pour tout $x\in X$ la famille $(u_i)_{i\in I_x}$ est sommable et si on note S_x sa somme, la famille $(S_x)_{x\in X}$ est sommable. On a alors :

$$\sum_{x\in X}\sum_{i\in I_x} u_i = \sum_{x\in X S_x} = \sum_{i\in I} u_i$$

Remarque 1 : Il y a un peu la même différence entre intégrale généralisée et intégrabilité qu'entre série et famille sommable.

1.2.2 Familles de réels

Définition 3 : Soit $(u_i)_{i\in I}$ une famille de réels, indicée par I fini ou dénombrable, non vide, cette famille est dite sommable si $(|u_i|)_{i\in I}$ est sommable, on définie alors sa somme par :

$$\sum_{i\in I} u_i = \sum_{i\in\{i\in I|u_i>0\}} |u_i| - \sum_{i\in\{i\in I|u_i<0\}} |u_i|$$

Proposition 7 : Si $(u_i)_{i\in I}$ est sommable toute famille $(u_i)_{i\in J}$ ou J est une partie de I est sommable.

Proposition 8 : Une famille $(u_i)_{i\in I}$ est sommable ssi il existe une bijection φ de $\mathbb{N}$ dans I telle que la série de terme général $u_{\varphi(n)}$ converge absolument.

Remarque 2 : Cette proposition prouve que les deux définitions 1 et 2 de familles sommables sont équivalentes.

Remarque 3 : $(u_n)_{n\in\mathbb{N}}$ est sommable ssi la série $\sum u_n$ est absolument convergente, on a alors la somme de la série et la somme de la famille qui sont égales.

Remarque 4 : Une des différences fondamentale qu'il y a entre famille sommable et série est l'ordre, dans une série les éléments sont ordonnés, alors que l'on peut regarder une famille sommable indicée par les rationnels $(\mathbb{Q})$.

Proposition 9 : Soit I fini ou dénombrable, non vide, et $(I_x)_{x\in X}$ une partition finie ou dénombrable de I. Dans le cas ou la famille $(u_i)_{i\in I_x}$ est sommable, on note S_x sa somme.
La famille $(u_i)_{i\in I}$ est sommable ssi pour tout $x\in X$ la famille $(u_i)_{i\in I_x}$ est sommable et la famille $(S_x)_{x\in X}$ est sommable. On a alors :

$$\sum_{x\in X} S_x = \sum_{x\in X}\sum_{i\in I_x} u_i = \sum_{i\in I} u_i$$

Proposition 10 : Soient I fini ou dénombrable, non vide, et $(I_n)_{n\in\mathbb{N}}$ une suite croissante de parties de I telle que $\cup I_n = I$ et $(u_i)_{i\in I}$ une famille sommable alors :

$$\lim_{n\to\infty}\sum_{i\in I_n} u_i = \sum_{i\in I} u_i$$

Preuve : Pour une famille positive, toute famille finie de I_n est une famille finie de I donc

$$\sum_{i\in I_n} u_i \leq \sum_{i\in I} u_i$$

Soit $\epsilon > 0$ alors il existe un J finie de I, tel que $\sum_{i\in J} u_i > \sum_{i\in I} u_i - \epsilon$, et il existe un N à partir duquel tout les I_n contiennent J. D'ou le résultat. Pour une famille quelconque, on regarde les termes positifs et les termes négatifs.

1.3 Généralités

Définition 4 : Une tribu sur un ensemble Ω est un ensemble $\mathcal{T}$ de parties de Ω qui vérifie :

1. $\varnothing \in \mathcal{T}$
2. Si une partie A est dans $\mathcal{T}$ alors son complémentaire appartient aussi à $\mathcal{T}$
3. Si une famille finie ou dénombrable de parties de Ω appartiennent à $\mathcal{T}$ alors leur réunion aussi

On appelle événement les éléments de la tribu.

Remarque 5 : Dans la théorie qui va suivre, on cherche à mesurer les parties de Ω, or pour pouvoir construire une théorie intéressante, il faut se limiter à certaines parties, ces parties vont être les éléments de la tribu. Une intersection finie ou dénombrable d'éléments d'une tribu appartient à la tribu.

Définition 5 : Une probabilité sur un ensemble Ω muni d'une tribu $\mathcal{T}$ est une application de $\mathcal{T}$ dans $\mathbb{R}^+$ qui a les 2 propriétés suivantes : $P(\Omega) = 1$ et si $(A_i)_{i\in I}$ est une famille finie ou dénombrable d'éléments de $\mathcal{T}$ deux à deux disjoints alors la famille $(P(A_i))_{i\in\mathbb{N}}$ est sommable et sa somme est égale à $P(\cup A_i)$.

Remarque 6 : La deuxième propriété est une propriété de mesure, penser aux aires, aux volumes, aux masses...

Proposition 11 : Propriétés d'une probabilité :
$P(\emptyset) = 0$, $P(\overline{A}) = 1 - P(A)$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Proposition 12 : Si $(A_n)_{n\in\mathbb{N}}$ est une famille croissante de $\mathcal{T}$ alors $\lim_n P(A_n) = P(\cup_n A_n)$

Preuve : On pose $A_0 = \varnothing$, on a alors $\cup A_n = \cup(A_n \setminus A_{n-1})$ et les ensembles $A_n \setminus A_{n-1}$ sont disjoints deux à deux on a donc

$$P(\cup A_n) = \sum P(A_n \setminus A_{n-1}) = \sum \big(P(A_n) - P(A_{n-1})\big) = \lim_{N\to\infty} \sum_{n=1}^{N} \big(P(A_n) - P(A_{n-1})\big)$$

or $\displaystyle\sum_{n=1}^{N} \big(P(A_n) - P(A_{n-1})\big) = P(A_N) - P(A_0)$ d'ou $P(\cup A_n) = \lim P(A_n)$

1.4 Probabilités conditionnelles et indépendance

Définition 6 : Si B est un événement de probabilité non nulle, et A un événement alors on définit la probabilité conditionnelle de A sachant B par

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Proposition 13 : Si P est une probabilité sur $(\Omega, \mathcal{T})$, et B un événement de probabilité non nulle, alors $P(.|B)$ définie une nouvelle probabilité sur $(\Omega, \mathcal{T})$.

Remarque 7 : Cette propriété justifie l'écriture $P_B(A) = P(A|B)$

Proposition 14 : Soit $(A_{i\in I})$ une partition finie ou dénombrable d'événements de probabilité non nulle, de Ω alors $P(A) = \sum_{i\in I} P(A \cap A_i) = \sum_{i\in I} P(A|A_i)P(A_i)$.

Définition 7 : Deux événements A et B sont indépendants si $P(A \cap B) = P(A)P(B)$.
Une famille d'événements $(A_i)_{i\in I}$ sont indépendants si pour toute partie finie J de I on a

$$P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j)$$

Remarque 8 : L'indépendance et l'indépendance deux à deux sont deux notions différentes, l'indépendance entraine l'indépendance deux à deux mais la réciproque est fausse.

Remarque 9 : Le terme indépendance est bien choisi car il permet de modéliser l'idée non mathématique d'indépendance de deux résultats, par exemple lorsque je lance deux dés, le résultat de l'un est indépendant du résultat de l'autre, il n'y a aucune notion mathématique ici, mais on va modéliser le problème de tel sorte que les événements lié à chacun des dés soient indépendants.

2 Variable aléatoire réelle

Définition 8 : Soit $(\Omega, \mathcal{T}, P)$ défini comme précédemment, (on parle d'un espace de probabilité ou d'un espace probabilisé). Une variable aléatoire est une fonction de Ω dans $\mathbb{R}$ telle que l'image réciproque de tout intervalle de $\mathbb{R}$ appartienne à $\mathcal{T}$. En théorie de l'intégration on parle de fonction mesurable .

Remarque 10 : Si $\mathcal{T}$ est l'ensemble des parties de Ω, alors toute fonction de Ω dans $\mathbb{R}$ est une variable aléatoire.

Définition 9 : On note $(X = a)$ pour $X^{-1}(\{a\})$, $(X \in I)$ pour $X^{-1}(I)$ et $(X < a)$ pour $X^{-1}(]-\infty, a[)$, etc . . .

Définition 10 : (hors programme) On appelle tribu borélienne de $\mathbb{R}$, et on note $\mathcal{B}^1$, la plus petite tribu de $\mathbb{R}$ qui contient tous les intervalles de $\mathbb{R}$.

Proposition 15 : (admis, hors programme) Pour qu'une fonction de Ω dans $\mathbb{R}$ soit une variable aléatoire il faut et il suffit que l'image réciproque de tout élément de la tribu borélienne $\mathcal{B}^1$, appartienne à $\mathcal{T}$, c'est en fait la bonne notion de variable aléatoire.

Définition 11 : Dans ce cas la fonction P_X définie sur $\mathcal{B}^1$, à valeur dans $[0, 1]$, par $P_X(A) = P(X^{-1}(A))$ est une probabilité sur $(\mathbb{R}, \mathcal{B}^1)$, on l'appelle loi de X. La loi d'une variable aléatoire est donc une probabilité sur $(\mathbb{R}, \mathcal{B}^1)$.

Remarque 11 : Très souvent, on travaille sur une variable aléatoire, "sans la connaître", il arrive même fréquemment que Ω ne soit pas explicité, on connaît juste la loi de la variable aléatoire, et c'est suffisant. Ainsi pour modéliser certains phénomènes, on peut soit utiliser un espace de probabilité, soit utiliser une variable aléatoire et travailler sur sa loi.

Proposition 16 : (admis) Soient X, Y deux variables aléatoires définies sur un même espace de probabilité, et λ un réel, alors $X + Y, \lambda X, XY, \max(X, Y), \min(X, Y)$ sont des variables aléatoires, si de plus f est une fonction continue par morceaux de $\mathbb{R}$ dans $\mathbb{R}$ alors $f \circ X$ est aussi une variable aléatoire.

Définition 12 : On appelle fonction de répartition de la variable aléatoire X la fonction F_X définie sur $\mathbb{R}$ par $F_X(t) = P(X \leq t)$.

Remarque 12 : La fonction de répartition est bien définie car $X^{-1}(]-\infty, t])$ appartient à la tribu.

Proposition 17 : (hors programme, énoncé dans toute sa généralité) La loi d'une variable aléatoire est entièrement déterminée par sa fonction de répartition. Ceci revient à dire que l'application qui a une loi donnée associe sa fonction de répartition est injective.

3 Variable aléatoire discrète

3.1 Définition

Définition 13 : Ce sont les variables aléatoires telles que $X(\Omega)$ est une partie finie ou dénombrable de $\mathbb{R}$.

Dans cette partie toutes les variables aléatoires sont discrètes.

Définition 14 : On appelle loi (de probabilité) de la variable aléatoire discrète X la fonction $p_X : X(\Omega) \to \mathbb{R}$, $p_X(x) = P(X = x)$. On appelle fonction de répartition de la variable aléatoire X la fonction F_X définie sur $\mathbb{R}$ par $F_X(t) = P(X \leq t)$.

Remarque 13 : La loi est bien définie car $X^{-1}([x, x])$ appartient à la tribu. On remarque que la fonction p_X permet de définir une probabilité P sur $\mathbb{R}$, $P(A) = \sum_{x \in A \cap X(\Omega)} p_X(x)$.Elle correspond à la définition 11 plus général de la loi d'une variable aléatoire réelle.

Exercice 1 : On jette un dé bleu et un dé vert, modéliser à l'aide d'un espace de probabilité les différents résultats possibles. On s'intéresse à la somme des deux dés, comment peut-on modéliser cette somme à l'aide d'une variable aléatoire X ?

Remarque 14 : $F_X(t) = \sum_{x \in]-\infty;t] \cap X(\Omega)} P(X = x)$ avec une somme au sens de la définition 1.

Proposition 18 : La fonction de répartition F_X d'une variable aléatoire discrète a les propriétés suivantes :

1. F_X est croissante.
2. F_X est continue à droite.
3. $\lim_{-\infty} F_X = 0$
4. $\lim_{+\infty} F_X = 1$
5. Deux V.A.D. ont même loi si et seulement si elles ont même fonction de répartition.

Preuve : La croissance est claire si $t < t'$ alors $(x \leq t) \subset (x \leq t')$.

La continuité à droite est plus délicate, comme F_X est croissante il suffit de montrer que pour tout point a, $\lim F_X(a + \frac{1}{n}) = F_X(a)$, or $F_X(a + \frac{1}{n}) = 1 - P(X > a + \frac{1}{n})$ et la famille $(X > a + \frac{1}{n})_n$ est une suite croissante d'ensembles, on peut donc appliquer la proposition 12 on a :

$$\lim P(X > a + \frac{1}{n}) = P\big(\cup (X > a + \frac{1}{n})\big) = P(X > a)$$

d'où le résultat. Idem pour les limites en $\pm\infty$.

5) Le sens direct est immédiat avec la remarque 14, la réciproque l'est un peu moins, si $F_X = F_Y$, on commence par remarquer que $F_X(a) - F_X(a - \frac{1}{n}) = P(a - \frac{1}{n} < X \leq a)$, comme précédemment à l'aide d'un passage à la limite on obtient que $P(X = a) = F_X(a) - \lim_n F_X(a - \frac{1}{n})$. L'égalité des fonctions de répartitions donne bien l'égalité des lois.

3.2 Espérance

Définition 15 : On suppose que $X(\Omega) = \{x_i | i \in I\}$ où I est soit fini, soit égal à $\mathbb{N}$ (ceci permet de se ramener à des séries plutôt qu'à des familles sommables), les x_i étant distincts deux à deux. Si la série $\sum_I |x_i| P(X = x_i)$ converge, X possède une espérance définie par $E(X) = \sum_{i \in I} x_i P(X = x_i)$. Pour $\alpha \in \mathbb{N}^*$, si la série $\sum |x_i|^\alpha P(X = x_i)$ est convergente, X possède un moment d'ordre $\alpha > 0$ défini par $\sum_i x_i^\alpha P(X = x_i)$. Si X possède un moment d'ordre 2, on définit la variance de X par $\mathrm{var}(X) = E((X - E(X))^2)$.

Remarque 15 : Si I est fini, ce ne sont pas des série mais juste des sommes finis, les remarques du début (prop. 2) sur la sommabilité montre que ces définitions ne dépendent pas de la façon dont on indice l'ensemble $X(\Omega)$.

Remarque 16 : Si Ω est fini ou dénombrable ce qui n'est pas en général le cas pour une variable aléatoire discrète, alors l'espérance de X n'est autre que $\sum_{\omega\in\Omega} X(\omega)P(\{\omega\})$, quitte à se restreindre à ce cas, les différentes propriétés que l'on va voir par la suite sont beaucoup plus simples à démontrer. C'est en un certain sens la bonne notion de l'espérance, que l'on généralise avec la théorie de Lebesgue.

Définition 16 : Une variable aléatoire est dite centrée si elle est d'espérance nulle, et réduite si elle est de variance égale à 1.

Exercice 2 : Soit X une variable aléatoire de loi uniforme sur $\{0,\ldots,n\}$, déterminer son espérance.

Proposition 19 : Si g est une application de $X(\Omega)$ dans $\mathbb{R}$ alors $g(X) = g\circ X$ est une V.A. discrète dont la loi est donnée par

$$\forall y \in g(X(\Omega)), P(g(X)=y) = \sum_{\{x\in X(\Omega)|g(x)=y\}} P(X=x)$$

Preuve : Soient I un intervalle de $\mathbb{R}$, et $F=(g\circ X)^{-1}(I) = X^{-1}(g^{-1}(I)) \subset X^{-1}(g^{-1}(I)\cap X(\Omega))$ or $X(\Omega)$ est fini ou dénombrable, donc F s'écrit comme une réunion dénombrable d'éléments de la tribu,

$$F = \bigcup_{a\in\left(X(\Omega)\cap g^{-1}(I)\right)} X^{-1}(\{a\})$$

c'est donc un élément de la tribu, car chaque $X^{-1}(\{a\})$ est l'image réciproque d'un intervalle par X.
De plus

$$\sum_{\{x|g(x)=y\}} P(X=x) = P\left(\bigcup_{\{x|g(x)=y\}} (X^{-1}(\{x\}))\right) = P(X^{-1}(g^{-1}(\{y\})))$$

Remarque 17 : Le résultat précédent reste vrai si $X_1,\ldots,X_n$ sont n variables aléatoires discrètes et g une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$,

$$P\big(g(X_1,\ldots,X_n)=(y_1,\ldots,y_p)\big) = \sum_{\{(x_1,\ldots,x_n)|g(x_1,\ldots,x_{n)}=(y_1,\ldots,y_p)\}} P(X_1=x_1,\ldots,X_n=x_{n)})$$

Théorème 1 : dit théorème de transfert
Soit X une variable aléatoire discrète, et f une application de $X(\Omega)=\{x_i|i\in I\}$ dans $\mathbb{R}$.
L'application composée $Y=f\circ X$ est une variable aléatoire qui possède une espérance ssi la série $\sum_i f(x_i)P(X=x_i)$ est absolument convergente et dans ce cas son espérance vaut :

$$E(Y) = \sum_{i\in I} f(x_i)P(X=x_i)$$

Preuve : Y prend ses valeurs dans $G=f(X(\Omega))$ qui est dénombrable. Pour $y\in G$ on pose $G_y=\{x\in X(\Omega)|f(x)=y\} = f^{-1}(\{y\})\cap X(\Omega)$ les G_y forment une partition de $X(\Omega)$, et $\forall y,\ P(Y=y)=\sum_{x\in G_y} P(X=x)$ donc d'après la proposition 3 ou 9 :

$$\sum_{y\in f(X(\Omega))} yP(Y=y) = \sum_{y\in f(X(\Omega))}\left(\sum_{x\in G_y} P(X=x)f(x)\right) = \sum_{x\in X(\Omega)} P(X=x)f(x)$$

Proposition 20 : Si X et Y possèdent une espérance, il en est de même pour $X+Y$ et $E(X+Y)=E(X)+E(Y)$, de même $E(\alpha X)=\alpha E(X)$ pour tout réel α et $E(1)=1$.

Preuve : Soit I un intervalle,

$$(X+Y)^{-1}(I) = \{\omega | X(\omega)+Y(\omega) \in I\} = \bigcup_{x\in X(\Omega)} \left(X^{-1}(\{x\}) \cap Y^{-1}(I-x)\right)$$

ou $I - x = \{t \in \mathbb{R} | \exists u \in I,\ t = u - x\}$, c'est donc un intervalle de $\mathbb{R}$, donc $(X+Y)^{-1}(I)$ s'écrit bien comme une réunion dénombrable d'éléments de la tribu, c'est un élément de la tribu
On peut remarquer que pour tout $x \in X(\Omega),\ P(X=x) = \sum_{y\in Y(\Omega)} P(X=x, Y=y)$

$$E(X)+E(Y) = \sum_{x\in X(\Omega)} xP(X=x) + \sum_{y\in Y(\Omega)} yP(Y=Y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} (x+y)P(X=x,Y=y)$$

Or on peut 'découper' $X(\Omega)\times Y(\Omega)$ en une partition définie par les valeurs prises par $X+Y$, c'est une partition finie ou dénombrable. $X(\Omega)\times Y(\Omega) = \coprod_{z\in Z(\Omega)} \{(x,y) \in X(\Omega)\times Y(\Omega) | x+y = z\}$, ce qui donne :

$$E(X)+E(Y) = \sum_{z\in Z(\Omega)} \left(\sum_{\{(x,y)\in X(\Omega)\times Y(\Omega)|x+y=z\}} zP(X=x,Y=y) \right)$$

Or $\sum_{\{(x,y)\in X(\Omega)\times Y(\Omega)|x+y=z\}} P(X=x,Y=y) = P(Z=z)$ d'ou le résultat annoncé.
Exercice 3 : Montrer que $\mathrm{var}(X) = E(X^2) - E(X)^2$.

Remarque 18 : Le théorème 1 reste vrai si $X_1, \ldots, X_n$ sont n variables aléatoires discrètes et g une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$, tq $\sum |g(x_{i_1}, \ldots, x_{i_p})|\ P\big((X_1 = x_{i_1}) \cap \ldots \cap (X_n = x_{i_n})\big)$ converge

$$E(g(X_1,\ldots,X_n)) = \sum_{(i_1,\ldots,i_n)} g(x_{i_1},\ldots,x_{i_n}) P(X_{i_1} = x_{i_1}, \ldots, X_{i_n} = x_{i_n})$$

Exercice 4 : Soit X une variable aléatoire de loi de Poisson de paramètre λ, $\left(\forall k \in \mathbb{N},\ P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}\right)$. Calculer $E(X(X-1))$, en déduire $\mathrm{var}(X)$.

3.3 Fonctions génératrices

<u>**Définition 17**</u> **:** Soit X une variable aléatoire telle que $X(\Omega)$ soit inclus dans $\mathbb{N}$, on pose pour $s \in [0,1],\ g_X(s) = E(s^X) = \sum_k s^k P(X=k)$. g_X est appelée fonction génératrice des moments de X.

Proposition 21 : X possède une espérance ssi g_X est dérivable en 1. Dans ce cas $E(X) = g'_X(1)$.

Preuve : Si X possède une espérance la série $\sum kP(X=k)$ converge, la série entière $\sum s^k P(X=k)$ ainsi que sa 'série dérivée' $\sum ks^{k-1}P(X=k)$ sont alors normalement convergentes sur $[0;1]$. g_X est alors dérivable de dérivée en 1 : $\sum kP(X=k) = E(X)$
Réciproquement : Supposons que g_X soit dérivable en 1 en remarquant que $\sum_{i=0}^{k-1} x^i = \frac{x^k-1}{x-1}$ on a :

$$\frac{g_X(s)-g_X(1)}{s-1} = \frac{1}{s-1}\sum_k (s^k-1)P(X=k) = \sum_k \sum_{0\le i<k} s^i P(X=k)$$

on a donc pour tout entier N :

$$\sum_{n=0}^{N} nP(X=n) = \lim_{s\to 1} \sum_{n=0}^{N} \sum_{0\le i<n} s^i P(X=n) \le \lim_{s\to 1} \sum_{n=0}^{\infty} \sum_{0\le i<n} s^i P(X=n) = g'_X(1)$$

Exercice 5 : Déterminer la fonction génératrice d'une variable aléatoire binomiale de paramètres (n,p), $P(X=k) = C_n^k p^k (1-p)^{n-k}$, en déduire son espérance et sa variance.

Remarque 19 : Une fonction génératrice définit entièrement la loi d'une variable aléatoire à valeur dans $\mathbb{N}$, en effet pour une série entière de rayon de convergence non nul, si $f(x) = \sum a_n x^n$, alors $a_n = \frac{1}{n!} f^{(n)}(0)$.

3.4 Lois usuelles

Uniforme, Bernoulli, binomiale, géométrique, Poisson.

4 Variable aléatoire à densité

Définition 18 : $f : \mathbb{R} \to \mathbb{R}$ est une densité si $f \geq 0$, intégrable sur $\mathbb{R}$ et $\int_{\mathbb{R}} f(t)\,\mathrm{d}t = 1$. Une variable aléatoire X admet f pour densité si pour tout intervalle I de $\mathbb{R}$, $P(X \in I) = \int_I f(t)\,\mathrm{d}t$

Exercice 6 : Soit X une variable aléatoire admettant pour densité

$$f_X(t) = \begin{cases} \alpha t^2 \text{ si } t \in [0;1] \\ 0 \text{ sinon} \end{cases}$$

Calculer α, puis $P(X \geq \frac{1}{2})$. Déterminer une autre densité pour X.

Dans cette partie la variable aléatoire X possède la densité f_X

Définition 19 : On appelle fonction de répartition de X la fonction définie sur $\mathbb{R}$ par

$$F_X(t) = P(X \leq t) = \int_{]-\infty;t]} f_X(t)\,\mathrm{d}t$$

. c'est la même définition que pour les variables aléatoires discrètes (déf. 14)

Proposition 22 : Soit X une variable aléatoire admettant f_X pour densité, on note F_X sa fonction de répartition, elle a les propriétés suivantes :

1. Pour tout $a < b$, $P(a < X \leq b) = F_X(b) - F_X(a)$.
2. $\lim_{-\infty} F_X = 0$ et $\lim_{+\infty} F_X = 1$.
3. F_X est croissante, continue, et $\forall x\ P(X = x) = 0$.
4. Si f_X est continue en x_0 alors F_X est dérivable en x_0 et $F'(x_0) = x_0$.
5. Si F_X est une fonction $\mathcal{C}^1$ alors X admet F'_X comme densité. (Ce résultat peut se généraliser à F_X est une fonction $\mathcal{C}^0$, et $\mathcal{C}^1$ sur $\mathbb{R}$ privé d'un nombre fini de points).
6. (admis) deux variables aléatoires à densité ayant même fonction de répartition, ont les même densités.

Exercice 7 : suite de l'exercice 6
Montrer que la variable aléatoire $\sqrt{X}$, possède une densité que l'on déterminera. **Preuve :** ...

4.1 Espérance

Définition 20 : Si la fonction définie sur $\mathbb{R}$ par $x \to x f_X(x)$ est intégrable sur $\mathbb{R}$ alors X possède une espérance définie par $E(X) = \int_{\mathbb{R}} x f_X(x)\,\mathrm{d}x$
Si la fonction définie sur $\mathbb{R}$ par $x \to x^k f_X(x)$ est intégrable sur $\mathbb{R}$ alors X possède un moment d'ordre k défini par $\int_{\mathbb{R}} x^k f_X(x)\,\mathrm{d}x$, de même pour la variance : var$X = E\big((X - E(X))^2\big)$.

Proposition 23 : (admis) Si X et Y sont deux variables aléatoires à densité ayant une espérance alors $X + Y$ et αX sont deux variables aléatoires qui ont une espérance, et $E(X + Y) = E(X) + E(Y)$, et $E(\alpha X) = \alpha E(X)$.

Remarque 20 : La propriété 23 qui est au programme de l'agrégation interne est un peu étrange car $X + Y$ n'est pas forcément une variable aléatoire à densité ou une variable aléatoire discrète, et dans ce cas l'espérance (au sens de l'agrégation) n'est pas définie, bien sur lorsqu'on connaît l'intégrale de Lebesgue il n'y a plus de problème. Comme cette propriété est très importante on comprend toutefois qu'elle soit au programme.

Exercice 8 : Donner un exemple très simple où X et Y sont à densité mais pas $X+Y$.
Montrer que var$(X) = E(X^2) - E(X)^2$.

Exercice 9 : Soit X une variable aléatoire exponentielle de paramètre λ :

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t} \text{ si } t \geq 0 \\ 0 \text{ sinon} \end{cases}$$

Déterminer l'espérance de X, la probabilité $P(-1 < X < 1)$ ainsi que $P(X = 1)$.

Proposition 24 : (admis) Soit g une application de $\mathbb{R}$ dans $\mathbb{R}$ continue par morceaux, telle que $x \to g(x)f_X(x)$ soit intégrable sur $\mathbb{R}$, alors $g(X)$ est une variable aléatoire dont l'espérance est donnée par $\int_{\mathbb{R}} g(x)f_X(x)\,\mathrm{d}x$.

Remarque 21 : On peut refaire la remarque 20. On peut facilement démontrer ce résultat pour une fonction g continue bijective.

Exercice 10 : Dans l'exercice 6, calculer de deux façons différentes $E(\sqrt{X})$.

4.2 Loi normale

<u>Définition 21</u> : On appelle variable aléatoire normale ou Gaussienne de paramètre (m, σ^2), une variable aléatoire admettant f_X comme densité :

$$f_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-m)^2}{2\sigma^2}}$$

Remarque 22 : Il n'existe pas de primitive de f_X qui s'écrive à l'aide des fonctions usuelles.

Proposition 25 : Soit $(a, b) \in \mathbb{R}^* \times \mathbb{R}$, si X est une variable aléatoire normale de paramètres (m, σ^2) alors $aX + b$ suit une loi normale de paramètre $(am + b, a^2\sigma^2)$.

Preuve : Il suffit de regarder la fonction de répartition puis de dériver.

Proposition 26 : Soit X est une variable aléatoire normale de paramètres (m, σ^2), on a $E(X) = m$ et var$(X) = \sigma^2$.

Preuve : Calculs.

5 Vecteurs aléatoires

<u>Définition 22</u> : Une application de Ω dans $\mathbb{R}^p$ est un vecteur aléatoire si chacune de ses composantes est une variable aléatoire.

5.1 Vecteurs aléatoires discrets

<u>Définition 23</u> : Un vecteur aléatoire est discret si chacune de ses composantes est une variable aléatoire discrète.

<u>Définition 24</u> : La loi d'un vecteur aléatoire discret $(X_1, \ldots, X_n)$ est la fonction p de $X_1(\Omega) \times \ldots \times X_n(\Omega)$ dans $\mathbb{R}$ telle que $p(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n)$.

<u>Définition 25</u> : $X_1, \ldots, X_n$ sont indépendantes si quelque soit $(x_1, \ldots, x_n)$, les événements $(X_i = x_i)$ sont indépendants c'est à dire :

$$P\left(\bigcap_{i=1}^{n}(X_i = x_i)\right) = \prod_{i=1}^{n} P(X_i = x_i)$$

Proposition 27 : Si X et Y sont des variables aléatoires indépendantes ayant une espérance alors XY a une espérance et $E(XY) = E(X)E(Y)$.

Preuve : $E(XY) = \sum_{z\in(XY)(\Omega)} zP(XY = z) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x, Y = y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x)P(Y = y) = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} xyP(X = x)P(Y = y) = E(X)E(Y)$

Exercice 11 : Montrer que si X et Y possèdent des variances et sont indépendantes alors $\text{var}(X + Y) = \text{var}X + \text{var}Y$

Définition 26 : Soient X et Y deux variables aléatoires, on définit la covariance de X et Y lorsqu'elle existe par $\text{cov}(X, Y) = E\big((X - E(X))(Y - E(Y)\big)$, et on définit le coefficient de corrélation $\rho_{X,Y}$ lorsque X et Y ont des variances non nuls par

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}X \ \text{var}Y}}$$

Proposition 28 : $\text{cov}(X, Y) = E(XY)-E(X)E(Y)$, si X et Y sont indépendantes alors $\text{cov}(X, Y) = 0$. La réciproque est fausse.

Exercice 12 : X suit une loi de Poisson de paramètre λ, et on définit la loi de Y à l'aide d'une probabilité conditionnelle : $P(Y = k|X = n) = C_n^k p^k(1 - p)^{n-k}$. Calculer l'espérance de XY puis $\text{cov}(X, Y)$, déterminer la loi de Y.

Proposition 29 : La variance d'une somme de variables aléatoires indépendantes ayant une variance, possède une variance égale à la somme des variances des différentes variables aléatoires.

Proposition 30 : Soient $f_1, \ldots, f_n : \mathbb{R} \mapsto \mathbb{R}$ des fonctions et $X_1, \ldots, X_n$ des variables aléatoires discrètes indépendantes alors les variables aléatoires $f_1 \circ X_1, \ldots, f_n \circ X_n$ sont indépendantes.

Proposition 31 : Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes à valeur dans $\mathbb{N}$, alors la somme $S_n = X_1 + \ldots + X_n$ a pour fonction génératrice le produit des fonctions génératrices :

$$g_{S_n}(t) = \prod_{k=1}^{n} g_{X_k}(t)$$

Preuve : $g_{S_n}(t) = E(t^{S_n}) = E(t^{X_1+\ldots+X_n}) = E(t^{X_1} \ldots t^{X_n}) = \prod_{k=1}^n E(t^{X_1}) \ldots E(t^{X_n}) = \prod_{k=1}^n g_{X_k}(t)$

5.2 Vecteurs aléatoires à densité

Définition 27 : Une fonction de $\mathbb{R}^p$ dans $\mathbb{R}^+$ est une densité sur $\mathbb{R}^p$ si elle est intégrable sur $\mathbb{R}^p$ d'intégrale égale à 1. Un vecteur aléatoire possède la densité f si pour tout intervalle $I_1, \ldots, I_p$,

$$P\big((X_1 \in I_1) \cap \ldots \cap (X_p \in I_p)\big) = \int_{I_1} \ldots \int_{I_p} f(x_1, \ldots, x_p) \ \mathrm{d}x_1 \ldots \mathrm{d}x_p$$

Proposition 32 : (admis) Si A est un domaine simple de $\mathbb{R}^p$ et f une densité du vecteur $(X_1, \ldots, X_p)$, alors $P(X \in A) = \int_A f(x_1, \ldots, x_p) \ \mathrm{d}x_1 \ldots \ \mathrm{d}x_p$.

Exercice 13 : Soit (X, Y) un couple de V.A. ayant f pour densité, et $\Delta = \{(x, x)|x \in \mathbb{R}\}$, montrer que $P((X, Y) \in \Delta) = 0$.

Définition 28 : On appelle indicatrice d'un sous-ensemble A d'un ensemble E, la fonction valant 1 sur A et 0 sur son complémentaire.

Proposition 33 : (admis) Comme dans la remarque 18 si g est le produit d'une fonction continue de $\mathbb{R}^p$ dans $\mathbb{R}$ et d'une indicatrice d'un domaine simple de $\mathbb{R}^p$ et sous réserve d'intégrabilité, on a

$$E(g(X_1, \ldots, X_p)) = \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} g(x_1, \ldots, x_p)f(x_1, \ldots, x_p) \ \mathrm{d}x_1 \ldots \mathrm{d}x_p$$

Définition 29 : Soient $X_1, \dots, X_p$ des variables aléatoires à densité, elles sont indépendantes si pour tous intervalles $I_1, \dots, I_p$, $P(\cap(X_i \in I_i)) = \prod P(X_i \in I_i)$

Proposition 34 : Si $X_1, \dots, X_p$ sont des variables aléatoires à densité, elles sont indépendantes ssi la fonction définie par $f(x_1, \dots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_p}(x_p)$ est une densité du vecteur $(X_1, \dots, X_p)$.

Preuve : Supposons que ce soit une densité du vecteur, alors $P(\cap(A_i \in I_i)) =$
$\int_{I_1} \dots \int_{I_p} f(x_1, \dots, x_p)\ \mathrm{d}x_1 \dots \mathrm{d}x_p = \int_{I_1} \dots \int_{I_p} f_{X_1}(x_1) f_{X_2}(x_2) \dots, f_{X_p}(x_p)\ \mathrm{d}x_1 \dots \mathrm{d}x_p$ et d'après le théorème de Fubini, ceci est égale au produit des intégrales $\int_{I_k} f_{X_k}(x_k)\ \mathrm{d}x_k$, d'ou l'indépendance.
réciproquement, on applique encore Fubini.

Définition 30 : Soient X et Y deux variables aléatoires, on définit la covariance de X et Y lorsqu'elle existe par $\mathrm{cov}(X, Y) = E\big((X - E(X))(Y - E(Y))\big)$, et on définit le coefficient de corrélation $\rho_{X,Y}$ lorsque X et Y ont des variances non nuls par

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}X\ \mathrm{var}Y}}$$

Proposition 35 : $\mathrm{cov}(X, Y) = E(XY) - E(X)E(Y)$
si X et Y sont indépendantes alors
$\mathrm{cov}(X, Y) = 0$. La réciproque est fausse.

Exercice 14 : Soit A le domaine du carré unité défini par

$$A = \{(x, y) \in [0, 1]^2 : x \geq y + \frac{1}{2} \text{ ou } x \leq y \leq x + \frac{1}{2}\}.$$

a) Déterminer $\int\int 1_A(x, y)\, \mathrm{d}x\, \mathrm{d}y$
b) Soit (X, Y) un couple de v.a. de densité uniforme sur A. Déterminer les lois marginales de X et de Y.
c) Calculer $P(X + Y \leq 1)$
d) Calculer $Cov(X, Y)$.

Proposition 36 : Soient X et Y deux variables aléatoires indépendantes admettant f_X et f_Y comme densité. Soient Φ et Ψ des fonctions continues par morceaux telles que $\Phi(X)$ et $\Psi(Y)$ aient une espérance alors $\Phi(X)\Psi(Y)$ possède une espérance égale à $E(\Phi(X))E(\Psi(Y))$

Preuve : Fubini

Remarque 23 : En particulier, toujours dans le cas de VA indépendantes, si $\Phi = \Psi = Id$, on a $E(XY) = E(X)E(Y)$ et $\mathrm{cov}(X, Y) = 0$.

Proposition 37 : Soient X et Y deux variables aléatoires indépendantes de lois normales, $X + Y$ est encore une variable aléatoire normale.

Preuve : Il suffit de regarder le cas ou X et Y sont de paramètre $(0, \frac{1}{2})$.

Déterminons la fonction de répartition de $X+Y$: F

$$\begin{aligned}
F(t)=P(X+Y\leq t) &= P((X,Y)\in\{(x,y)|x+y\leq t\})\\
&= \int\int_{\{(x,y)|x+y\leq t\}} f_{(X,Y)}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\
&= \int\int_{\{(x,y)|x+y\leq t\}} f_X(x)f_Y(y)\,\mathrm{d}x\,\mathrm{d}y\\
&= \int_{[-\infty;t]}\int_{\mathbb{R}} f_X(u-y)f_Y(y)\,\mathrm{d}y\,\mathrm{d}u\\
&= \frac{1}{\pi}\int_{[-\infty;t]}\int_{\mathbb{R}} e^{-(u-y)^2-y^2}\,\mathrm{d}y\,\mathrm{d}u\\
&= \frac{1}{\pi}\int_{[-\infty;t]} e^{-\frac{1}{2}u^2}\int_{\mathbb{R}} e^{-2(y-\frac{1}{2}u)^2}\,\mathrm{d}y\,\mathrm{d}u\\
&= \frac{1}{\pi}\int_{[-\infty;t]} e^{-\frac{1}{2}u^2}\int_{\mathbb{R}} e^{-2(v)^2}\,\mathrm{d}v\,\mathrm{d}u\\
&= \frac{1}{\pi}\int_{[-\infty;t]} e^{-\frac{1}{2}u^2}\frac{1}{2}\sqrt{2\pi}\,\mathrm{d}u\\
&= \frac{1}{\sqrt{2\pi}}\int_{[-\infty;t]} e^{-\frac{1}{2}u^2}\,\mathrm{d}u
\end{aligned}$$

Donc la fonction de répartition est dérivable, et sa dérivée est la densité d'une loi normale de paramètre $(0,1)$. Donc $X+Y$ suit une loi normale centrée réduite.

Remarque 24 : On peut trouver une formule du même genre dans le cas général, pour deux variables aléatoires à densité, indépendantes, leur somme admet pour densité le produit de convolution de deux densités :

$$f_X * f_Y(t) = \int_{\mathbb{R}} f_X(t-x)f_Y(x)\;dx$$

6 Théorèmes limites

Proposition 38 : Inégalité de Bienaymé Tchebychev
Soit X une variable aléatoire discrète ou à densité ayant une espérance m et une variance σ^2. On a :

$$\forall\epsilon>0,\;P(|X-m|\geq\epsilon)\leq\frac{\sigma^2}{\epsilon^2}$$

Remarque 25 : Cette inégalité permet de contrôler, l'éloignement de la variable aléatoire X par rapport à sa valeur moyenne m.

Preuve : Dans le cas de VA à densité.

$$\begin{aligned}
\mathrm{var}(X) &= \int_{\mathbb{R}}(x-m)^2 f_X(x)\;dx\\
&\geq \int_{\mathbb{R}\setminus[m-\epsilon,m+\epsilon]}(x-m)^2 f_X(x)\;dx\\
&\geq \int_{\mathbb{R}\setminus[m-\epsilon,m+\epsilon]}\epsilon^2 f_X(x)\;dx\\
&\geq \epsilon^2 P(X\notin[m-\epsilon,m+\epsilon])
\end{aligned}$$

D'ou le résultat, dans le cas des VA à densité.

Théorème 2 : Loi faible des grands nombres

Soit (X_n) une suites de variables aléatoires indépendantes et de même loi (discrète ou à densité) ayant pour espérance m et pour variance σ^2. On a :

$$\forall \epsilon > 0, \lim_{n\to\infty} P(|\frac{1}{n}\sum_{k=1}^{n} X_k - m| > \epsilon) = 0$$

Remarque 26 : Cela est une façon de dire que la variable aléatoire définie comme la moyenne arithmétique de n variables aléatoires indépendantes de même loi, est très proche pour n grand de l'espérance (c'est à dire une sorte de moyenne théorique) de chacune de ces variables aléatoires.

Preuve : Il suffit d'appliquer l'inégalité de Bienaymé Tchebychev, à la variable aléatoire $\frac{1}{n}\sum_{k=1}^{n} X_k$ qui a pour espérance m et pour variance $\frac{\sigma^2}{n}$.

Théorème 3 : Loi forte des grands nombres (admis)

Soit (X_n) une suites de variables aléatoires indépendantes et de même loi (discrète ou à densité) ayant pour espérance m. alors l'ensemble

$$A = \{\omega \in \Omega | \lim \frac{X_1(\omega) + X_2(\omega) + \ldots + X_n(\omega)}{n} = m\}$$

appartient à la tribu de Ω, de plus sa probabilité est égale à 1.

Théorème 4 : Théorème de la limite centrale (admis)

Soient (X_n) une suites de variables aléatoires indépendantes, de même loi (discrète ou à densité) ayant pour espérance m, pour variance σ^2, Y une variable aléatoire de loi normale centrée réduite, et

$$Z_n = \left(\frac{\frac{X_1+X_2+\ldots+X_n}{n} - m}{\sigma}\right)\sqrt{n} = \frac{X_1 + X_2 + \ldots + X_n - nm}{\sigma\sqrt{n}}$$

pour tout intervalle I on a :

$$\lim_{n\to\infty} P(Z_n \in I) = P(Y \in I)$$

7 Exercices

7.1 Modélisation

Exercice 1 : On dispose de 3 urnes et de 3 boules. On répartit au hasard les boules dans les urnes. Modéliser les cas suivants à l'aide d'un ensemble Ω et d'une probabilité P sur Ω :
1) Les urnes et les boules sont distinguables (discernables).
2) Les urnes sont distinguables mais non les boules.
3) Les boules sont distinguables mais non les urnes.
4) Les boules et les urnes sont indistinguables.

Exercice 2 : On constitue une file d'attente en attribuant au hasard des numéros d'ordre à n personnes. Quelle est la probabilité que deux amis soient distants de r places (c'est à dire séparés par $r-1$ personnes). On résoudra l'exercice en utilisant trois modélisations différentes :
a) $\Omega = \{(k_1, k_2) | k_1 \neq k_2\}$
b) $\Omega = \{\{k_1, k_2\} | k_1 \neq k_2\}$
c) $\Omega = \{(k_1, k_2, ..., k_n) | \forall i, j,\ i \neq j \Longrightarrow k_i \neq k_j\}$

Exercice 3 : On jette un dé, et on veut modéliser le nombre de jets nécessaire pour qu'un 6 apparaisse
a) à l'aide d'une probabilité sur $\mathbb{N}^*$.
b) A l'aide d'une variable aléatoire N.
c) Calculer $E(N)$, interprétation.

Exercice 4 : Montrer qu'il n'existe pas de probabilité uniforme sur $\mathbb{N}$.

7.2 Indépendance

Exercice 5 : (Foata-Fuchcs p66)
Soient (ω, P) un espace de probabilité, et A et B deux événements de cet espace, notons α, β, γ, et δ les probabilités des événements $A \cap B$, $A \cap \overline{B}$, $\overline{A} \cap B$, et $\overline{A} \cap \overline{B}$.
a) Calculer $\alpha + \beta + \gamma + \delta$.
b) Montrer que $P(A \cap B) - P(A)P(B) = \alpha\delta - \beta\gamma$
c) Montrer que $|P(A \cap B) - P(A)P(B)| \leq \frac{1}{4}$.
d) Donner des exemples d'égalités.

Exercice 6 : (Foata-Fuchs p59 remarque 1)
On lance deux dés, on note A l'événement le premier dé donne un nombre pair, B l'événement le deuxième dé donne un nombre impair, C l'événement la parité des deux dés est la même . Modéliser l'expérience aléatoire à l'aide d'une mesure de probabilité. Discuter les indépendances de A et B, de A et C, de B et C et de A, B et C.

Exercice 7 : (Foata-Fuchs 8p63) On fait l'hypothèse que dans une famille les sexes des enfants sont indépendants les uns des autres et que chaque enfant a la probabilité $\frac{1}{2}$ d'être un garçon et la probabilité $\frac{1}{2}$ d'être une fille.
a) Une famille a deux enfants dont l'un au moins est un garçon . Quelle est la probabilité pour que les deux enfants soient des garçons ?
b) Une famille a deux enfants. L'aîné est un garçon, quelle est la probabilité pour que les deux enfants soient des garçons ?

Exercice 8 : (Il y en partout par ex Précis)HEC Proba stat Belin (Degrave)
Un dépistage systématique est effectué sur une population dont 0.1% des individus présentent une certaine affection A non apparente. Ce dépistage est débuté par un test qui donne 95% de résultats positifs pour les personnes atteintes par A, et 1% de résultat positifs pour les personnes non atteintes.

Quelle est la probabilité conditionnelle qu'une personne prise au hasard soit atteinte par A sachant que le test a donné un résultat positif ? soit indemne sachant que le test a donné un résultat négatif ?

Exercice 9 : Dress, corrigé peu convainquant p 29
Devant un certain tableau clinique, on estime que 6 personnes sur 10 sont atteintes d'une maladie M. On effectue alors deux tests biologiques dont les résultats sont indépendants (toute la difficulté réside dans la modélisation de cette hypothèse). Les deux tests donnent 95% de résultats positifs pour les personnes atteintes par M, et 10% de résultat positifs pour les personnes non atteintes.
a) Quelle est la probabilité que la personne soit malade si les deux tests sont positifs ?
b) Quelle est la probabilité que le deuxième test soit positif si le premier.

Exercice 10 : Harari p33 non corrigé
On s'intéresse à la transmission d'une information binaire, c'est à dire ne pouvant prendre que deux valeurs. On admet que le procédé de transmission directe entre deux individus A et B est tel que, lorsque A émet une valeur de l'information à destination de B, ce dernier reçoit la valeur émise par A avec la probabilité p, et donc l'autre valeur avec la probabilité $q = 1 - p$, on suppose que $0 < p < 1$. On considère des individus successifs $i_0, i_1, ..., i_n$ avec $n \in \mathbb{N}$. L'information émise par i_0 est transmise à i_1, qui transmet la valeur reçue à i_2, et ainsi de suite jusqu'à i_n.
Entre deux individus, i_k et i_{k+1}, la transmission de l'information suit la loi décrite plus haut. On note p_k la probabilité que la valeur de l'information reçue par i_k soit identique à celle émise par i_0, et on pose $p_0 = 1$.
a) Exprimer p_{k+1} en fonction de p_k.
b) On rappelle que pour étudier une suite arithmético-géométrique du genre $u_{n+1} = au_n + b$ on pose $v_n = u_n + \alpha$ avec α tel que (v_n) soit une suite géométrique. En déduire une expression de p_n en fonction de n et de p.
c) Déterminer $\lim_{k\to\infty} p_k$
d) Déterminer un p tel que $p_{100} > 99,9\%$

Exercice 11 : (Degrave Précis) En fait c'est un exemple très simple de chaîne de Markov
Trois enfants A,B et C jouent avec une balle. Lorsque A a la balle, la probabilité qu'il l'envoie à C est 0,25. Lorsque B a la balle, il l'en voie respectivement à A et a C avec les probabilités 0,75 et 0,25. C envoie toujours la balle à B. On désigne par A_n (resp. B_n, C_n)les probabilités pour qu'à l'issue du nième lancer, ce soit A (resp. b et C) qui ait la balle.
a) Montrer qu'il existe une matrice carrée M telle que

$$\begin{pmatrix} A_{n+1} \\ B_{n+1} \\ C_{n+1} \end{pmatrix} = M \begin{pmatrix} A_n \\ B_n \\ C_n \end{pmatrix}$$

b) Diagonaliser M et calculer M^n pour $n \in \mathbb{N}$.
c) Calculer les limites lorsque n tend vers l'infini des probabilités A_n, B_n et C_n. On vérifiera que ces limites sont indépendantes de l'enfant qui avait la balle au début du jeu.

7.3 Variables aléatoires discrètes

Exercice 12 : Soit X une variable aléatoire de Poisson de paramètre λ :

$$\forall n \in \mathbb{N},\ P(X = n) = e^{-\lambda}\frac{\lambda^n}{n!}$$

a) Déterminer son espérance, à l'aide de la définition puis de la fonction génératrice.
b) Déterminer sa variance, à l'aide de la définition puis de la fonction génératrice.

Exercice 13 : Soit T une v.a. entière définie sur $(\Omega, \mathcal{A}, P)$. On suppose que $\forall n \in N$, $P(T > n) \neq 0$ et $\forall (p, n) \in N^2$, $P(T \geq n + p | T \geq n) = P(T \geq p)$.

a) Montrer que T suit une loi géométrique (Ce sont les variables aléatoires de la forme $P(T=n)=a^n(1-a)$).
b) Donner un exemple d'expérience modélisée par une telle loi.
c) Déterminer la fonction génératrice d'une variable aléatoire géométrique, en déduire son espérance et sa variance.

Exercice 14 : (Précis 111 p 104)
On joue à pile ou face avec une pièce non équilibrée. A chaque lancer la probabilité d'obtenir pile est égale à $\frac{2}{3}$ tandis que celle d'obtenir face est égale à $\frac{1}{3}$.Les lancers sont supposés indépendants. On note X une variable aléatoire modélisant le nombre de lancers nécessaires pour l'obtention, pour la première fois, de 2 piles consécutifs. soit $n \in \mathbb{N}^*$, on note p_n la probabilité de l'événement $(X=n)$.
a) Expliciter les événements $(X=2),(X=3),(X=4)$. déterminer p_1,p_2,p_3,p_4.
b) Á l'aide de la formule des probabilités totales, en distinguant deux cas selon le résultat du premier lancer, montrer que :

$$\forall n \geq 3;\; p_n = p_n = \frac{2}{9}p_{n-2} + \frac{1}{3}p_{n-1}$$

c) En déduire l'expression de p_n en fonction de n.
d) Déterminer $E(X)$.

Exercice 15 : Soient $X_1,\dots X_n$ des variables aléatoires indépendantes de même lois de Bernoulli de paramètre p. Quelle est la loi de la somme $S=\sum_{k=1}^n X_k$, quelle est l'espérance de cette somme et sa variance ? Montrer qu'il n'existe pas de variables aléatoires Y indépendantes des X_i telle que $S-Y$ suivent une loi binomiale de paramètre $(n-1,p)$. Que pensez vous de la question si on ne suppose plus l'indépendance ?

Exercice 16 : Soient $X_1,\dots X_n,\dots$ des variables aléatoires indépendantes de lois de Bernoulli de paramètre p_n telle que $\lim np_n=\lambda>0$.
a) Quelle est la limite de $P(X_n=k)$ lorsque n tend vers l'infini ? Interprétation.
b) Application : La proportion de maisons touchées par la foudre une année donnée est $0,001\%$, quelle est la probabilité qu'une assurance assurant 100 000 maisons, ait moins de 3 maisons touchées par la foudre, on précisera les hypothèses faites et on les critiquera.

Exercice 17 : Précis p 106
Soient X et Y des variables aléatoires indépendantes de même loi uniforme sur $\{1,\dots,n\}$. Calculer $P(X=Y),P(x\geq Y)$, déterminer la loi de $X-Y$. Déterminer la covariance de $X-Y$ et $X+Y$.

7.4 Variable aléatoires à densité

Exercice 18 : Soit X une variable aléatoire, de fonction de répartition F_X.
a) Montrer que si X est une variable aléatoire à densité alors F_X est continue.
b) Montrer que si F_X est continue alors $P(X=a)=0$ pour tout réel a.

Exercice 19 : partout
Soit X une variable aléatoire uniforme sur $[a;b]$, déterminer son espérance et sa variance.

Exercice 20 : Soit X une variable aléatoire de densité $f_X(t)=\alpha t 1_{[0;\beta]}(t)$, avec $\beta >$.
1) Déterminer une relation entre α et β.
2) Représenter f_X et la fonction de répartition de X : F_X.
3) Déterminer α et β pour que X ait une espérance égale à 2.

Exercice 21 : en partie traité dans précis
Soit X une variable aléatoire uniforme sur $]-1;1[$, déterminer la loi et l'espérance de $U=|X|$, $V=X^2$, et $W=\frac{1}{2}\ln\frac{1+X}{1-X}$.

Exercice 22 : Une variable aléatoire est dite exponentielle de paramètre λ si elle admet pour densité :

$$f_X(x)=\lambda e^{-\lambda x}1_{\mathbb{R}^+}(x)$$

Soient X et Y deux variables aléatoires exponentielles indépendantes de même paramètre, déterminer la loi de $X+Y$.

Exercice 23 : Précis p 210
Soit Z une variable aléatoire de densité $f_X(t)=\alpha e^{-|t|}$.
a) Déterminer α.
b) Déterminer la loi de la variable aléatoire $Z=|T|$.
c) Soient X et Y deux variables aléatoires indépendantes de loi exponentielle de paramètre 1, déterminer la lois de $X-Y$.

Exercice 24 : partout Gaultier p 51
Une variable aléatoire est dite de Cauchy de paramètre $a>0$ si elle admet pour densité :

$$f_X(x)=\frac{a}{\pi}\frac{1}{a^2+x^2}$$

a) Montrer que f_X est bien une densité.
b) Montrer qu'une variable aléatoire de Cauchy ne possède pas d'espérance.
c) Si X est une variable aléatoire de Cauchy de paramètre a quelle est la loi de $2X$.

Exercice 25 : Soient X_1 et X_2 deux v.a. indépendantes de même loi uniforme sur $[0,1]$. Déterminer les lois des variables aléatoires $U=\max\{X_1,X_2\}$ et $V=\min\{X_1,X_2\}$. Que vaut $E(\mid X_1-X_2\mid)$?

Exercice 26 : Soit A le domaine du carré unité défini par

$$A=\{(x,y)\in[0,1]^2 : x\geq y+\frac{1}{2}\text{ ou } x\leq y\leq x+\frac{1}{2}\}.$$

a) Déterminer $\int\int 1_A(x,y)\,dxdy$
b) Soit (X,Y) un couple de v.a. de densité uniforme sur A. Déterminer les lois marginales de X et de Y.
c) Calculer $Cov(X,Y)$.

Exercice 27 : Série Schaum p 57
La densité de probabilité jointe de (X,Y) est $f_{(X,Y)}$:

$$f_{(X,Y)}(x,y)=cxy\text{ si }(x,y)\in[0;4]\times[1,5]\text{ sinon }0$$

Évaluer c, calculer $P(X\geq 3,Y\leq 2)$, déterminer la loi de X, calculer la covariance de (X,Y).
X et Y sont-elles indépendantes ?

7.5 Loi normale

Exercice 28 : Précis exemple 3 p 189 et p 202
Soient X une variable aléatoire admettant pour densité une fonction f continue sauf en un nombre fini de points, et $(a,b)\in\mathbb{R}^*\times\mathbb{R}$, on pose $Y=aX+b$, montrer que Y est à densité et déterminer sa densité.
Application : Montrer que si X suit une loi normale de paramètre $(m;\sigma^2)$ alors $X^*=\frac{1}{\sigma}(X-m)$ suit une loi normale de paramètre $(0;1)$.

Exercice 29 : soient X et Y deux variables aléatoires indépendantes, ayant des densités f_X et f_Y continues sur $\mathbb{R}$.
a) Déterminer la fonction de répartition de la variable aléatoire $S=X+Y$
b) On suppose de plus que f_X est bornée, montrer que S est une variable aléatoire à densité, on donnera une forme intégrale de cette densité.
c) Si X et Y deux variables aléatoires indépendantes de même loi normale centrée réduite, déterminer la loi de $X+Y$.
d) Même question si les paramètres de X et Y sont (m,σ^2) et (m',σ'^2).

7.6 Théorèmes limites

Exercice 30 : Convergence en loi, de la loi hypergéométrique vers la loi binomiale.
Soit (X_k) une suite de variables aléatoires hypergéomètriques de paramètres (n, M_k, N_k) telle que

$$\lim_{k\to+\infty} M_k = +\infty \text{ et } \lim_{k\to+\infty} \frac{M_k}{N_k + M_k} = p \in]0;1[$$

montrer que :

$$\lim_{k\to\infty} P(X_k = l) = C_n^l p^l (1-p)^{n-l}$$

interpréter le résultat.

Exercice 31 : Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes de loi uniforme sur $[0, A]$. Il faut avoir en tête que l'on ne connaît pas A et qu'on cherche à le déterminer. On pose $M = \frac{2}{n}(X_1 + \ldots + X_n)$ et $S = max(X_1, \ldots, X_n)$.

1. Déterminer l'espérance et la variance de M.
2. Déterminer la loi et l'espérance de S.
3. On choisit α tel que $T = \alpha S$ soit d'espérance A, déterminer la variance de T.
4. Sur un site de fouille, on retrouve des vis de différentes longueurs, tout laisse supposer que les habitants du lieu, utilisaient un outil permettant de fabriquer des vis de longueurs variables de 1 à A cm, et qu'ils fabriquaient ces vis sans qu'il y ait de longueurs privilégiées. On retrouve des vis de longueurs 3,1 ;5,3 ;1,6 et 5,7 cm, comment peut-on estimer A ?

Exercice 32 : de ce genre partout Précis p 250
Une entreprise compte 300 employés, chacun d'eux téléphone en moyenne 6 minutes par heures. Quel est le nombre de lignes que l'entreprise doit installer pour que la probabilité que toutes les lignes soient utilisées au même instant soit au plus égale à 0,025.

Exercice 33 : Précis p 251
Soit (X_n) une suite de variables aléatoires à valeur dans $\mathbb{N}$ telle que pour tout $k \in \{1, \ldots, n\}$,

$$P(X_n \leq k) = \frac{k^2}{n^3}(3n - 2k)$$

On pose $Y_n = \frac{X_n}{n}$.
a) Soit x un réel donné. Déterminer la limite de la suite $\frac{[nx]^2}{n}$.
b) Montrer que pour tout $x \in [0;1]$:

$$P(X_n \leq k) = \frac{[nx]^2}{n^3}(3n - 2[nx])$$

c) en déduire la limite de $(P(Y_n \leq x)$ pour tout $x \in \mathbb{R}$.

Bibliographie :

- Pour les exercices
  - Foata-Fuchs, calcul des probabilités, DUNOD édition 2
  - Degrave, Précis de Mathématiques HEC Probabilités et statistiques, Bréal
  - Harari Personnaz, cours de mathématiques 3.Probabiliés, HEC Belin
  - Dress, Probabilités statistique Deug sciences, Dunod
  - Gaultier, Probabilités 70 exercices corrigés, Vuibert
  - Murray R. Spiegel, Probabilité et statistique, série Schaum
- Pour le cours, en plus des précédents et particulièrement le premier
  - Ouvrard, Probabilités 1, Cassini
  - Lannuzel, Probabilités et statistiques, Masson

UNIVERSITÉ de Cergy-Pontoise

Problème de probabilité

8 février 2006

Soit X une variable aléatoire à densité de densité f_X, on appelle fonction caractéristique de X la fonction φ définie par :

$$\forall t \in \mathbb{R},\ \varphi(t) = E(e^{itX})$$

1. Généralités
   (a) Montrer que φ est une fonction définie et continue sur $\mathbb{R}$.
   (b) Déterminer $\varphi(0)$, et montrer que $|\varphi|$ est majorée par $|\varphi(0)|$.
   (c) Soient $a \neq 0$ et b des réels, déterminer la fonction caractéristique de la variable aléatoire $Y = aX + b$, φ_Y en fonction de celle de X.
   (d) Montrer que si X admet une espérance alors φ est une fonction de classe $\mathcal{C}^1$, et $E(X) = \frac{1}{i}\varphi'(0)$.
   (e) Montrer que si X admet un moment d'ordre k alors φ est une fonction de classe $\mathcal{C}^k$, et :
   $$\mu_k(X) = \frac{1}{i^k}\varphi^{(k)}(0)$$
   (f) Montrer que si X possède une variance alors :
   $$\varphi(t) = 1 + iE(X)t - \frac{1}{2}E(X^2)t^2 + o(t^2)$$
   (g) Soient Y et Z deux variables aléatoires à densité, indépendantes, montrer que :
   $$\forall t \in \mathbb{R},\ \varphi_{(Y+Z)}(t) = \varphi_X(t)\varphi_Y(t)$$
2. Quelques fonctions caractéristiques
   (a) Déterminer la fonction caractéristique d'une variable aléatoire uniforme sur un intervalle $[a;b]$.
   (b) Déterminer la fonction caractéristique d'une variable aléatoire exponentielle de paramètre λ.
   (c) Soit X une variable aléatoire normale, centrée, réduite, on note φ sa fonction caractéristique, montrer que $\forall t,\ \varphi'(t) = -t\varphi(t)$. En déduire φ.
   (d) Déterminer la fonction caractéristique d'une variable aléatoire normale de paramètres m, σ^2.
3. Quelques applications
   (a) Déterminer l'espérance et la variance d'une variable aléatoire normale de paramètres m, σ^2.
   (b) On admet le résultat suivant très important qui justifie l'appellation de fonction caractéristique, deux variables aléatoires ayant la même fonction caractéristique ont même loi (par exemple elles ont même fonction de répartition)
   Déterminer la loi de la somme de deux variables aléatoires normales indépendantes de paramètre (m_1, σ_1^2) et (m_2, σ_2^2).
   (c) Soient (X_n) une suite de variables aléatoires à densité, et X une variable aléatoire à densité, on note F_n et F les fonctions de répartitions et φ_n et φ les fonctions caractéristiques. On admet le résultat suivant, la suite de fonction (F_n) converge simplement vers F ssi la suite de fonctions (φ_n) converge simplement vers φ.
   Démontrer le théorème de la limite centrale, pour une suite de variables aléatoires indépendantes de même loi exponentielle de paramètre 1.

Index