

Les probabilités à l'agrégation externe de mathématiques :

Guide pour une révision

~~EXCLU DU PRET~~

IREM de LYON
BIBLIOTHEQUE
Université Claude Bernard -LYON I
43, Bd du 11 Novembre 1918
69622 VILLEURBANNE Cedex

Yves DUCÉL
IREM de BESANÇON

Les probabilités à l'agrégation externe de mathématiques : Guide pour une révision

IREM de LYON
BIBLIOTHEQUE
Université Claude Bernard - LYON I
43, Bd du 11 Novembre 1918
69622 VILLEURBANNE Cedex

Yves DUCCEL
IREM de BESANÇON

Cette brochure reprend un travail réalisé à l'Université de Franche-Comté dans le cadre de la préparation à l'option "Probabilités et Statistiques" du concours externe de l'agrégation de mathématiques.

Ce document se veut avant-tout un guide pour l'étudiant agrégatif dans sa révision des principales notions au programme de l'épreuve écrite optionnelle et de l'aider efficacement à maîtriser une bibliographie utile également dans la préparation des leçons de probabilités et d'analyse de l'épreuve orale. A cet effet les principaux théorèmes sont accompagnés de références précises des ouvrages où l'étudiant pourra se reporter pour trouver une ou plusieurs démonstrations du résultat. Le recours à plusieurs auteurs doit permettre au candidat de disposer d'approches différentes sur une même notion et ainsi de mieux la cerner.

Bon nombre de résultats cités sont démontrés de façon très détaillée dans les manuels d'exercices corrigés en théorie de la mesure, de l'intégration et des probabilités, [1] et [2], co-rédigés avec Jean-Pascal Ansel, ainsi que dans les ouvrages [3] et [4] en préparation.

L'ensemble de ce travail a bénéficié des nombreuses discussions avec Jean-Pascal Ansel ainsi que de ses compétences informatiques. Je tiens à l'en remercier.

Table des Matières

Introduction	3
1 Le formalisme des probabilités	7
1.1 Loi d'une variable aléatoire	7
1.2 Moments d'une variable aléatoire	10
1.3 Fonction caractéristique d'une v.a.	12
2 L'indépendance stochastique	15
2.1 Indépendance d'événements et de tribus	15
2.2 Indépendance de v.a. à valeurs vectorielles	17
2.3 Probabilité et espérance conditionnelles	20
2.4 Uniforme intégrabilité et martingales	22
2.5 Vecteurs gaussiens et statistiques de v.a.r.	22
3 Convergences de variables aléatoires	23
3.1 Convergences dans L^0	23
3.2 Comparaison des convergences dans L^0	25
3.3 Théorèmes-limites dans L^0	27
4 Convergence de lois de probabilité	31
4.1 Convergence étroite dans $M_1(\mathbb{R})$	31
4.2 Convergence en loi de v.a.r.	35
4.3 Théorèmes-limites en loi	36
Annexe : Programmes de l'agrégation	40
Bibliographie	42

Chapitre 1

Le formalisme des probabilités

Le formalisme de la théorie des probabilités utilise les outils de la théorie de la mesure en adoptant un vocabulaire spécifique aux probabilités. Il s'agit donc dans un premier temps de préciser ce nouveau langage avant d'introduire des concepts propres à cette théorie. Pour une justification du choix du formalisme et de sa signification en tant que modèle d'une "réalité", on pourra se reporter à [7] chapitre I, [5] p. 93 et 132, [17] p. 56. Pour une présentation de la théorie de la mesure et de l'intégration, on pourra consulter avec profit pour le cours [11], [17] et pour les exercices [1].

1.1 Loi d'une variable aléatoire

Une **probabilité** $\mathbb{P}$ sur un espace mesurable $(\Omega, \mathcal{F})$ est une mesure positive sur $(\Omega, \mathcal{F})$ telle que $\mathbb{P}(\Omega) = 1$. On dit alors que $(\Omega, \mathcal{F}, \mathbb{P})$ est un **espace de probabilité**. Ω s'appelle l'espace des **réalisations** ou des **épreuves**. Les éléments de la tribu $\mathcal{F}$ s'appellent les **événements de Ω** . Une propriété vraie $\mathbb{P}$ -presque-partout est dite **$\mathbb{P}$ -presque-sûre**, en abrégé $\mathbb{P}$ -p.s. ou plus simplement p.s.

Dans la suite $(\Omega, \mathcal{F}, \mathbb{P})$ désignera toujours un espace de probabilité et $(E, \mathcal{E})$ un espace mesurable.

Une application $(\mathcal{F}, \mathcal{E})$ -mesurable X de Ω dans E s'appelle une **variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs E** , en abrégé v.a. à valeurs E .

Si $E = \mathbb{R}$ (resp. $E = \overline{\mathbb{R}}$, $E = \mathbb{R}^d$ où $d \in \mathbb{N}^*$) on dit que X est une v.a. **réelle** (resp. **numérique**, **de dimension d**), en abrégé v.a.r., v.a.n., v.a. $\mathbb{R}^d$.

D'après la théorie de la mesure :

Proposition 1 - 1 *Une application X de Ω dans $\mathbb{R}^d$ est une v.a. si, et seulement si, ses composantes sont des v.a.r.*

L'image de la probabilité $\mathbb{P}$ par la v.a. X est une probabilité sur $(E, \mathcal{E})$, notée $\mathbb{P}_X$, et appelée la **loi (de probabilité) de la v.a. X** . Pour tout $A \in \mathcal{E}$, on a donc $\mathbb{P}_X(A) = \mathbb{P}(X \in A) := \mathbb{P}(\{\omega \in \Omega; X(\omega) \in A\})$.

Par le théorème du transfert il vient :

Proposition 1 - 2 Soient X une v.a. à valeurs E et f une application $\mathcal{E}$ -mesurable de E dans $\mathbb{R}^+$ alors,

$$\int_{\Omega} f(X(\omega)) d\mathbb{P}(\omega) = \int_E f(x) d\mathbb{P}_X(x).$$

L'existence des v.a., de loi donnée a priori, est assurée par :

Proposition 1 - 3 Soient $(E, \mathcal{E})$ un espace mesurable et μ une probabilité sur $(E, \mathcal{E})$. Alors il existe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans E et de loi de probabilité μ .

On rappelle le théorème d'extension de Hahn-Carathéodory adapté au cas des probabilités (cf. [1] p. 34, [9] p. 403, [15] p. 276, [17] p. 54, [11] pp 110-116) :

Proposition 1 - 4 Théorème de prolongement

Soient E un ensemble, $\mathcal{A}$ une famille de parties de E contenant E , stable par réunion finie et passage au complémentaire i.e. $\mathcal{A}$ est une **algèbre** sur E , μ une application de $\mathcal{A}$ dans $[0, +\infty]$ vérifiant $\mu(E) = 1$ et $\mu(\bigcup_n A_n) = \sum_n \mu(A_n)$ pour toute suite $(A_n)_{\mathbb{N}}$ d'éléments deux à deux disjoints de $\mathcal{A}$ telle que $\bigcup_n A_n \in \mathcal{A}$. Alors il existe une unique probabilité $\hat{\mu}$ sur $(E, \sigma(\mathcal{A}))$ telle que, pour tout $A \in \mathcal{A}$, $\hat{\mu}(A) = \mu(A)$.

Le résultat suivant est très utile dans la détermination des lois de v.a.r. :

Proposition 1 - 5 Soit $(E, \mathcal{E})$ un espace mesurable et μ une probabilité sur $(E, \mathcal{E})$. Alors une v.a. X à valeurs E sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ admet μ pour loi de probabilité si, et seulement si, pour toute application $\mathcal{E}$ -mesurable et positive h de E dans $[0, +\infty]$,

$$\int_{\Omega} h(X(\omega)) d\mathbb{P}(\omega) = \int_E h(x) d\mu(x).$$

Une mesure μ sur $\mathbb{R}^d$ est dite **discrète** si elle s'écrit $\mu = \sum_{n \in \mathbb{N}} p_n \delta_{e_n}$ où $(p_n)_{\mathbb{N}}$ est une suite de réels positifs ou nuls et $(e_n)_{\mathbb{N}}$ une suite d'éléments de $\mathbb{R}^d$.

Une v.a. X à valeurs dans $\mathbb{R}^d$ est dite **discrète** si sa loi est discrète. Dans ce cas, pour tout $n \in \mathbb{N}$, $p_n = \mathbb{P}(X = e_n)$.

Les v.a. de Poisson, de Bernoulli, binomiale, hypergéométrique sont des exemples de v.a.r. discrètes. De façon précise les v.a.r. discrètes sont les v.a.r. presque-sûrement à valeurs dans un ensemble dénombrable :

Proposition 1 - 6 (cf. exercice chapitre I dans [2])

Une v.a. X à valeurs $\mathbb{R}^d$ est discrète si, et seulement si, il existe une partie D de $\mathbb{R}^d$ dénombrable telle que $IP(X \in D) = 1$. Dans ce cas la loi de la v.a.r. X est

$$IP_X = \sum_{x \in D} IP(X = x) \delta_x.$$

On dit aussi que la loi de X est **portée par D** .

Les v.a. discrètes constituent une famille de v.a.r importante dans les applications des probabilités, une autre classe de v.a.r. très importante aussi est celle des v.a.r. absolument continues.

On dit qu'une mesure μ sur $\mathbb{R}^d$ est **absolument continue sur $\mathbb{R}^d$** ou encore à **densité sur $\mathbb{R}^d$** si elle admet une densité f par rapport à la mesure de Lebesgue $\lambda^{(d)}$ de $\mathbb{R}^d$, i.e. f est une application de $\mathbb{R}^d$ dans $[0, +\infty[$ intégrable au sens de Lebesgue sur $\mathbb{R}^d$ et, pour tout A borélien de $\mathbb{R}^d$, on a $\mu(A) = \int_A f d\lambda^{(d)}$. On note alors $\mu = f \cdot \lambda^{(d)}$.

On dit qu'une v.a. à valeurs $\mathbb{R}^d$ est **absolument continue sur $\mathbb{R}^d$** ou encore à **densité sur $\mathbb{R}^d$** si sa loi est absolument continue sur $\mathbb{R}^d$.

Les v.a.r. de Gauss-Laplace, exponentielle, uniforme sur un intervalle de $\mathbb{R}$, de Cauchy sont des exemples de v.a.r. absolument continues sur $\mathbb{R}$.

Pour un inventaire des lois classiques se reporter à [5] p. 144 et 367, pour une étude détaillée des lois discrètes classiques à [12] pp 145-171 et des lois absolument continues classiques à [12] pp 174-187.

Si μ est une probabilité sur $\mathbb{R}$, on appelle **fonction de répartition de μ** (en abrégé f.r.), et on note F_μ , l'application de $\mathbb{R}$ dans $[0, 1]$ définie, pour tout $x \in \mathbb{R}$, par $F_\mu(x) := \mu(]-\infty, x])$. Si X est une v.a.r., la f.r. de la loi de X est appelée **fonction de répartition de X** , on la note alors F_X .

La f.r. d'une probabilité a les propriétés suivantes (cf. exercice II-6 [1] p. 39) :

Proposition 1 - 7 La f.r. de μ est croissante, continue à droite sur $\mathbb{R}$ avec $\lim_{x \rightarrow +\infty} F_\mu(x) = 1$ et $\lim_{x \rightarrow -\infty} F_\mu(x) = 0$. Pour tout $a \in \mathbb{R}$, F_μ est continue en a si, et seulement si, $\mu(\{a\}) = 0$. De plus l'ensemble des points de discontinuité d'une f.r. est dénombrable.

Réciproquement par application du théorème de prolongement on a :

Proposition 1 - 8 Soit F une application de $\mathbb{R}$ dans $[0, 1]$, croissante, continue à droite sur $\mathbb{R}$ avec $\lim_{x \rightarrow +\infty} F(x) = 1$ et $\lim_{x \rightarrow -\infty} F(x) = 0$. Alors il existe une probabilité et une seule μ sur $\mathbb{R}$ telle que $F_\mu = F$. La probabilité μ est déterminée, pour tous réels a, b avec $a < b$, par $\mu(]a, b]) = F(b) - F(a)$.

On trouvera une démonstration de la proposition suivante dans cf. [9] p. 7 et [2] exercice chapitre II :

Proposition 1 - 9 *Soit F une application de $\mathbb{R}$ dans $[0, 1]$, croissante, continue à droite sur $\mathbb{R}$ avec $\lim_{x \rightarrow +\infty} F(x) = 1$ et $\lim_{x \rightarrow -\infty} F(x) = 0$. Alors il existe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et une v.a.r. X sur $(\Omega, \mathcal{F}, \mathbb{P})$ telle que $F_X = F$.*

On notera que dans ce cas la v.a.r. X n'est pas unique.

Le résultat suivant est souvent utilisé pour prouver l'unicité de certaines probabilités (cf. [8] p. 63, [5] p.103, [15] p. 267, [17] p. 48) :

Proposition 1 - 10 Lemme d'unicité des probabilités

Soit $(\Omega, \mathcal{F})$ un espace mesurable, $\mathcal{C}$ une famille stable par intersection finie et engendrant la tribu $\mathcal{F}$. Si deux probabilités sur $(\Omega, \mathcal{F})$ coïncident sur $\mathcal{C}$, alors elles sont égales.

1.2 Moments d'une variable aléatoire

Comme $\mathbb{P}$ est une mesure finie sur $(\Omega, \mathcal{F})$ on a, pour tout p et q dans $[1, +\infty[$ avec $p < q$, la chaîne d'inclusions (cf. [1] p. exercices IV-8 et IV-11) :

$$L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \subseteq L^q(\Omega, \mathcal{F}, \mathbb{P}) \subseteq L^p(\Omega, \mathcal{F}, \mathbb{P}) \subseteq L^1(\Omega, \mathcal{F}, \mathbb{P}).$$

Si X est une v.a.r. sur $(\Omega, \mathcal{F}, \mathbb{P})$ positive ou intégrable sur Ω , on appelle **espérance mathématique de X** , et on note $\mathbb{E}(X)$, la quantité (éventuellement infinie)

$$\mathbb{E}(X) := \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

Plus généralement, si $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ et $p \in \mathbb{N}^*$, on appelle **moment d'ordre p de X** , resp. **moment centré d'ordre p de X** , le nombre réel

$$m_p := \int_{\Omega} X^p(\omega) d\mathbb{P}(\omega), \text{ resp. } m'_p := \int_{\Omega} (X(\omega) - m_1)^p d\mathbb{P}(\omega).$$

Le moment centré d'ordre 2 s'appelle la **variance de X** , et se note σ_X^2 ou $\text{Var}(X)$. Sa racine carrée σ_X s'appelle l'**écart-type de X** .

Si X et Y sont deux v.a.r. de carré intégrable, on appelle **covariance de X et Y** , et on note $\text{Cov}(X, Y)$, le réel $\text{Cov}(X, Y) := \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.

Par application du théorème de transfert il vient :

Proposition 1 - 11 *Sous les conditions d'existence,*

$$m_1 := \mathbb{E}(X) = \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int_{\mathbb{R}} x d\mathbb{P}_X(x).$$

$$\begin{aligned}
\sigma_X^2 &:= \int_{\Omega} (X(\omega) - m_1)^2 dIP(\omega) = \int_{\mathbb{R}} (x - m_1)^2 dIP_X(x). \\
\sigma_X^2 &= \text{Cov}(X, X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2. \\
m_p &:= \int_{\Omega} X^p(\omega) dIP(\omega) = \int_{\mathbb{R}} x^p dIP_X(x). \\
m'_p &:= \int_{\Omega} (X(\omega) - m_1)^p dIP(\omega) = \int_{\mathbb{R}} (x - m_1)^p dIP_X(x).
\end{aligned}$$

Le résultat suivant établit un lien entre les moments et la fonction de répartition d'une v.a.r. positive. On notera qu'il n'est plus valable si la v.a.r. n'est plus supposée positive (cf. [1] exercice VI-8 p. 102 ou exercice I-7 de [2]).

Proposition 1 - 12 *Si est X une v.a.r. positive de fonction de répartition F_X , alors pour tout $n \in \mathbb{N}^*$, l'application, qui à $t \in \mathbb{R}$ associe $nt^{n-1}IP(X > t)$, est intégrable au sens de Riemann sur tout intervalle compact de $\mathbb{R}$ et vérifie la relation*

$$\mathbb{E}[X^n] = \int_0^{+\infty} nt^{n-1} IP(X > t) dt = \int_0^{+\infty} nt^{n-1} (1 - F_X(t)) dt.$$

Soit X une v.a. à valeurs $\mathbb{R}^d$ de composantes $X_1, X_2, \dots, X_d$ intégrables sur Ω . On appelle **espérance de X** , et on note $\mathbb{E}(X)$, le vecteur de $\mathbb{R}^d$,

$$\mathbb{E}(X) := (\mathbb{E}(X_1), \mathbb{E}(X_2), \dots, \mathbb{E}(X_d)).$$

Si X est une v.a. à valeurs matricielles, on appelle **espérance de X** , et on note $\mathbb{E}(X)$, la matrice dont les coefficients sont les espérances de ceux de X .

Soit X une v.a. à valeurs $\mathbb{R}^d$ de composantes $X_1, X_2, \dots, X_d$ de carré intégrable sur Ω . On appelle **matrice de dispersion** de X l'espérance de la matrice carrée d'ordre d , $(X - \mathbb{E}(X))(X - \mathbb{E}(X))^*$, où $*$ désigne l'opération de transposition des matrices. On dit aussi **matrice des covariances de X** et on la note D_X .

La démonstration de la proposition suivante se trouve dans [5] p. 119 et [2] chapitre I :

Proposition 1 - 13 *Si X est une v.a. à valeurs $\mathbb{R}^d$ et M une matrice (déterministe) à coefficients réels à c lignes et d colonnes, alors*

1.

$$D_{[X - \mathbb{E}(X)]} = D_X, \quad \mathbb{E}(MX) = M\mathbb{E}(X), \quad D_{MX} = MD_X M^*.$$

2. D_X est une matrice symétrique. D_X est de type positif i.e., pour tout $u \in \mathbb{R}^d$, $u^* D_X u \geq 0$. Le coefficient d'indice (i, j) de D_X est la covariance $\text{Cov}(X_i, X_j)$ des v.a.r. X_i et X_j . Les éléments diagonaux de D_X sont les variances des composantes de X .

3. Le déterminant de D_X est nul si, et seulement si, il existe des réels non tous nuls $a_0, a_1, \dots, a_d$ tels que $a_1 X_1 + a_2 X_2 + \dots + a_d X_d = a_0$ p.s.

Donnons deux inégalités souvent utilisées en probabilité (cf. [5] p. 118, [2] chapitre I) :

Proposition 1 - 14 Inégalité de Markov

Soit X une v.a. à valeurs $\mathbb{R}^d$, φ une application borélienne de $\mathbb{R}^d$ dans $[0, +\infty[$, alors pour tout réel $a > 0$,

$$\mathbb{P}(\varphi(X) \geq a) \leq \frac{1}{a} \mathbb{E}(\varphi(X)).$$

Proposition 1 - 15 Inégalité de Bienaymé-Chebyshev

Soit X une v.a.r. dans L^p où $p \in [1, +\infty[$, alors pour tout réel $a > 0$,

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{1}{a^p} \mathbb{E}(|X - \mathbb{E}(X)|^p).$$

En particulier si X est une v.a.r. de carré intégrable, on a

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{1}{a^2} \text{Var}(X).$$

1.3 Fonction caractéristique d'une v.a.

On pourra se reporter pour plus de détails au thème d'étude-chapitre VII de [1] p. 117, ou à [5] pp 138-144, [9] pp 76-93, [15] pp 147-169.

On appelle **fonction caractéristique** d'une probabilité μ sur $\mathbb{R}^d$ (en abrégé f.c.), et on note Φ_μ , l'application de $\mathbb{R}^d$ dans $\mathbb{C}$ définie, pour tout $u \in \mathbb{R}^d$, par

$$\Phi_\mu(u) := \int_{\mathbb{R}^d} e^{i\langle u, x \rangle} d\mu(x),$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien de $\mathbb{R}^d$.

Si X est une v.a. à valeurs $\mathbb{R}^d$, on appelle **fonction caractéristique de X** la f.c. de sa loi $\mathbb{P}_X$. On la note Φ_X . On a donc la relation :

Proposition 1 - 16 Pour toute v.a. X à valeurs $\mathbb{R}^d$ et tout $u \in \mathbb{R}^d$,

$$\Phi_X(u) = \mathbb{E}(e^{i\langle u, X \rangle}).$$

La f.c. d'une v.a. à valeurs $\mathbb{R}^d$ est un outil très utile dans l'identification des lois de v.a. en vertu du théorème suivant qui résulte des formules d'inversion de Perron-Sieltjès :

Proposition 1 - 17 Théorème d'injectivité

Deux v.a. à valeurs $\mathbb{R}^d$ ont la même loi si, et seulement si, elles ont la même fonction caractéristique.

La notion de f.c. est parfois utile dans les calculs de moments de v.a.r. grâce aux deux propositions suivantes :

Proposition 1 - 18 *Si X est une v.a.r. et n un entier naturel tels que $\mathbb{E}(|X^n|) < +\infty$, alors la f.c. Φ_X de X est de classe $\mathcal{C}^n$ sur $\mathbb{R}$ et, pour tout $u \in \mathbb{R}$, sa dérivée d'ordre n en u est donnée par*

$$\Phi_X^{(n)}(u) = \int_{\mathbb{R}} (ix)^n e^{iux} d\mathbb{P}_X(x) = i^n \mathbb{E}(X^n e^{iuX}).$$

En particulier la dérivée d'ordre n de Φ_X en 0 est $\Phi_X^{(n)}(u) = i^n \mathbb{E}(X^n)$ et Φ_X admet un développement limité en 0 à l'ordre n donné par :

$$\Phi_X(u) = \sum_{k=0}^{k=n} \frac{i^k \mathbb{E}(X^k)}{k!} u^k + o(u^n).$$

La proposition précédente montre qu'il y a un lien entre l'intégrabilité de la v.a.r. $|X|^n$ et la régularité de la f.c. de X : Si X est dans L^n , Φ_X est de classe $\mathcal{C}^n$. La réciproque est fautive. En effet, par exemple, il est possible de trouver une v.a.r. X non intégrable dont la f.c. est de classe $\mathcal{C}^1$ (cf [18] p. 65 exemple 8.7).

Cependant, on peut donner une réciproque partielle au résultat précédent sous la forme :

Proposition 1 - 19 *Soient X une v.a.r. et n un entier naturel tels que Φ_X soit dérivable en 0 à l'ordre n . Alors $\mathbb{E}(X^{2p}) < +\infty$ où $2p$ est le plus grand entier pair inférieur ou égal à n . En particulier, X admet des moments jusqu'à l'ordre $2p$. De plus, si Φ_X admet un développement limité en 0 à l'ordre n s'écrivant*

$$\Phi_X(u) = 1 + \sum_{k=1}^{k=n} a_k u^k + o(u^n)$$

alors, pour tout $1 \leq k \leq 2p$, $\mathbb{E}(X^k) = (-i)^k a_k k!$.

Chapitre 2

L'indépendance stochastique

$(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace de probabilité et I un ensemble quelconque supposé non vide. Si $(X_i)_I$ est une famille quelconque de v.a. sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, on notera $\sigma(X_i, i \in I)$ la plus petite (au sens de l'inclusion) des sous-tribus $\mathcal{A}$ de $\mathcal{F}$ telles que, pour tout $i \in I$, X_i est $\mathcal{A}$ -mesurable.

2.1 Indépendance d'événements et de tribus

Une famille quelconque d'événements, $(A_i)_I$, est dite **(mutuellement) indépendante pour $\mathbb{P}$** , ou plus simplement **indépendante** si, pour toute sous-famille finie $(A_i)_K$, $K \subseteq I$ et K fini,

$$\mathbb{P} \left(\bigcap_{i \in K} A_i \right) = \prod_{i \in K} \mathbb{P}(A_i).$$

On dit aussi que les événements $(A_i)_I$ sont **indépendants**.

On remarquera que la notion d'indépendance dépend de la probabilité $\mathbb{P}$ choisie sur $(\Omega, \mathcal{F})$. De plus l'indépendance mutuelle d'une famille d'événements entraîne leur indépendance deux à deux, cependant il faut noter que la réciproque est fausse (cf. [18] exemple 3-1 à 3-3).

Proposition 2 - 1 *Si la famille d'événements $(A_i)_I$ est indépendante, il en est de même de la famille d'événements $(B_i)_I$ où, pour tout $i \in I$, $B_i := A_i$ ou $B_i := A_i^c$.*

Une famille quelconque de sous-tribus de $\mathcal{F}$, $(\mathcal{A}_i)_I$, est dite **(mutuellement) indépendante pour $\mathbb{P}$** si toute famille $(A_i)_I$ vérifiant, pour tout $i \in I$, $A_i \in \mathcal{A}_i$ est indépendante.

La proposition suivante est une reformulation de la proposition précédente :

Proposition 2 - 2 *La famille d'événements $(A_i)_I$ est indépendante si, et seulement si, la famille des sous-tribus $(\sigma(A_i))_I$ est indépendante.*

Le résultat suivant est utile pour prouver l'indépendance de tribus à partir de leurs systèmes de générateurs (cf. [8] p. 94) :

Proposition 2 - 3 *Soient $\mathcal{C}$ et $\mathcal{D}$ deux familles d'événements stables par intersection finie. On suppose que, pour tout $(A, B) \in \mathcal{C} \times \mathcal{D}$, A et B sont indépendants, alors les sous-tribus engendrées respectivement par $\mathcal{C}$ et $\mathcal{D}$ sont indépendantes.*

En particulier, on a le corollaire suivant (cf. [2] exercice II-19) :

Proposition 2 - 4 *Soit $(\mathcal{F}_i)_I$ une famille indépendante de sous-tribus de $\mathcal{F}$. Si K et J sont deux parties disjointes de I , alors les tribus engendrées respectivement par les familles $\bigcup_{i \in J} \mathcal{F}_i$ et $\bigcup_{i \in K} \mathcal{F}_i$ sont indépendantes.*

Les deux résultats suivants sont une application de cette proposition :

Proposition 2 - 5 Loi du Tout ou Rien de Kolmogorov (cas des tribus, cf. [8] p. 95 et [2] problème du chapitre II)

Soit $(\mathcal{F}_n)_{\mathbb{N}}$ une suite de sous-tribus indépendantes. Notons, pour tout $n \in \mathbb{N}$, $\mathcal{A}_n$ la tribu engendrée par la famille $\bigcup_{k \geq n} \mathcal{F}_k$ i.e. $\mathcal{A}_n := \sigma(\bigcup_{k \geq n} \mathcal{F}_k)$. Si A est mesurable par rapport à toutes les tribus $\mathcal{A}_n$, alors $\mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$.

Proposition 2 - 6 Loi du Tout ou Rien de Kolmogorov (cas des événements, cf. [5] p.127)

Soit $(A_n)_{\mathbb{N}}$ une suite d'événements indépendants. Notons, pour tout $n \in \mathbb{N}$, $\mathcal{A}_n$ la tribu engendrée par la suite $(A_k)_{k \geq n}$ i.e. $\mathcal{A}_n := \sigma(A_n, A_{n+1}, \dots, A_{n+k}, \dots)$. Si A est mesurable par rapport à toutes les tribus $\mathcal{A}_n$, alors $\mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$.

On remarquera que la mesurabilité par rapport à toutes les tribus $\mathcal{A}_n$ équivaut à la mesurabilité par rapport à la tribu $\bigcap_n \mathcal{A}_n$ appelée **tribu asymptotique** associée à la suite de tribus $(\mathcal{A}_n)_{\mathbb{N}}$.

Si $(A_n)_{\mathbb{N}}$ est une suite d'événements, on note $\limsup_n(A_n)$ l'ensemble des $\omega \in \Omega$ tels que $\{n \in \mathbb{N} / \omega \in A_n\}$ est infini. Dire que l'événement $\limsup_n(A_n)$ est réalisé signifie qu'une infinité d'événements de la suite $(A_n)_{\mathbb{N}}$ sont réalisés. On note aussi $\liminf_n(A_n)$ l'ensemble des $\omega \in \Omega$ tels que $\{n \in \mathbb{N} / \omega \notin A_n\}$ est fini. Dire que l'événement $\liminf_n(A_n)$ est réalisé signifie que tous les événements de la suite $(A_n)_{\mathbb{N}}$, sauf éventuellement un nombre fini d'entre eux, sont réalisés. On a donc $\liminf_n(A_n) \subseteq \limsup_n(A_n)$.

Il résulte de la loi du Tout ou Rien que, pour une suite indépendante $(A_n)_{\mathbb{N}}$ d'événements, on a $\mathbb{P}(\limsup_n(A_n)) = 0$ ou $\mathbb{P}(\limsup_n(A_n)) = 1$. Le résultat suivant très important en probabilité précise la remarque précédente (cf. [5] p. 38) :

Proposition 2 - 7 Lemme de Borel-Cantelli

1. Soit $(A_n)_\mathbb{N}$ une suite (non nécessairement indépendante) d'événements.
Si la série numérique $\sum_n \mathbb{P}(A_n)$ converge dans $\mathbb{R}$, alors

$$\mathbb{P}(\limsup_n (A_n)) = 0,$$

cela signifie que presque-sûrement le nombre des événements réalisés de la suite $(A_n)_\mathbb{N}$ est fini.

2. Soit $(A_n)_\mathbb{N}$ une suite indépendante d'événements.
Si la série numérique $\sum_n \mathbb{P}(A_n)$ diverge dans $\mathbb{R}$, alors

$$\mathbb{P}(\limsup_n (A_n)) = 1,$$

cela signifie que presque-sûrement le nombre des événements réalisés de la suite $(A_n)_\mathbb{N}$ est infini.

On notera que si la suite $(A_n)_\mathbb{N}$ n'est pas indépendante, on peut avoir $\mathbb{P}(\limsup_n (A_n)) = 0$ avec la série de terme général $\mathbb{P}(A_n)$ divergente (cf. [18] ex.4-2).

2.2 Indépendance de v.a. à valeurs vectorielles

On dit qu'une famille quelconque $(X_i)_I$ de v.a. X_i à valeurs dans $\mathbb{R}^{d_i}$, $i \in I$, est (mutuellement) indépendante pour $\mathbb{P}$ si la famille de sous-tribus $(\sigma(X_i))_I$ est indépendante.

Le lien entre les notions d'indépendance pour les événements, les v.a. et les tribus est mis en évidence par les deux propositions suivantes :

Proposition 2 - 8 La famille d'événements $(A_i)_I$ est indépendante si, et seulement si, la famille des v.a.r. $(\mathbf{1}_{A_i})_I$ est indépendante.

Proposition 2 - 9 Si $\mathcal{A}$ et $\mathcal{B}$ sont deux sous-tribus de $\mathcal{F}$ indépendantes, X et Y deux v.a.r. respectivement $\mathcal{A}$ -mesurable et $\mathcal{B}$ -mesurable, alors les v.a.r. X et Y sont indépendantes.

Le résultat suivant affirme l'existence de suites (finies ou infinies) de v.a.r. indépendantes de lois données a priori (cf. [5] p. 125 et 136, pour le cas des suites finies [2] exercice II-2).

Proposition 2 - 10 Soit $(\mu_i)_I$, où $I \subseteq \mathbb{N}$, une famille de probabilités sur $\mathbb{R}$. Alors il existe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et $(X_i)_I$ une suite de v.a.r. indépendantes sur $(\Omega, \mathcal{F}, \mathbb{P})$ tels que, pour tout $i \in I$, μ_i soit la loi de X_i .

Proposition 2 - 11 Théorème d'indépendance des fonctions de v.a.r.

Si $(X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m)$ est une suite finie indépendante de v.a.r., f (resp. g) une application borélienne de $\mathbb{R}^n$ (resp. $\mathbb{R}^m$) dans $\mathbb{R}^q$ (resp. $\mathbb{R}^p$), alors le couple de v.a. $(f(X_1, X_2, \dots, X_n), g(Y_1, Y_2, \dots, Y_m))$ est indépendant.

La réciproque est fausse (cf. [18] exemple 7-5).

Pour les v.a., la loi du Tout ou Rien de Kolmogorov devient :

Proposition 2 - 12 Loi du Tout ou Rien de Kolmogorov (cas des v.a.)

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. indépendantes. Notons, pour tout $n \in \mathbb{N}$, $\mathcal{A}_n$ la tribu engendrée par la suite de v.a. $(X_k)_{k \geq n}$ i.e. $\mathcal{A}_n := \sigma(X_n, X_{n+1}, \dots, X_{n+k}, \dots)$. Si A est mesurable par rapport à toutes les tribus $\mathcal{A}_n$, alors $\mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$.

Donnons maintenant quelques critères d'indépendance ([8] p. 93, [5] p. 125, [9] p. 23, [15] p. 121) :

Proposition 2 - 13 Soit $(X_1, X_2, \dots, X_n)$ une suite de v.a. à valeurs respectivement dans $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}, \dots, \mathbb{R}^{d_n}$. Les assertions suivantes sont équivalentes :

1. La suite de v.a. $(X_1, X_2, \dots, X_n)$ est indépendante.
2. $\mathbb{P}_{(X_1, X_2, \dots, X_n)} = \mathbb{P}_{X_1} \otimes \mathbb{P}_{X_2} \otimes \dots \otimes \mathbb{P}_{X_n}$.
3. Pour toutes applications f_i boréliennes bornées (ou positives) de $\mathbb{R}^{d_i}$ dans $\mathbb{R}$, $i = 1, 2, \dots, n$,

$$\mathbb{E}[f_1(X_1)f_2(X_2)\dots f_n(X_n)] = \mathbb{E}[f_1(X_1)]\mathbb{E}[f_2(X_2)]\dots \mathbb{E}[f_n(X_n)].$$

4. Pour tout $(u_1, u_2, \dots, u_n) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$,

$$\Phi_{(X_1, X_2, \dots, X_n)}(u_1, u_2, \dots, u_n) = \Phi_{X_1}(u_1)\Phi_{X_2}(u_2)\dots \Phi_{X_n}(u_n)$$

où Φ_{X_i} désigne la fonction caractéristique de X_i .

On notera que si $X_1, X_2, \dots, X_n$ est une suite indépendante de v.a.r. intégrables, alors

$$\mathbb{E}(X_1 X_2 \dots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)\dots \mathbb{E}(X_n).$$

Cependant la réciproque est fausse (cf. [18] exemples 7-9 et 7-10).

Pour les v.a. absolument continues, le critère suivant est souvent commode, où la notation $\otimes$ signifie que, pour tout $(x_1, x_2, \dots, x_n) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$,

$$f_1 \otimes f_2 \otimes \dots \otimes f_n(x_1, x_2, \dots, x_n) := f_1(x_1)f_2(x_2)\dots f_n(x_n).$$

Proposition 2 - 14 Soient $X_1, X_2, \dots, X_n$ des v.a. à valeurs respectivement dans $\mathbb{R}^{d_1}, \mathbb{R}^{d_2}, \dots, \mathbb{R}^{d_n}$.

1. Si les v.a. $X_1, X_2, \dots, X_n$ sont indépendantes et absolument continues de densités respectives $f_1, f_2, \dots, f_n$, alors la v.a. $X := (X_1, X_2, \dots, X_n)$ à valeurs dans $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$ est absolument continue de densité $f := f_1 \otimes f_2 \otimes \dots \otimes f_n$.
2. Réciproquement, si la v.a. $X := (X_1, X_2, \dots, X_n)$ à valeurs dans $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$ est absolument continue de densité $f := f_1 \otimes f_2 \otimes \dots \otimes f_n$, alors les v.a. $X_1, X_2, \dots, X_n$ sont indépendantes et absolument continues. Dans ce cas, pour tout $i = 1, 2, \dots, n$, la densité de la v.a. X_i est proportionnelle à f_i .

On notera que dans la réciproque les fonctions f_i ne sont pas nécessairement les densités des v.a. X_i . De plus si X et Y sont deux v.a. telles que la v.a. (X, Y) soit absolument continue, alors elles sont aussi absolument continues. La réciproque de ce résultat est fausse (cf. [18] exemple 5-6)

Enfin un dernier critère valable pour les vecteurs gaussiens :

Proposition 2 - 15 Soit $X := (X_1, X_2, \dots, X_d)$ une v.a. gaussienne à valeurs dans $\mathbb{R}^d$. Alors les v.a.r. $X_1, X_2, \dots, X_d$ sont indépendantes si, et seulement si, la matrice de dispersion de la v.a. X est diagonale.

Si le vecteur aléatoire $X := (X_1, X_2, \dots, X_d)$ n'est pas gaussien mais a ses composantes de carré intégrable, la condition nécessaire de la proposition précédente reste vraie, par contre la condition suffisante devient fausse (cf. [18] exemple 10-4 iii). La condition nécessaire s'énonce souvent sous la forme :

Proposition 2 - 16 Si $X_1, X_2, \dots, X_d$ sont des v.a.r. indépendantes et de carré intégrable, alors, pour tout $1 \leq i < j \leq d$, $\text{Cov}(X_i, X_j) = 0$.

Deux v.a.r. X et Y telles que $\text{Cov}(X, Y) = 0$ sont dites **non-corrélées**. Plus généralement, on dira qu'une suite de v.a.r. $(X_n)_N$ est **non-corrélée** si, pour tout $(i, j) \in \mathbb{N}^2$ avec $i \neq j$, X_i et X_j sont non-corrélées. D'après la remarque ci-dessus, l'indépendance d'un couple de v.a.r. implique sa non-corrélation, mais la réciproque est fausse.

Donnons quelques propriétés des sommes de v.a. indépendantes :

Proposition 2 - 17 1. Si X et Y sont deux v.a. indépendantes à valeurs $\mathbb{R}^d$ alors, pour tout $u \in \mathbb{R}^d$,

$$\Phi_{X+Y}(u) = \Phi_X(u) \cdot \Phi_Y(u)$$

où Φ_X désigne la fonction caractéristique de X .

2. Si X et Y sont des v.a.r. indépendantes de carré intégrable, alors

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Ces deux conditions nécessaires d'indépendance n'admettent pas de réciproque (cf. [18] exemple 7-6 ii). Pour la relation sur les variances, la non-corrélation des v.a.r. est en fait suffisante. La proposition suivante donne une application du résultat précédent sur les fonctions caractéristiques :

- Proposition 2 - 18** 1. Si X et Y sont deux v.a. indépendantes binomiales de paramètre p et de tailles respectives n et m , alors $X + Y$ est une v.a. binomiale de paramètre p et de taille $n + m$.
2. Si X et Y sont deux v.a. indépendantes de Poisson de paramètres respectifs α et β , alors $X + Y$ est une v.a. de Poisson de paramètre $\alpha + \beta$.
3. Si X et Y sont deux v.a. indépendantes de lois normales de moyennes respectives m , m' et de variances σ^2 , σ'^2 , alors $X + Y$ est une v.a. de moyenne $m + m'$ et de variance $\sigma^2 + \sigma'^2$.

Dans la proposition suivante la notation $*$ désigne le **produit de convolution** de deux fonctions f et g positives boréliennes de $\mathbb{R}^d$ dans $\mathbb{R}$ défini, pour tout $x \in \mathbb{R}^d$, par

$$f * g(x) := \int_{\mathbb{R}^d} f(x - u)g(u)d\lambda^{(d)}(u).$$

Proposition 2 - 19 Soient X et Y deux v.a. à valeurs $\mathbb{R}^d$ indépendantes et absolument continues de densités respectives f_X et f_Y , alors la v.a. $X + Y$ est absolument continue de densité $f_{X+Y} = f_X * f_Y$.

Pour une étude plus détaillée de la convolution et du cas de plusieurs v.a. on pourra se reporter à [8] p. 99, [17] p. 184, [15] p. 124.

L'inégalité suivante généralise l'inégalité de Chebycheff (cf. [9] p. 48, [2] exercice II-18) :

Proposition 2 - 20 Inégalité de Kolmogorov

Soient $X_1, X_2, \dots, X_n$ des v.a.r. centrées, indépendantes, de carré intégrable et $\varepsilon > 0$, alors

$$\mathbb{P}\left(\max_{1 \leq k \leq n} |X_1 + X_2 + \dots + X_k| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \sum_{k=1}^{k=n} \mathbb{E}(X_k^2) = \frac{1}{\varepsilon^2} \sum_{k=1}^{k=n} \text{var}(X_k).$$

2.3 Probabilité et espérance conditionnelles

Soit B un événement de probabilité non nulle. Pour tout $A \in \mathcal{F}$, notons

$$\mathbb{P}^B(A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

On utilise aussi la notation $\mathbb{P}(A|B)$ à la place de $\mathbb{P}^B(A)$.

Proposition 2 - 21 IP^B est une mesure de probabilité sur $(\Omega, \mathcal{F})$.

La probabilité IP^B est appelée **probabilité conditionnelle relativement à l'événement B** .

Proposition 2 - 22 Formule des probabilités totales

Soit $(B_k)_I$, où $I \subseteq \mathbb{N}$, une partition de Ω formée d'événements de probabilités non nulles. Pour tout $A \in \mathcal{F}$,

$$IP(A) = \sum_{k \in I} IP(B_k) IP^{B_k}(A)$$

qu'on peut encore écrire

$$IP = \sum_{k \in I} IP(B_k) IP^{B_k}.$$

Une autre relation facile à établir est :

Proposition 2 - 23 Formule de T. Bayes

Soit $(B_k)_I$, où $I \subseteq \mathbb{N}$, une partition de Ω formée d'événements de probabilités non nulles. Pour tout $A \in \mathcal{F}$ de probabilité non nulle et tout $i \in I$,

$$IP(B_i|A) = \frac{IP(B_i) IP(A|B_i)}{\sum_{k \in I} IP(B_k) IP(A|B_k)}.$$

Soit X une v.a. à valeurs $\mathbb{R}^d$, B un événement tel que $IP(B) > 0$. On appelle **loi conditionnelle de X relativement à l'événement B** la loi $\widetilde{IP}_X$ de la v.a. X par rapport à la probabilité $\widetilde{IP} := IP^B$ c'est-à-dire, pour tout borélien A ,

$$\widetilde{IP}_X(A) = IP_X^B(A) = \frac{IP(\{X \in A\} \cap B)}{IP(B)}.$$

Si Y est une v.a. à valeurs $\mathbb{R}^n$ et $y \in \mathbb{R}^n$ tel que $IP(Y = y) > 0$, on appelle **loi conditionnelle de X relativement à $Y = y$** , la loi conditionnelle de X relativement à l'événement $\{Y = y\}$.

Proposition 2 - 24 Avec les notations précédentes, si de plus X est une v.a. discrète à valeurs $\mathbb{R}^d$ portée par une partie dénombrable D de $\mathbb{R}^d$, la loi conditionnelle de X relativement à $Y = y$ est

$$\sum_{x \in D} \widetilde{IP}(X = x) \delta_x = \sum_{x \in D} IP(X = x | Y = y) \delta_x = \sum_{x \in D} \frac{IP(\{X = x\} \cap \{Y = y\})}{IP(Y = y)} \delta_x.$$

A titre d'exemples on pourra se reporter à [2] exercice II-11 pour montrer la proposition suivante :

Proposition 2 - 25 Soient X, U deux v.a.r. de Poisson indépendantes de paramètres respectifs $\alpha > 0$, $\beta > 0$ et n un entier naturel. Posons $Y := X + U$. La loi conditionnelle de X relativement à $Y = n$ est la loi binomiale de taille n et de paramètre $p = \frac{\alpha}{\alpha + \beta}$.

Pour la notion d'espérance conditionnelle étudier intégralement le chapitre VIII de [5] pp. 195-216.

On pourra consulter aussi [2] chapitre III, [9] pp 190-200, [15] pp. 127-146 et [16] pp 91-98.

2.4 Uniforme intégrabilité et martingales

Pour ces notions étudier intégralement le paragraphe 5 du chapitre IV de [9] pp 223-228.

On pourra consulter aussi le paragraphe sur l'équi-intégrabilité dans [15] pp 192-195.

Il est conseillé aussi d'étudier quelques notions sur les martingales et de se reporter pour cela au chapitre XII de [5] pp 311-323. On pourra consulter aussi [9] pp 200-237.

2.5 Vecteurs gaussiens et statistiques de v.a.r.

Pour la notion de vecteur gaussien, étudier intégralement le paragraphe 9 du chapitre VI de [5] pp 153-157.

On pourra consulter aussi [2] chapitre VII, et [9] pp 190-237.

Pour des résultats classiques sur les statistiques de v.a.r. , on étudiera avec profit les exercices de [10] : II-16, III-7, III-9, III-13, III-15, IV-19, VI-6, VI-9.

Chapitre 3

Convergences de variables aléatoires

La plupart des résultats et définitions de ce chapitre peuvent s'étendre aux v.a. à valeurs dans $\mathbb{R}^d$ en prenant pour $|\cdot|$ la norme euclidienne de $\mathbb{R}^d$.

$(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace de probabilité.

Les espaces $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ et $L^p(\Omega, \mathcal{F}, \mathbb{P})$ sont introduits en théorie de la mesure pour $p \in [1, +\infty]$. On rappelle (cf. [1] p. 67 exercices IV-8 et IV-11) que, pour tout $1 < p < q < +\infty$,

$$L^\infty \subseteq L^q \subseteq L^p \subseteq L^1, \quad \|\cdot\|_{L^p} \leq \|\cdot\|_{L^q}, \quad \bigcap_{p \geq 1} L^p \neq L^\infty, \quad \bigcup_{p \geq 1} L^p \neq L^1.$$

On note $L^0(\Omega, \mathcal{F}, \mathbb{P})$ ou plus simplement L^0 , l'ensemble des classes d'équivalence $\mathbb{P}$ -presque-sûre des v.a. à valeurs $\overline{\mathbb{R}}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$. On notera $\overline{X}$ la classe d'équivalence $\mathbb{P}$ -presque-sûre de la v.a. X . On a, pour tout $p \in [1, +\infty]$, $L^p \subseteq L^0$.

Dans la suite, $(\overline{X}_n)_\mathbb{N}$ est une suite d'éléments de L^0 et $\overline{X}$ un élément de L^0 .

3.1 Convergences dans L^0

Différents modes de convergence peuvent être considérés pour les suites d'éléments de L^0 .

On dit que $(\overline{X}_n)_\mathbb{N}$ **converge $\mathbb{P}$ -presque-sûrement** (ou plus simplement **presque-sûrement**) **vers** $\overline{X}$ si l'ensemble des $\omega \in \Omega$ tels que $(X_n(\omega))_\mathbb{N}$ ne converge pas dans $\overline{\mathbb{R}}$ vers $X(\omega)$ est $\mathbb{P}$ -négligeable.

On note alors $(\overline{X}_n)_\mathbb{N} \xrightarrow{\mathbb{P}\text{-p.s.}} \overline{X}$ et $\overline{X} = p.s.\text{-}\lim_n(\overline{X}_n)$.

En général il n'existe pas de métrique d sur L^0 telle que, pour toute suite $(\overline{X}_n)_\mathbb{N}$ d'éléments de L^0 et $\overline{X} \in L^0$, on ait : $\lim_n d(\overline{X}_n, \overline{X}) = 0$ si, et seulement si, $(\overline{X}_n)_\mathbb{N}$

converge $\mathbb{P}$ -presque-sûrement vers $\bar{X}$. On peut montrer qu'il existe une topologie associée à la convergence presque-sûre si, et seulement si, la probabilité $\mathbb{P}$ est discrète (cf. [18] exemple 14-13, [1] p. 86 exercice V-10).

Un nouveau mode de convergence est introduit dans L^0 de la façon suivante :

On dit que $(\bar{X}_n)_\mathbb{N}$ converge en probabilité vers $\bar{X}$ si, pour tout $\delta > 0$, $\lim_n \mathbb{P}(|X_n - X| \geq \delta) = 0$.

On note alors $(\bar{X}_n)_\mathbb{N} \xrightarrow{\mathbb{P}} \bar{X}$ et $\bar{X} = \mathbb{P}\text{-}\lim_n(\bar{X}_n)$.

L'application $d : (\bar{X}, \bar{Y}) \in L^0 \times L^0 \mapsto d(\bar{X}, \bar{Y}) := \inf\{\delta > 0, \mathbb{P}(|X - Y| \geq \delta) \leq \delta\}$ définit une métrique sur L^0 appelée **métrique de Ky-Fan**. De plus (L^0, d) est complet et $(\bar{X}_n)_\mathbb{N}$ converge en probabilité vers $\bar{X}$ si, et seulement si, $\lim_n d(\bar{X}_n, \bar{X}) = 0$ (cf. [1] p. 86 exercice V-9).

La topologie d'espace vectoriel normé de L^p permet de définir un autre mode de convergence pour les suites d'éléments de L^0 qui sont dans L^p .

Si $p \in [1, +\infty]$, on dit que $(\bar{X}_n)_\mathbb{N}$ converge dans L^p vers $\bar{X}$ si $\bar{X} \in L^p$ et pour tout $n \in \mathbb{N}$, $\bar{X}_n \in L^p$ avec $\lim_n \|\bar{X}_n - \bar{X}\|_{L^p} = 0$.

On note alors $(\bar{X}_n)_\mathbb{N} \xrightarrow{L^p} \bar{X}$ et $\bar{X} = L^p\text{-}\lim_n(\bar{X}_n)$.

Proposition 3 - 1 1. Les définitions ci-dessus sont cohérentes, i.e. ne dépendent pas des représentants choisis dans les expressions.

2. Lorsqu'elle existe, la p.s.-limite (resp. $\mathbb{P}$ -limite, L^p -limite) d'une suite d'éléments de L^0 est unique dans L^0 .

On est amené à commettre certains abus de langage ou d'écriture. Ces abus sont très courants en théorie de la mesure ou en probabilités et doivent être utilisés en connaissance de cause si on veut éviter des expressions très lourdes ou des erreurs. On note souvent de la même façon un élément de L^0 et un de ses représentants. Ainsi on écrit $X \in L^p$ au lieu de $\bar{X} \in L^p$. De même on écrit $(X_n)_\mathbb{N} \xrightarrow{\mathbb{P}\text{-p.s.}} X$ au lieu de $(\bar{X}_n)_\mathbb{N} \xrightarrow{\mathbb{P}\text{-p.s.}} \bar{X}$ (même chose pour les autres convergences). On écrira aussi $\|X\|_{L^p}$ au lieu de $\|\bar{X}\|_{L^p}$, " X est égale à Y dans L^0 " ou " X est égale à Y p.s." pour $\bar{X} = \bar{Y}$. Enfin si $\mathcal{P}$ est une propriété d'éléments de L^0 , on dira qu'il existe "une unique v.a. X dans L^0 vérifiant $\mathcal{P}$ " ou que " X vérifiant $\mathcal{P}$ est unique presque-sûrement" pour exprimer qu'il existe un unique élément $\bar{X} \in L^0$ avec X v.a. vérifiant $\mathcal{P}$.

A partir de maintenant on fera les abus de notation signalés ci-dessus.

Donnons un résultat permettant d'effectuer des opérations sur les limites en probabilité et p.s. (cf. [1] p. 84 exercice V-4) :

Proposition 3 - 2 Soit f une application continue de $\mathbb{R}^d$ dans $\mathbb{R}$. Si la suite de v.a. à valeurs $\mathbb{R}^d$ $(X_n)_\mathbb{N}$ converge en probabilité (resp. p.s.) vers la v.a. X à valeurs

$\mathbb{R}^d$, alors la suite de v.a.r. $(f(X_n))_{\mathbb{N}}$ converge en probabilité (resp. p.s.) vers la v.a.r. $f(X)$.

Cette proposition n'est plus vraie si l'application f n'est plus supposée continue (cf. [18] exemple 14-11).

L'inégalité de Tchébycheff est utile dans l'étude des questions de convergence en probabilité. Le lemme de Borel-Cantelli, quant à lui, est souvent utilisé dans les questions de convergence p.s. grâce à la proposition suivante (cf. [2] exercice chapitre IV) :

Proposition 3 - 3 1. La suite $(X_n)_{\mathbb{N}}$ converge p.s. vers la v.a.r. X si, et seulement si, pour tout $\varepsilon > 0$, $IP(\limsup_n \{|X_n - X| \geq \varepsilon\}) = 0$.

2. Si pour tout $\varepsilon > 0$, la série $\sum_n IP(\{|X_n - X| \geq \varepsilon\})$ converge dans $\mathbb{R}$, alors la suite $(X_n)_{\mathbb{N}}$ converge p.s. vers la v.a.r. X .

3. Si la suite d'événements $(\{|X_n - X| \geq \varepsilon\})_{\mathbb{N}}$ est indépendante, alors la suite $(X_n)_{\mathbb{N}}$ converge p.s. vers la v.a.r. X si, et seulement si, pour tout $\varepsilon > 0$, la série $\sum_n IP(\{|X_n - X| \geq \varepsilon\})$ converge dans $\mathbb{R}$.

On remarquera que dans le troisième item l'indépendance est une hypothèse essentielle. On trouvera un contre-exemple dans ([18] exemple 14-4). Le premier item souligne le lien existant entre les convergences p.s. et en probabilité en s'énonçant :

Proposition 3 - 4 La suite $(X_n)_{\mathbb{N}}$ converge p.s. vers X si, et seulement si, la suite $(M_n)_{\mathbb{N}}$, définie pour tout $n \in \mathbb{N}$ par $M_n := \sup_{k \geq n} |X_k - X|$, converge en probabilité vers 0.

On ne confondra pas la convergence dans L^p avec la convergence de la suite des espérances mathématiques. Plus précisément :

Proposition 3 - 5 Si la suite $(X_n)_{\mathbb{N}}$ converge dans L^p vers la v.a.r. X , alors la suite des espérances $(IE(X_n))_{\mathbb{N}}$ converge dans $\mathbb{R}$ vers $IE(X)$.

La réciproque est fausse, mais on a le résultat suivant (cf. [1] exercice V-5 p. 84) :

Proposition 3 - 6 Si la suite $(X_n)_{\mathbb{N}}$ converge p.s. vers la v.a.r. X et si, pour tout entier n , $X_n \geq 0$, alors $(X_n)_{\mathbb{N}}$ converge dans L^1 vers la v.a.r. X si, et seulement si, $(IE(X_n))_{\mathbb{N}}$ converge dans $\mathbb{R}$ vers $IE(X)$.

3.2 Comparaison des convergences dans L^0

Proposition 3 - 7 Si $1 < p < q < +\infty$ et si la suite $(X_n)_{\mathbb{N}}$ converge dans L^q vers la v.a.r. X , alors la suite $(X_n)_{\mathbb{N}}$ converge dans L^p vers la v.a.r. X .

La réciproque est fausse (cf. [18] exemple 14-5). La topologie de L^q contient la topologie induite sur L^q par celle de L^p , mais on n'a pas l'égalité des topologies.

Concernant le lien entre les convergences presque-sûre et dans L^p , en général la convergence presque-sûre d'une suite de v.a. de L^p n'entraîne pas la convergence de cette suite dans L^p (cf. [18] exemple 14-8). Cependant on obtient les deux résultats suivants :

Proposition 3 - 8 *Si $(X_n)_\mathbb{N}$ est dominée dans L^p , $1 \leq p < +\infty$, et converge presque-sûrement vers X , alors elle converge dans L^p vers X .*

La réciproque de cette proposition est fausse, la convergence dans L^p n'entraîne pas la convergence presque-sûre (cf. [18] exemple 14-7, [1] p. 83 exercice V-1-(3)), mais

Proposition 3 - 9 *Si $(X_n)_\mathbb{N}$ converge dans L^p vers X , $1 \leq p < +\infty$, alors il existe une sous-suite convergeant presque-sûrement vers X .*

Concernant le lien entre les convergences dans L^p et en probabilité :

Proposition 3 - 10 *Si $(X_n)_\mathbb{N}$ converge dans L^p vers X , $1 \leq p < +\infty$, alors elle converge en probabilité vers X .*

La réciproque est fausse, la convergence en probabilité n'entraîne pas la convergence dans L^p (cf. [18] exemple 14-6, [1] p. 83 exercice V-1-(5)). Cependant (cf. [1] p. 85 exercice V-6) :

Proposition 3 - 11 *Si $(X_n)_\mathbb{N}$ est dominée dans L^p , $1 \leq p < +\infty$ alors $(X_n)_\mathbb{N}$ converge en probabilité vers X si, et seulement si, $(X_n)_\mathbb{N}$ converge dans L^p vers X .*

L'hypothèse de domination peut être affaiblie en introduisant la notion de d'uniforme-intégrabilité (cf. [15] p.192, [5] p. 316, [9] p. 223).

Une partie $\mathcal{H}$ de L^p , $1 \leq p < +\infty$, est dite **uniformément intégrable dans L^p** , en abrégé **u.i.**, si

$$\lim_{a \rightarrow +\infty} \sup_{X \in \mathcal{H}} \mathbb{E} \left(|X|^p \mathbb{1}_{\{|X| > a\}} \right) = 0.$$

On montre alors :

Proposition 3 - 12 *Si $(X_n)_\mathbb{N}$ est uniformément intégrable dans L^p , $1 \leq p < +\infty$, alors $(X_n)_\mathbb{N}$ converge dans L^p vers X si, et seulement si, $(X_n)_\mathbb{N}$ converge en probabilité vers X .*

Comme application de la notion d'uniforme intégrabilité, on peut montrer (cf. [10] exercice VII-13) :

Proposition 3 - 13 *Si $(X_n)_\mathbb{N}$ est une suite de v.a.r. gaussiennes convergeant en probabilité vers une v.a.r. X , alors X est gaussienne et la suite $(X_n)_\mathbb{N}$ converge dans L^p vers X , pour tout $p \in [1, +\infty[$.*

Enfin le lien entre les convergences en probabilité et presque-sûre est précisé par :

Proposition 3 - 14 *Si $(X_n)_\mathbb{N}$ converge presque-sûrement vers X , alors $(X_n)_\mathbb{N}$ converge en probabilité vers X .*

Ce dernier résultat n'est plus vrai si la mesure n'est pas finie ([?] p. 85 exercice V-8). Quant à la réciproque, elle n'est pas vraie, la convergence en probabilité n'entraîne pas la convergence presque-sûre (cf. [18] exemple 14-3). Cependant on a les résultats suivants :

Proposition 3 - 15 *Si $(X_n)_\mathbb{N}$ converge en probabilité vers X , alors il existe une sous-suite convergeant presque-sûrement vers X .*

Le résultat suivant montre à quelle condition les notions de convergences en probabilité et presque-sûre coïncident (cf. [1] p. 86 exercice V-10) :

Proposition 3 - 16 *Les notions de convergences en probabilité et presque-sûre coïncident si, et seulement si, la probabilité $\mathbb{P}$ est discrète.*

Dans le cas des séries de v.a.r. indépendantes, on peut montrer (cf. [10] ex V-8, [2] chapitre IV pour des v.a.r. positives, [6] p. 45, [2] chapitre IV pour des v.a.r. quelconques) :

Proposition 3 - 17 Lemme de Lévy

Si $(X_n)_\mathbb{N}$ est une suite indépendante de v.a.r., alors la série $\sum_n X_n$ converge en probabilité vers X si, et seulement si, elle converge presque-sûrement vers X .

3.3 Théorèmes-limites dans L^0

Donnons quelques résultats classiques précisant la convergence de suites ou de séries de v.a.r. Ces théorèmes sont connus sous le nom générique de **théorèmes-limites**. Parmi ceux-ci, les théorèmes portant sur la convergence en probabilité (resp. presque-sûre) des moyennes empiriques de v.a.r. sont connus sous le nom de **lois faibles des grands nombres** (resp. **lois fortes des grands nombres**).

On appelle **moyenne empirique** des n v.a.r. $X_1, X_2, \dots, X_n$ la v.a.r.

$$\frac{1}{n}(X_1 + X_2 + \dots + X_n).$$

Dans le cas où $(X_n)_\mathbb{N}$ est une suite indépendante de v.a.r. de Bernoulli de paramètre $p \in]0, 1[$, i.e. $\mathbb{P}_{X_n} := p\delta_1 + (1-p)\delta_0$, la loi faible des grands nombres est une conséquence immédiate du résultat plus précis suivant (cf. [8] pp 32-33) :

Proposition 3 - 18 Théorème des grandes déviations pour les v.a.r. de Bernoulli

Si $(X_n)_N$ est une suite indépendante de v.a.r. de Bernoulli de même paramètre $p \in]0, 1[$, alors, pour tout $\varepsilon > 0$, il existe une constante $C_\varepsilon > 0$ telle que, pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - p\right| \geq \varepsilon\right) \leq 2 \exp(-nC_\varepsilon).$$

En réalité cette proposition permet de montrer que la loi est forte (cf. [8] p. 52) :

Proposition 3 - 19 Loi forte des grands nombres pour des v.a.r. de Bernoulli

Si $(X_n)_N$ est une suite indépendante de v.a.r. de Bernoulli de même paramètre $p \in]0, 1[$, alors la suite des moyennes empiriques $(\frac{X_1 + X_2 + \dots + X_n}{n})_{N^*}$ converge presque-sûrement vers p .

Le théorème des grandes déviations pour les v.a.r. de Bernoulli est un cas particulier du théorème plus général suivant qui permet aussi, sous ses hypothèses, de démontrer grâce au lemme de Borel-Cantelli une loi forte (cf. [10] exercice V-14).

Pour cela introduisons une définition. On appelle **transformée de Laplace** d'une v.a.r. X l'application M_X de $\mathbb{R}$ dans $[0, +\infty]$ définie, pour tout $t \in \mathbb{R}$, par $M_X(t) := \mathbb{E}(e^{tX})$.

Proposition 3 - 20 Théorème des grandes déviations de Chernoff

Soit $(X_n)_N$ une suite indépendante de v.a.r. intégrables et de même loi. On suppose qu'il existe $a \in \mathbb{R}$ tel que $\mathbb{E}(X_1) < a$, $\mathbb{P}(\{X_1 > a\}) > 0$ et que la transformée de Laplace de X_1 est finie dans un voisinage de 0. Alors

$$\lim_n \frac{1}{n} \ln \mathbb{P}\left(\frac{X_1 + X_2 + \dots + X_n}{n} \geq a\right) = - \inf_{t \in \mathbb{R}} \{\ln M_{X_1}(t) - at\}.$$

L'application qui, à tout réel u associe le nombre réel $\inf_{t \in \mathbb{R}} \{\ln M_X(t) - ut\}$ s'appelle la **transformée de Cramer** de la v.a.r. X (cf. [8] p. 78 et 102).

Plus généralement on peut caractériser les suites de v.a.r. indépendantes et de même loi qui vérifient la loi forte des grands nombres (cf. [9] pp 51-55 et 43, [15] p. 202, [6] p. 51) :

Proposition 3 - 21 Loi forte des grands nombres de Kolmogorov

Si $(X_n)_N$ est une suite indépendante de v.a.r. de même loi, alors la suite des moyennes empiriques $(\frac{X_1 + X_2 + \dots + X_n}{n})_{N^*}$ converge presque-sûrement dans $\mathbb{R}$ si, et seulement si, $X_1 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Dans ce cas la p.s.-limite est $\mathbb{E}(X_1)$.

On notera que des résultats de ce type peuvent être démontrés sans l'hypothèse de l'identité des lois. Par exemple (cf. [10] exercices V-9, V-5) :

Proposition 3 - 22 Loi forte des grands nombres pour des v.a.r. uniformément bornées

Si $(X_n)_N$ est une suite de v.a.r. indépendantes, centrées et telles qu'il existe $M > 0$ vérifiant, pour tous $n \in \mathbb{N}$ et $\omega \in \Omega$, $|X_n(\omega)| \leq M$, alors la suite des moyennes empiriques $(\frac{X_1+X_2+\dots+X_n}{n})_{N^*}$ converge presque-sûrement vers 0.

Dans [9] p. 54 et [2] problème du chapitre IV, on trouvera une démonstration, proposée par Etemadi, de condition suffisante de loi forte des grands nombres pour des v.a.r. indépendantes deux à deux (et non mutuellement indépendantes) intégrables et de même loi. En revanche, elle cesse d'être vraie si les v.a.r. sont seulement supposées non-corrélées deux à deux.

On trouvera des applications de la loi des grands nombres à des démonstrations de théorèmes d'analyse dans [10] exercice V-12, [8] p. 103, [5] p. 170 et 261, [2] chapitre IV.

D'autres théorèmes portent sur la convergence des séries de v.a.r. indépendantes. Par application de la loi du Tout-ou-Rien, il vient (cf. [10] exercices V-1 et V-3, [2] problème du chapitre II) :

Proposition 3 - 23 Si $(X_n)_N$ est une suite indépendante de v.a.r., alors la série de v.a.r. $\sum_n X_n$ converge-p.s. dans $\mathbb{R}$ ou diverge-p.s. et la série entière à coefficients aléatoires $\sum_n X_n z^n$ a un rayon de convergence p.s. constant.

Par application de l'inégalité de Kolmogorov on obtient une condition suffisante (cf. [15] p. 201, [9] p. 49), mais non nécessaire (cf. [18] exemple 18-1-i) de convergence p.s. d'une série de v.a.r. indépendantes :

Proposition 3 - 24 Soit $(X_n)_N$ une suite indépendante de v.a.r. centrées. Si la série $\sum_n \sigma_{X_n}^2$ converge dans $\mathbb{R}$, alors la série de v.a.r. $\sum_n X_n$ converge presque-sûrement dans $\mathbb{R}$.

On peut caractériser les séries de v.a.r. indépendantes qui convergent p.s. dans $\mathbb{R}$ par (cf. [10] exercice VIII-11, [9] p. 50 et 102) :

Proposition 3 - 25 Théorème de Kolmogorov des trois séries

Soit $(X_n)_N$ une suite de v.a.r. indépendantes. La série de v.a.r. $\sum_n X_n$ converge presque-sûrement dans $\mathbb{R}$ si, et seulement si, il existe $a > 0$ tel que les trois séries $\sum_n \mathbb{P}(|X_n| > a)$, $\sum_n \mathbb{E}(X_n^{(a)})$ et $\sum_n \sigma_{X_n^{(a)}}^2$ convergent dans $\mathbb{R}$ où on a posé $X_n^{(a)} := X_n \mathbb{1}_{\{|X_n| \leq a\}}$.

Terminons par un résultat précisant le comportement des sommes de v.a.r. indépendantes et de même loi (cf. [6] p. 64 et 291, [9] p. 394) :

Proposition 3 - 26 Loi du logarithme itéré de Hartman-Wintner (1941)

Si $(X_n)_N$ est une suite indépendante de v.a.r. de même loi, de carré intégrable,

centrées et de variance σ^2 , alors, pour IP -presque-tout $\omega \in \Omega$,

$$\limsup_{n \rightarrow +\infty} \frac{X_1(\omega) + X_2(\omega) + \dots + X_n(\omega)}{\sqrt{2n\sigma^2 \ln(\ln n)}} = 1.$$

Chapitre 4

Convergence de lois de probabilité

Dans ce chapitre on notera $M_1(\mathbb{R})$, ou plus simplement M_1 , l'ensemble des mesures de probabilité sur $\mathbb{R}$. On se propose de munir cet ensemble d'une structure topologique appelée **topologie de la convergence étroite**. On notera $C_b(\mathbb{R})$ (resp. $C_k(\mathbb{R})$) l'espace fonctionnel des applications, appelées **fonctions-test**, continues et bornées (resp. continues et à support compact) de $\mathbb{R}$ dans $\mathbb{R}$.

4.1 Convergence étroite dans $M_1(\mathbb{R})$

On dira qu'une suite $(\mu_n)_N$ de probabilités sur $\mathbb{R}$ **converge étroitement** vers une probabilité μ sur $\mathbb{R}$ si, pour tout $f \in C_b(\mathbb{R})$, la suite de réels $(\int_{\mathbb{R}} f d\mu_n)_N$ converge dans $\mathbb{R}$ vers $\int_{\mathbb{R}} f d\mu$.

On note alors $(\mu_n)_N \xrightarrow{e} \mu$ et $\mu = e\text{-}\lim_n(\mu_n)$.

On peut munir l'ensemble $M_1(\mathbb{R})$ des probabilités sur $\mathbb{R}$ d'une structure d'espace métrique dont la topologie associée est celle de la convergence étroite (cf. [2] exercice chapitre V, [9] p. 75 et thème d'étude chapitre VI de [2] qui détaille le lien avec la notion de topologie faible).

Proposition 4 - 1 *L'application d de $M_1 \times M_1$ dans $\mathbb{R}^+$ définie par*

$$d(\mu, \nu) := \inf\{\varepsilon > 0; \forall x \in \mathbb{R}, F_\mu(x - \varepsilon) - \varepsilon \leq F_\nu(x) \leq F_\mu(x + \varepsilon) + \varepsilon\}$$

définit une métrique, dite de Lévy, sur M_1 telle que la suite $(\mu_n)_N$ de probabilités converge étroitement vers la probabilité μ si, et seulement si, $(d(\mu_n, \mu))_N$ converge vers 0.

Il peut être avantageux de pouvoir remplacer $C_b(\mathbb{R})$ par d'autres espaces de fonctions-test dans la définition de la convergence étroite. Par exemple (cf. [5] p. 179 et [2] chapitre VI-thème d'étude) :

Proposition 4 - 2 Soient $(\mu_n)_\mathbb{N}$ une suite de probabilités sur $\mathbb{R}$ et μ une probabilité sur $\mathbb{R}$. Alors la suite $(\mu_n)_\mathbb{N}$ converge étroitement vers μ sur $\mathbb{R}$ si, et seulement si, pour tout $f \in C_k(\mathbb{R})$, la suite réelle $(\int_{\mathbb{R}} f d\mu_n)_\mathbb{N}$ converge dans $\mathbb{R}$ vers $\int_{\mathbb{R}} f d\mu$.

On notera que, si on ne suppose pas que μ est une probabilité, la condition suffisante précédente est fausse (considérer la suite $(\delta_n)_\mathbb{N}$).

Donnons un critère de convergence étroite utilisant les f. de r. (cf. [15] p. 176 et [2] chapitre VI) :

Proposition 4 - 3 Critère des fonctions de répartition

La suite $(\mu_n)_\mathbb{N}$ de probabilités converge étroitement vers la probabilité μ si, et seulement si, la suite $(F_{\mu_n}(x))_\mathbb{N}$ converge vers $F_\mu(x)$ en tout point de continuité $x \in \mathbb{R}$ de la fonction de répartition F_μ .

On remarquera que la suite $(F_{\mu_n}(x))_\mathbb{N}$ ne converge pas nécessairement vers $F_\mu(x)$ aux points $x \in \mathbb{R}$ où la fonction de répartition F_μ de μ n'est pas continue (cf. [9] p. 68 exercice 2-3).

De plus si $(\mu_n)_\mathbb{N}$ converge étroitement vers la probabilité μ , et si A est un borélien de $\mathbb{R}$, la suite $(\mu_n(A))_\mathbb{N}$ ne converge pas nécessairement vers $\mu(A)$ (cf. [10] exercice IV-2 et IV-6, [18] exemple 16-3).

Plus généralement (cf. [9] p. 72 et 144, [14] p. 422, [5] p. 177) :

Proposition 4 - 4 La suite $(\mu_n)_\mathbb{N}$ de probabilités converge étroitement vers la probabilité μ si, et seulement si, pour toute application f de $\mathbb{R}$ dans $\mathbb{R}$ borélienne, bornée, d'ensemble de discontinuité μ -négligeable, la suite réelle $(\int_{\mathbb{R}} f d\mu_n)_\mathbb{N}$ converge dans $\mathbb{R}$ vers $\int_{\mathbb{R}} f d\mu$.

Dans le cas des probabilités discrètes il est pratique d'utiliser le critère suivant (cf. [10] exercice IV-4 et [2] chapitre VI) :

Proposition 4 - 5 Critère des probabilités discrètes

Soient, pour tout $n \in \mathbb{N}$, μ_n et μ des probabilités discrètes portées par $\mathbb{N}$. Alors la suite $(\mu_n)_\mathbb{N}$ converge étroitement vers μ si, et seulement si, pour tout $k \in \mathbb{N}$, la suite $(\mu_n(\{k\}))_\mathbb{N}$ converge vers $\mu(\{k\})$.

On notera qu'une suite de probabilités discrètes peut converger étroitement vers une probabilité non-discrète et que le résultat ci-dessus est faux si les probabilités sont portées par une partie dénombrable de $\mathbb{R}$ dont les points ne sont pas isolés (prendre $\mu_n := \delta_{\frac{1}{n}}$ et $\mu := \delta_0$).

Dans le cas des probabilités à densité la condition suivante est suffisante mais non nécessaire pour la convergence étroite (cf. [10] exercice IV-5) :

Proposition 4 - 6 Théorème de Scheffé

Soient, pour tout $n \in \mathbb{N}$, μ_n et μ des probabilités absolument continues sur $\mathbb{R}$ de densités respectives f_n et f par rapport à la mesure de Lebesgue λ . Si la suite des densités $(f_n)_{\mathbb{N}}$ converge λ -presque-partout vers la densité f , alors la suite $(\mu_n)_{\mathbb{N}}$ converge étroitement vers μ .

Si une suite de fonctions de répartition converge, sa limite n'est pas nécessairement une fonction de répartition (cf. [10] exercice IV-2). Cependant (cf. [10] exercice IV-8-a, [9] p. 73) :

Proposition 4 - 7 Théorème de Helly

Si $(F_n)_{\mathbb{N}}$ est une suite de fonctions de répartition de probabilités sur $\mathbb{R}$, alors il existe une application F de $\mathbb{R}$ dans $[0, 1]$ continue à droite, croissante et une sous-suite $(F_{n_k})_{k \in \mathbb{N}}$ de la suite $(F_n)_{\mathbb{N}}$ telle que $(F_{n_k}(x))_{k \in \mathbb{N}}$ converge vers $F(x)$ en tout point $x \in \mathbb{R}$ où F est continue.

En particulier, ce résultat a pour conséquence :

Proposition 4 - 8 *La limite, si elle existe, d'une suite de fonctions de répartition est une application de $\mathbb{R}$ dans $[0, 1]$ continue à droite et croissante.*

Donnons maintenant une condition nécessaire et suffisante pour que la limite, si elle existe, d'une suite de fonctions de répartition soit la fonction de répartition d'une probabilité. Pour cela introduisons une notion nouvelle.

On dira qu'une partie $\mathcal{P}$ de $M_1(\mathbb{R})$ est **tendue** si, pour tout $\varepsilon > 0$, il existe un compact K de $\mathbb{R}$ tel que, pour tout $\mu \in \mathcal{P}$, $\mu(K) \geq 1 - \varepsilon$.

On a alors la caractérisation des parties tendues de M_1 (cf. [10] exercice IV-8) :

Proposition 4 - 9 Théorème de Prohorov

Une partie $\mathcal{P}$ de $M_1(\mathbb{R})$ est tendue si, et seulement si, toute suite d'éléments de $\mathcal{P}$ contient une sous-suite convergeant étroitement vers un élément de $M_1(\mathbb{R})$.

On notera que la sous-suite ne converge pas nécessairement dans $\mathcal{P}$. On a alors le corollaire (cf. [10] exercice IV-7) :

Proposition 4 - 10 *Soit $(F_{\mu_n})_{\mathbb{N}}$ une suite de fonctions de répartition convergeant simplement vers F . Alors F est la fonction de répartition d'une probabilité si, et seulement si, la partie $\{\mu_n, n \in \mathbb{N}\}$ est tendue.*

Il peut être intéressant de connaître une condition suffisante pour savoir si une suite est tendue (cf. [9] p. 74) :

Proposition 4 - 11 *Soit $(\mu_n)_{\mathbb{N}}$ une suite de probabilités sur $\mathbb{R}$. S'il existe une application borélienne φ de $\mathbb{R}$ dans $\mathbb{R}$, positive, telle que $\lim_{|x| \rightarrow +\infty} \varphi(x) = +\infty$ et $\sup_n (\int_{\mathbb{R}} \varphi d\mu_n) < +\infty$, alors la partie $\{\mu_n, n \in \mathbb{N}\}$ est tendue.*

Comme application du théorème de Prohorov on a (cf. [9] p. 82, [15] p. 178, [10] exercice IV-9) :

Proposition 4 - 12 Théorème de continuité de Lévy

Soit $(\mu_n)_\mathbb{N}$ une suite de probabilités sur $\mathbb{R}$. La suite des f.c. $(\Phi_{\mu_n})_\mathbb{N}$ converge simplement sur $\mathbb{R}$ vers une application φ de $\mathbb{R}$ dans $\mathbb{C}$ continue en 0 si, et seulement si, il existe une probabilité μ sur $\mathbb{R}$, de fonction caractéristique $\Phi_\mu = \varphi$, telle que la suite $(\mu_n)_\mathbb{N}$ converge étroitement vers μ .

L'hypothèse de la continuité de φ en 0 est essentielle (cf. [18] exemple 16-6). Dans la pratique cela donne un critère de convergence étroite à l'aide des fonctions caractéristiques (pour une démonstration directe cf. [5] p. 179) :

Proposition 4 - 13 Critère des fonctions caractéristiques

Soient $(\mu_n)_\mathbb{N}$ une suite de probabilités sur $\mathbb{R}$ et μ une probabilité sur $\mathbb{R}$. La suite $(\mu_n)_\mathbb{N}$ converge étroitement vers μ si, et seulement si, la suite des f.c. $(\Phi_{\mu_n})_\mathbb{N}$ converge simplement sur $\mathbb{R}$ vers la f.c. Φ_μ .

Une application φ de $\mathbb{R}$ dans $\mathbb{C}$ est dite **définie-positive** ou **de type-positif** si, pour tout $n \in \mathbb{N}$, tout n -uplet de réels $(t_1, t_2, \dots, t_n)$ et tout n -uplet de nombres complexes $(z_1, z_2, \dots, z_n)$,

$$\sum_{1 \leq i, j \leq n} \varphi(t_i - t_j) z_i \overline{z_j} \in \mathbb{R}^+.$$

Par application du théorème de Lévy on peut caractériser alors les fonctions caractéristiques par (cf. [10] exercice IV-24) :

Proposition 4 - 14 Théorème de Bochner (1932)

Une application φ de $\mathbb{R}$ dans $\mathbb{C}$ est la fonction caractéristique d'une probabilité sur $\mathbb{R}$ si, et seulement si, elle est continue en 0, définie-positive et vérifie $\varphi(0) = 1$.

Donnons une autre application classique du théorème de continuité de Lévy.

Si a et σ sont des réels avec $\sigma \neq 0$, on note $\mathcal{N}_1(a, \sigma^2)$ la **probabilité de Gauss-Laplace sur $\mathbb{R}$** i.e. la mesure sur $\mathbb{R}$ de densité $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-\frac{1}{2\sigma^2}(x-a)^2]$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$. On note $\mathcal{N}_1(a, 0)$ la mesure de Dirac au point a .

Proposition 4 - 15 *Soient $(a_n)_\mathbb{N}$ et $(\sigma_n)_\mathbb{N}$ deux suites de réels.*

Alors la suite $(\mathcal{N}_1(a_n, \sigma_n^2))_\mathbb{N}$ converge étroitement si, et seulement si, les suites $(a_n)_\mathbb{N}$ et $(\sigma_n)_\mathbb{N}$ convergent dans $\mathbb{R}$.

Dans ce cas la limite-étroite de $(\mathcal{N}_1(a_n, \sigma_n^2))_\mathbb{N}$ est la probabilité $\mathcal{N}_1(\lim_n a_n, \lim_n \sigma_n^2)$.

4.2 Convergence en loi de v.a.r.

Faisons le lien avec les v.a.r. (cf. [9] p. 70 et [2] chapitre VI-thème d'étude) :

Proposition 4 - 16 *Si $(\mu_n)_\mathbb{N}$ est une suite de probabilités convergeant étroitement vers la probabilité μ , alors il existe une suite de v.a.r. $(X_n)_\mathbb{N}$ convergeant presque-sûrement vers une v.a.r. X telle que $\mathbb{P}_X = \mu$ et, pour tout $n \in \mathbb{N}$, $\mathbb{P}_{X_n} = \mu_n$.*

Introduisons une nouvelle définition. Soient $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)_\mathbb{N}$ et $(\Omega, \mathcal{F}, \mathbb{P})$ une suite d'espaces de probabilités. On dit qu'une suite de v.a.r. $(X_n)_\mathbb{N}$ définies respectivement sur $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)_\mathbb{N}$ **converge en loi vers une v.a.r. X** définie sur $(\Omega, \mathcal{F}, \mathbb{P})$, si la suite des lois $(\mathbb{P}_{X_n})_\mathbb{N}$ converge étroitement vers la loi $\mathbb{P}_X$ de X .

La proposition précédente a donc pour corollaire :

Proposition 4 - 17 *Si une suite de v.a.r. $(X_n)_\mathbb{N}$ converge en loi vers une v.a.r. X , alors il existe une suite de v.a.r. $(Y_n)_\mathbb{N}$ et une v.a.r. Y telles que :*

1. *pour tout $n \in \mathbb{N}$, Y_n a même loi que X_n ,*
2. *la v.a.r. Y a même loi que X ,*
3. *la suite $(Y_n)_\mathbb{N}$ converge presque-sûrement (et en loi) vers la v.a.r. Y .*

Précisons maintenant le lien entre les convergences en loi et en probabilité.

Proposition 4 - 18 *Si une suite de v.a.r. $(X_n)_\mathbb{N}$ converge en probabilité vers une v.a.r. X , alors elle converge en loi vers X .*

La réciproque est fausse (cf. [10] exemple V-6-b). Cependant :

Proposition 4 - 19 *Une suite de v.a.r. $(X_n)_\mathbb{N}$ converge en loi vers une constante $a \in \mathbb{R}$ si, et seulement si, elle converge en probabilité vers a .*

Dans le cas général on ne peut pas effectuer d'opérations sur les limites-en-loi (cf. [10] exercice V-1). On a cependant (cf. [10] exercice V-7, [9] p. 72, [2] exercice chapitre V) :

Proposition 4 - 20 Théorème de Slutsky

Soient $(X_n)_\mathbb{N}$ une suite de v.a.r. convergeant en loi vers une v.a.r. X et $(Y_n)_\mathbb{N}$ une suite de v.a.r. convergeant en loi vers 0, alors $(X_n + Y_n)_\mathbb{N}$ converge en loi vers X et $(X_n Y_n)_\mathbb{N}$ converge en loi (et aussi en probabilité) vers 0.

4.3 Théorèmes-limites en loi

Donnons quelques théorèmes classiques concernant la convergence étroite des probabilités et la convergence en loi des v.a.r..

Notons comme indiqué ci-dessous les probabilités discrètes suivantes où $n \in \mathbb{N}^*$, $N \in \mathbb{N}^*$, $\lambda \in]0, +\infty[$, $p \in]0, 1[$ et $q := 1 - p$:

probabilité **binomiale** : $\mathcal{B}(n, p) := \sum_{k=0}^n C_n^k p^k q^{n-k} \delta_k$,

probabilité de **Bernoulli** : $\mathcal{B}(p) := \mathcal{B}(1, p)$,

probabilité de **Poisson** : $\mathcal{P}(\lambda) := \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$,

probabilité **hypergéométrique** : $\mathcal{H}(N, n, p) := \sum_{k=0}^n \frac{C_N^k C_{N-n}^{n-k}}{C_N^n} \delta_k$, où $n \leq N$ et p tel que $Np \in \mathbb{N}$.

On a alors (cf. [12] p. 263) :

Proposition 4 - 21 *Si $(p_n)_{n \in \mathbb{N}^*}$ est une suite d'éléments de $]0, 1[$ telle que $\lim_n (np_n) = \lambda$ où $\lambda \in]0, +\infty[$, alors $(\mathcal{B}(n, p_n))_{n \in \mathbb{N}^*}$ converge étroitement vers $\mathcal{P}(\lambda)$.*

En remarquant qu'une v.a.r. de loi binomiale $\mathcal{B}(n, p)$ a même loi que la somme de n v.a.r. indépendantes de loi $\mathcal{B}(p)$, la proposition précédente est un cas particulier du résultat suivant (cf. [9] p. 125 et 117, [2] exercice chapitre V) :

Proposition 4 - 22 Loi des événements rares

Soient $(X_k^{(n)})_{(k,n) \in \mathbb{N}^2}$ une suite de v.a.r. indépendantes à valeurs dans $\mathbb{N}$. On pose $p_k^{(n)} := IP(X_k^{(n)} = 1)$ et $\varepsilon_k^{(n)} := IP(X_k^{(n)} \geq 2)$. On suppose de plus que :

1. $\lim_n (p_1^{(n)} + p_2^{(n)} + \dots + p_n^{(n)}) = \lambda \in]0, +\infty[$,
2. $\lim_n (\max(p_1^{(n)}, p_2^{(n)}, \dots, p_n^{(n)})) = 0$,
3. $\lim_n (\varepsilon_1^{(n)} + \varepsilon_2^{(n)} + \dots + \varepsilon_n^{(n)}) = 0$.

Alors la suite de v.a.r. $(X_1^{(n)} + X_2^{(n)} + \dots + X_n^{(n)})_{n \in \mathbb{N}^*}$ converge en loi vers une v.a.r. de loi $\mathcal{P}(\lambda)$.

Donnons une approximation de la loi hypergéométrique (cf. [12] p. 262 et [2] exercice chapitre V) :

Proposition 4 - 23 *Soient $n \in \mathbb{N}^*$ et $p \in]0, 1[$. Notons $S := \{N \in \mathbb{N}; n \leq N \text{ et } Np \in \mathbb{N}\}$. Alors $(\mathcal{H}(N, n, p))_{N \in S}$ converge étroitement vers $\mathcal{B}(n, p)$.*

Beaucoup de théorèmes d'approximation en loi font intervenir la loi de Gauss-Laplace. Ce sont souvent des applications du théorème-limite central. On peut montrer directement ce théorème pour les v.a.r. de Bernoulli qu'on énonce sous la forme suivante (cf. [8] pp 35-38, [12] p. 266, [9] p. 66) :

Proposition 4 - 24 Théorème de De Moivre-Laplace

Soient $p \in]0, 1[$ et, pour tout $n \in \mathbb{N}^*$, une v.a.r. S_n de loi $\mathcal{B}(n, p)$. Alors la suite de v.a.r. $\left(\frac{1}{\sqrt{npq}}(S_n - np)\right)_{n \in \mathbb{N}^*}$ converge en loi vers une v.a.r. de loi $\mathcal{N}_1(0, 1)$.

Si X est une v.a.r. de carré intégrable, d'espérance m et d'écart-type $\sigma > 0$, la v.a.r.

$$X^* := \frac{1}{\sigma}(X - m)$$

s'appelle la v.a. centrée-réduite associée à X . On montre alors (cf. [8] p. 102 et 133, [9] p. 93) :

Proposition 4 - 25 Théorème-limite central

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. indépendantes, de carré intégrable et de même loi. Alors $\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right)_{n \in \mathbb{N}^*}$ converge en loi vers une v.a.r. de loi $\mathcal{N}_1(0, 1)$.

Comme corollaire de ce théorème, il vient :

Proposition 4 - 26 Soient $\lambda \in]0, +\infty[$ et, pour tout $n \in \mathbb{N}^*$, une v.a.r. S_n de loi $\mathcal{P}(n\lambda)$. Alors la suite de v.a.r. $\left(\frac{1}{\sqrt{n\lambda}}(S_n - n\lambda)\right)_{n \in \mathbb{N}^*}$ converge en loi vers une v.a.r. de loi $\mathcal{N}_1(0, 1)$.

Citons une généralisation du théorème-limite central (cf. [9] p. 98) :

Proposition 4 - 27 Théorème de Lindeberg-Feller

Soient $(X_k^{(n)})_{(k,n) \in \mathbb{N}^2}$ une suite de v.a.r. indépendantes, de carré intégrable et centrées. On suppose de plus que :

1. $\lim_n (\sigma_{X_1^{(n)}}^2 + \sigma_{X_2^{(n)}}^2 + \dots + \sigma_{X_n^{(n)}}^2) = \sigma^2 \in]0, +\infty[$,
2. pour tout réel $a > 0$, $\lim_n \sum_{k=0}^{k=n} \mathbb{E}[|X_k^{(n)}|^2 \mathbf{1}_{\{|X_k^{(n)}| > a\}}] = 0$.

Alors la suite de v.a.r. $(X_1^{(n)} + X_2^{(n)} + \dots + X_n^{(n)})_{n \in \mathbb{N}^*}$ converge en loi vers une v.a.r. de loi $\mathcal{N}_1(0, \sigma^2)$.

Les théorèmes-limites jouent un rôle théorique important en statistique et dans la vérification des modèles probabilistes de phénomènes aléatoires. En particulier, par application de la loi forte de grands nombres on montre (cf. [5] p. 191 exercice VII-23 et p. 356, [2] problème du chapitre V) :

Proposition 4 - 28 Théorème fondamental de la statistique

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. indépendantes et de même loi. Alors, pour IP-presque-tout $\omega \in \Omega$, la suite des lois empiriques de $(X_1, X_2, \dots, X_n)$,

$$\left(\frac{1}{n} \sum_{k=1}^{k=n} \delta_{X_k(\omega)}\right)_{n \in \mathbb{N}^*},$$

converge étroitement vers la loi de X_1 .

On en déduit que, pour $\mathbb{P}$ -presque-tout $\omega \in \Omega$, la suite $(F_{n,\omega}(x))_{\mathbb{N}^*}$, où $F_{n,\omega}$ est la fonction de répartition de la loi empirique de $(X_1, X_2, \dots, X_n)$ appelée **fonction de répartition empirique** de $(X_1, X_2, \dots, X_n)$, converge vers $F_{X_1}(x)$ en tout point x de continuité de la f.r. de X_1 .

En fait, on obtient le résultat plus précis suivant (cf. [8] p. 104, [9] p. 56) :

Proposition 4 - 29 Théorème de Glivenko-Cantelli

Soit $(X_n)_{\mathbb{N}}$ une suite de v.a.r. indépendantes et de même loi. Alors, pour $\mathbb{P}$ -presque-tout $\omega \in \Omega$, la suite $(F_{n,\omega})_{\mathbb{N}^*}$ des fonctions de répartition empiriques de $(X_1, X_2, \dots, X_n)$, converge uniformément sur $\mathbb{R}$ vers F_{X_1} .

Annexe : Programmes de probabilité de l'agrégation

I - Épreuve optionnelle écrite

(B.O.E.N. numéro 21 du 21 mai 1992)

C3. Composition de probabilités et statistiques

Cette épreuve portera sur les questions suivantes :

I. Concepts généraux et instruments du calcul des probabilités

Notion de catégorie d'épreuves, d'événement. Tribu d'événements. Espace probabilisable. Mesure de probabilité. Espace probabilisé.

Système de constituants (partition de l'espace probabilisé).

Épreuves composées : espace-produit, distribution conjointe et distributions marginales.

Notion de variable aléatoire. Loi de distribution, fonction de répartition, densité.

Valeur moyenne, variance, écart-type, moments. Types de variables aléatoires. Inégalité de Bienaïmé — Tchébichev. Exemples de loi de distribution : loi dégénérée, loi binomiale, loi de Poisson, loi uniforme sur un segment (a,b), loi de Laplace-Gauss, loi de Cauchy.

Système de n variables aléatoires : loi de distribution conjointe, lois marginales. Exemple : loi de Laplace-Gauss à n dimensions. Covariance, coefficient de corrélation, droites de régression : application aux lois de Laplace-Gauss à deux dimensions.

Fonction caractéristique ; fonction génératrice.

II. Probabilités conditionnelles. Indépendance stochastique

Probabilité conditionnelle d'un événement par rapport à un événement de probabilité non nulle. Mesure de probabilité conditionnelle.

Formule de Bayes de « probabilité des causes ».

Distribution conditionnelle dans le cas d'épreuves composées. Tirage sans remplacement, loi hypergéométrique.

Espérance mathématique conditionnelle d'une variable aléatoire par rapport à un événement, par rapport à un système de constituants, par rapport à une tribu. Distributions conditionnelles. Courbes de régression.

Couple d'événements indépendants. Indépendance de deux ou plusieurs systèmes de constituants. Indépendance de deux ou plusieurs tribus d'événements.

Épreuves indépendantes. Épreuves répétées, problème de Bernoulli.

Variables aléatoires indépendantes : somme de variables aléatoires, produit de convolution de leurs lois et produit de leurs fonctions caractéristiques. Loi du χ^2 , loi de Student.

III. Notion de convergence. Loi des grands nombres. Théorèmes limites.

Convergence en loi ; convergence en fonction caractéristique ; convergence faible. Équivalence entre ces trois types de convergence.

Notions de convergence stochastique, convergence presque sûre, convergence en moyenne d'ordre 2 pour une suite de variables aléatoires.

Lois (faible et forte) des grands nombres.

Loi du tout ou rien, lemme de Borel-Cantelli.

Théorèmes limites ; convergences vers la loi de Gauss (théorème de Moivre-Laplace) ; convergence vers la loi de Poisson.

IV. Statistiques

Notions élémentaires de statistique.

Problème de l'échantillonnage — Estimation des paramètres — Méthode du maximum de vraisemblance. Tests du χ^2 et de Student.

Comparaison entre une distribution expérimentale et une distribution théorique.

II - Épreuve orale : leçons de probabilités

(B.O.E.N. numéro 21 du 21 mai 1992)

.. Instructions

L'agrégation est un concours de recrutement d'enseignants dont la culture mathématique doit permettre de dominer l'ensemble des programmes du second degré et ceux des classes préparatoires aux grandes écoles.

L'oral permet d'apprécier :

— par la construction d'un plan sur un sujet donné, les qualités d'organisation et de synthèse des candidats ;

— par une discussion relative à ce plan, la précision et la rigueur de leur raisonnement, quel que soit le niveau de l'étude proposée ;

— par l'exposé détaillé d'une partie du plan, leurs qualités d'exposition et de présentation, sur l'importance desquelles il apparaît opportun d'insister.

Les modalités sont les suivantes pour chacune des deux épreuves orales :

1. Le candidat tire au sort une enveloppe contenant deux sujets au choix. À l'issue de trois heures de préparation, il indique au jury celui des deux sujets qu'il a choisi.

Pendant la préparation, le candidat peut utiliser les ouvrages qui se trouvent sur place (bibliothèque de l'agrégation). Il peut également utiliser des ouvrages de référence qu'il a apportés lui-même. Ces ouvrages doivent être imprimés, vendus dans le commerce, et ne pas comporter de notes manuscrites. Ils seront contrôlés au moment de l'épreuve par le jury qui peut s'opposer à l'utilisation de certains ouvrages s'il juge qu'il y a risque de dénaturer le travail de préparation.

La liste des livres de la bibliothèque de l'agrégation est publiée chaque année dans le rapport du concours précédent. Tout achat de livres entre deux sessions pour cette bibliothèque est signalé aux candidats avec leur convocation aux épreuves orales.

2. Sur le sujet choisi le candidat n'a pas à bâtir une leçon détaillée destinée à une classe d'un niveau déterminé ou correspondant à un nombre limité d'heures de cours. Il lui est demandé surtout une étude de synthèse construite à partir d'une base de connaissance ne dépassant pas les limites du programme d'oral. Le candidat a le libre choix du niveau auquel il place son exposé : le niveau d'une classe de terminale risque cependant d'être insuffisant et d'autre part les connaissances exposées doivent être réellement maîtrisées.

3. L'épreuve commence par la présentation, en quinze à vingt minutes, d'un plan d'étude qui ne doit être ni une énumération de paragraphes, ni un exposé complet avec développement des démonstrations.

Il s'agit de définir avec précision les notions introduites, de donner des énoncés complets des résultats fon-

damentaux, de citer des exemples et des applications, et d'insister sur l'enchaînement des idées.

4. Après la présentation du plan le candidat est invité à fournir au jury une liste d'au moins deux points qu'il juge importants dans son étude. C'est parmi ces points que le jury choisit le thème d'un exposé, qui peut être soit le développement détaillé d'une partie bien délimitée du plan, soit la démonstration d'un théorème, soit la présentation d'un exemple significatif. La netteté et la clarté de cet exposé, l'aisance et la sûreté avec lesquelles il est présenté, constituent pour le jury un facteur important d'appréciation.

5. L'exposé est suivi d'une discussion au cours de laquelle le jury s'assure de la solidité des connaissances du candidat sur les questions abordées dans le plan et l'exposé, et éventuellement sur tout autre point en rapport avec le sujet et figurant au programme de l'oral. Cette discussion permet ainsi au candidat de développer, de justifier et d'illustrer son point de vue, en même temps qu'elle met en valeur sa culture mathématique.

Un ou plusieurs exercices peuvent être proposés par le jury.

6. Les candidats sont invités, notamment pour illustrer et compléter une leçon, à utiliser leurs connaissances en matière de méthodes numériques, d'algorithme, et de programmation des ordinateurs.

Seconde épreuve (analyse, probabilités) (coefficient 4)

— le programme complémentaire suivant, consacré aux probabilités.

I. Espaces de probabilité

Notions d'épreuve, d'événement. Tribu d'événements.

Espace probabilisable. Mesure de probabilité. Espace probabilisé.

Cas où l'ensemble des événements élémentaires est fini ou dénombrable.

II. Indépendance et conditionnement

Probabilité conditionnelle d'un événement par rapport à un événement de probabilité non nulle. Mesure de probabilité conditionnelle.

Formule de Bayes de « probabilité des causes ».

Couple d'événements indépendants. Indépendance de deux ou plusieurs tribus d'événements.

III. Variables aléatoires

Notion de variable aléatoire à valeurs réelles. Loi de distribution ; fonction de répartition. On se limitera au d'une loi discrète et à celui d'une loi définie par densité continue par morceaux.

Valeur moyenne, variance, écart-type, moments. Inégalité de Bienaimé-Tchébichev.

Exemples de lois de distribution : loi dégénérée, loi binomiale, loi de Poisson, loi uniforme sur un segment b), loi exponentielle, loi de Laplace-Gauss.

Indépendance de deux ou plusieurs variables aléatoires. Produit de variables aléatoires indépendantes. Variance d'une somme.

IV. Théorèmes limites

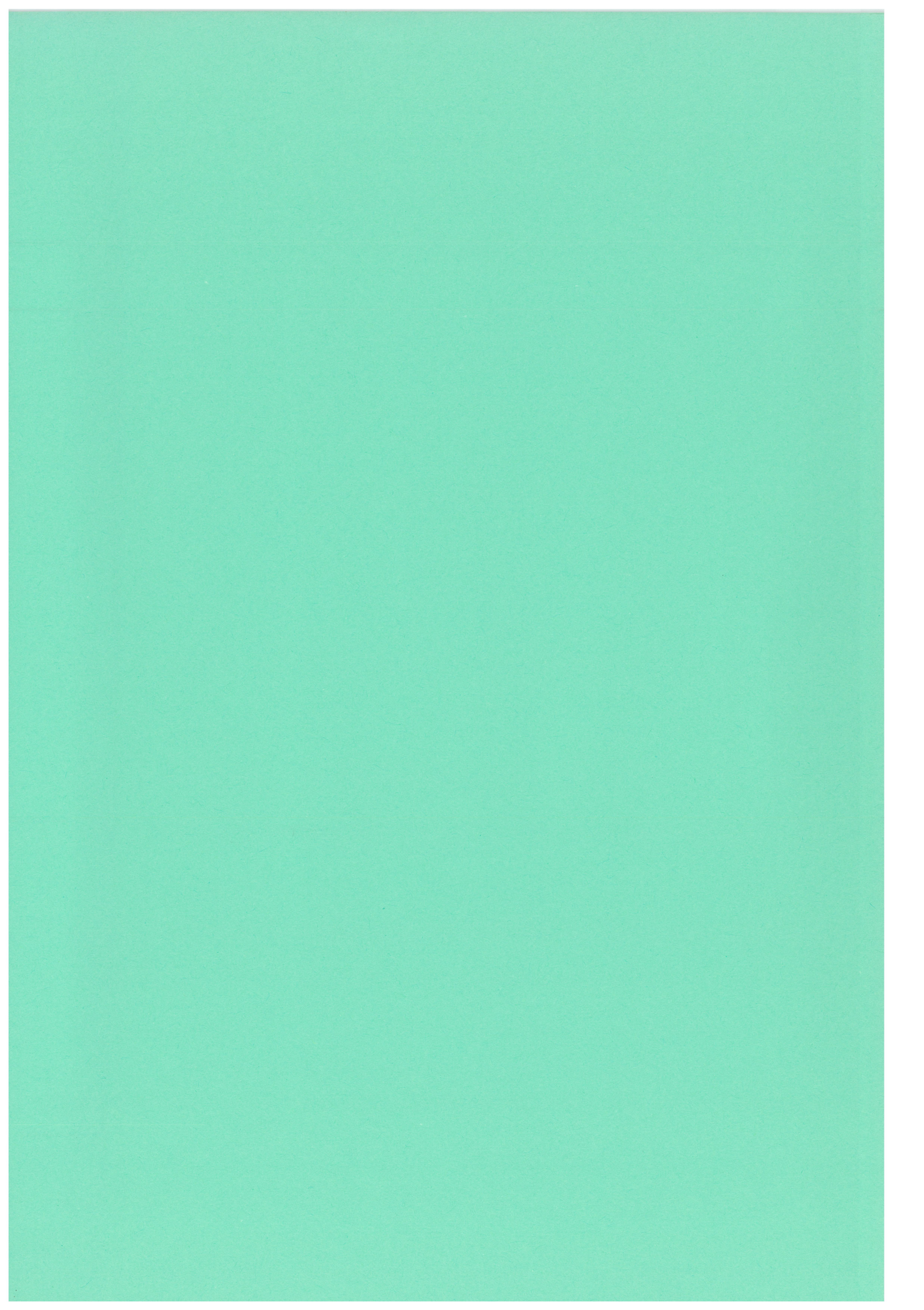
Loi faible des grands nombres.

Énoncé et exemples d'utilisation d'un théorème de convergence vers la loi normale.

Bibliographie

- [1] Ansel J.P.- Ducel Y., Exercices corrigés en théorie de la mesure et de l'intégration, Ellipses, 1995.
- [2] Ansel J.P.- Ducel Y., Exercices corrigés en théorie des probabilités, Ellipses, 1996.
- [3] Ansel J.P.- Ducel Y., Les probabilités dans les Concours d'Enseignement : Agrégation externe de mathématiques, Ellipses, En préparation.
- [4] Ansel J.P.- Ducel Y., Les probabilités dans les Concours d'Enseignement : Agrégation Interne et CAPES externe de mathématiques, Ellipses, En préparation.
- [5] Bouleau N., Probabilités de l'ingénieur : variables aléatoires et simulation, Hermann, 1986.
- [6] Breiman L., Probability, Addison-Wesley , 1968.
- [7] Brémaud P., Introduction aux probabilités : modélisation des phénomènes aléatoires, Springer-Verlag, 1988.
- [8] Dacunha-Castelle D.- Duflo M., Probabilités et statistiques, tome 1, Problèmes à temps fixe, Masson, 1990.
- [9] Durrett R., Probability : theory and examples, Wadsworth and Brooks/Cole, 1991.
- [10] Cottrell M.- Duhamel C.- Genon-Catalot V., Exercices de probabilités avec rappels de cours, Belin-DIA, 1980.
- [11] Gramain A., Intégration, Hermann-Méthodes, 1994.
- [12] Leboeuf C.- Roque J.L.- Guegand J., Cours de probabilités et de statistiques, Ellipses, 2ème édition 1983.
- [13] Leboeuf C.- Roque J.L.- Guegand J., Exercices corrigés de probabilités, Ellipses, 1987.
- [14] Marle C., Mesures et probabilités, Hermann, 1974.

- [15] Métivier M., Notions fondamentales de la théorie des probabilités, Dunod, 1972.
- [16] Métivier M., Probabilités : dix leçons d'introduction, Ellipses, 1987.
- [17] Revuz D., Mesure et intégration, Hermann-Méthodes, 1994.
- [18] Stoyanov J., Counterexamples in probability, John Wiley and Sons, 1989.



I.R.E.M. de Franche-Comté
UFR des Sciences et Techniques
16, route de Gray, La Bouloie
F-25030 BESANÇON cedex
Téléphone : 81.66.61.92 - Télécopie : 81.66.61.99
Cour. électr. : iremfc@math.univ-fcomte.fr

TITRE : Les probabilités à l'agrégation externe de mathématiques : guide pour une révision.

AUTEUR : Yves DUCCEL

DATE : Mars 1996

MOTS CLÉS : agrégation, mathématiques, probabilités, concours, préparation, enseignement.

RÉSUMÉ : Réalisée à partir d'un enseignement donné par l'auteur, dans le cadre de la préparation à l'agrégation externe de mathématiques, cette brochure a pour objectif de guider le candidat dans sa révision des notions de probabilité au programme de l'épreuve optionnelle écrite et de l'aider efficacement à maîtriser une bibliographie utile également dans la préparation des leçons de probabilité ou d'analyse de l'épreuve orale.

Format A4 - Nombre de pages : 42 - Poids : 130 g

IREM DE BESANÇON

Dépot Légal: 99/96

Numéro ISBN : 2-909963-08-X