Guide de survie : Probabilités

Remerciements : Tristan PHAM-MARIOTTI, Yann THANWERDAS, David EL MALIH, Camille RAFFIN, Santi GATILLON, Salim ABOUDOU, Julien PARIS, Jérôme NOWAK, Cécile GONTIER

1 Probabilités discrètes

1.1 Caractérisation par les singletons

Théorème 1 (Caractérisation par les singletons)

- *Dans l'espace probabilisé dénombrable $(\Omega, \mathcal{T}, \mathbb{P})$, $\mathbb{P}$ est entièrement caractérisée par ses valeurs sur les singletons $\mathbb{P}(\{\omega\})$ où $\omega \in \Omega$.*
- *Réciproquement, soit $\Omega = \{\omega_n / n \in \mathbb{N}\}$ et (p_n) suite de réels. Alors il existe une mesure de probabilité $\mathbb{P}$ sur la tribu $\mathcal{P}(\Omega)$ vérifiant $\mathbb{P}(\{\omega_n\}) = p_n$ si et seulement si $p_n \geq 0$ et $\sum_{n=0}^{\infty} p_n = 1$.*

1.2 Espérance

Définition 1 (Espérance)
X v.a.r. Si $\sum_{\omega\in\Omega} |X(\omega)|\mathbb{P}(\{\omega\}) < \infty$ alors :

$$E(X) = \sum_{\omega\in\Omega} X(w)\mathbb{P}(\{\omega\}) = \sum_{i\in\mathbb{N}} x_i \mathbb{P}(X = x_i)$$

Théorème 2 (Théorème de transfert)
Soit g fonction mesurable telle que $\sum_{i\in\mathbb{N}} |g(x_i)|\mathbb{P}(X = x_i) < \infty$. Alors :

$$E(g(X)) = \sum_{i\in\mathbb{N}} g(x_i)\mathbb{P}(X = x_i)$$

1.3 Probabilités conditionnelles

Définition 2 (Probabilité conditionnelle)
$(\Omega, \mathcal{F}, \mathbb{P})$ espace probabilisé et $B \in \mathcal{F}$ tel que $\mathbb{P}(B) > 0$. La probabilité de $A \in \mathcal{F}$ sachant B est :

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

Théorème 3 (Equation de partition)
$(\Omega, \mathcal{F}, \mathbb{P})$ espace probabilisé et $(E_n)_{n\in\mathbb{N}}$ partition dénombrable de Ω. On a pour tout $A \in \mathcal{F}$:

$$\mathbb{P}(A) = \sum_{n\in\mathbb{N}} \mathbb{P}(A|E_n)\mathbb{P}(E_n)$$

Application : Ex 1.6, ex 2.4

1.4 Quelques lois discrètes

Loi discrète uniforme : $\Omega = \{1, 2, ..., n\}, \mathcal{T} = \mathcal{P}(\Omega)$.

$$\forall k \in \Omega \quad \mathbb{P}(X = k) = \frac{1}{n}$$

Loi de Bernoulli : $\Omega = \{0, 1\}$, $\mathcal{T} = \mathcal{P}(\Omega), p \in [0, 1]$.

$$\mathbb{P}(X = 1) = p \quad \mathbb{P}(X = 0) = 1 - p$$

$$E(X) = p \qquad Var(X) = p(1 - p)$$

Loi binomiale : $\Omega = \{0, 1, ..., n\}$, $\mathcal{T} = \mathcal{P}(\Omega), p \in [0, 1]$.

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$E(X) = np \qquad Var(X) = np(1 - p)$$

Remarque : Une variable aléatoire X suivant une loi binomiale peut s'écrire comme somme de variables de Bernoulli.

Loi géométrique : $\Omega = \mathbb{N}, \mathcal{T} = \mathcal{P}(\Omega), p \in [0, 1]$.

$$\mathbb{P}(X = k) = p^k(1 - p) \quad E(X) = \frac{p}{1 - p} \quad Var(X) = \frac{p}{(1 - p)^2}$$

Loi de Poisson : $\Omega = \mathbb{N}, \mathcal{T} = \mathcal{P}(\Omega), \lambda > 0$.

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad E(X) = \lambda \quad Var(X) = \lambda$$

2 Probabilités continues : Généralités

Contrairement au cas discret, on ne peut plus caractériser une probabilité par sa valeur sur les singletons. On introduit 2 nouveaux outils (la fonction de répartition et la fonction caractéristique) pour la caractériser.

2.1 Caractérisation par la fonction de répartition

Définition 3 (Fonction de répartition)
Espace $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$.
La fonction de répartition de $\mathbb{P}$ est l'application définie par $\forall x \in \mathbb{R} \quad F(x) = \mathbb{P}(] - \infty, x])$.

Théorème 4 (Caractérisation par la fonction de répartition)
Sur l'espace $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$, la mesure de probabilité $\mathbb{P}$ est entièrement caractérisée par sa fonction de répartition.

Théorème 5 (Propriétés caractéristiques de la fonction de répartition)
Une fonction F est une fonction de répartion d'une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P})$ si et seulement si les 3 conditions suivantes sont vérifiées :

- *F est croissante*
- *F est continue à droite*
- $\lim_{x\to-\infty} F(x) = 0$ *et* $\lim_{x\to+\infty} F(x) = 1$

Dans certains cas, la fonction de répartition peut s'écrire comme une intégrale par rapport à la mesure de Lebesgue, on définit alors la fonction densité.

Définition 4 (Densité de $\mathbb{P}$)
Si une application $f : \mathbb{R} \to \mathbb{R}$ est positive, sommable et $\int_{\mathbb{R}} f(t)\lambda(dt) = 1$
Alors l'application $F : x \to \int_{]-\infty,x]} f(t)\lambda(dt)$ est la fonction de répartion d'une mesure de probabilité $\mathbb{P}$.
f est appelée densité de $\mathbb{P}$.

Application : Ex 4.6.1)

2.2 Espérance

Par rapport au cas discret, les sommes sont remplacées par des intégrales.

Définition 5 (Espérance)
X v.a.r. telle que $\int_\Omega |X(\omega)|\mathbb{P}(d\omega) < \infty$. Alors :

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

Théorème 6 (Théorème de transfert)
La loi de X v.a. est la mesure de probabilité P_X caractérisée par :

$$E(h(X)) = \int_E h(x)\mathbb{P}_X(dx)$$

pour toute application mesurable bornée $h : E \to \mathbb{R}$.

Remarque : On a bien une caractérisation, c'est à dire que si on trouve que $E(h(X)) = \int_E h(x)f(x)\lambda(dx)$ alors on sait que $P_X(dx) = f(x)\lambda(dx)$. (voir nombreux exercices dessus)

Application : Ex 2.1.3)b), ex 3.1, ex 4.1, ex 4.2, ex 5.2.2).b)

Théorème 7 (Inégalité de Markov)
X v.a.r. Pour tout réel $a > 0$ on a :

$$\mathbb{P}(|X| \geq a) \leq \frac{E(|X|)}{a}$$

2.3 Caractérisation par la fonction caractéristique

Définition 6 (Fonction caractéristique)
La fonction caractéristique de X v.a. à valeur dans $\mathbb{R}^N$ est l'application φ définie par :

$$\forall t \in \mathbb{R}^N \; \varphi(t) = E(e^{i<t,X>}) = \int_{\mathbb{R}^N} e^{i<t,x>}\mathbb{P}_X(dx)$$

Propriété 1 (Propriétés simples)

- *φ est continue*
- $\forall \lambda \in \mathbb{R} \quad \forall a \in \mathbb{R}^N \quad \varphi_{\lambda X+a}(t) = e^{iat}\varphi_X(\lambda t)$

Théorème 8 (Caractérisation par la fonction caractéristique)

$$\mathbb{P}_X = \mathbb{P}_Y \Leftrightarrow \varphi_X = \varphi_Y$$

2.4 Variance, covariance

Définition 7 (Variance)
$Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$

Définition 8 (Covariance)
$Cov(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$

Définition 9 (Matrice de covariances)
Soit $X = (X_1, ..., X_N)$ vecteur aléatoire. La matrice de covariances de X est la matrice $(cov(X_i, X_j))_{1\leq i,j \leq N}$.

Propriété 2

- *$Cov : L^2 \times L^2 \to \mathbb{R}$ est bilinéaire*
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
- *Toute matrice de covariances est symétrique et positive*

2.5 Quelques lois

Loi uniforme : $\Omega = \mathbb{R}, \mathcal{T} = \mathcal{B}(\mathbb{R}), (a,b) \in \mathbb{R}^2$.

$$\forall x \in \mathbb{R} \quad f_X(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$$

Loi exponentielle : $\Omega = \mathbb{R}, \mathcal{T} = \mathcal{B}(\mathbb{R}), \lambda > 0$.

$$\forall x \in \mathbb{R} \quad f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{\mathbb{R}+}(x) \quad E(X) = \frac{1}{\lambda} \quad Var(X) = \frac{1}{\lambda^2}$$

Loi normale : $\Omega = \mathbb{R}, \mathcal{T} = \mathcal{B}(\mathbb{R}), (m,\sigma) \in \mathbb{R} \times \mathbb{R}^+$.

$$\forall x \in \mathbb{R} \quad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}exp(-\frac{(x-m)^2}{2\sigma^2})$$

3 Indépendance

3.1 Généralités

Théorème 9 (Indépendance de deux v.a.)
X v.a. dans $(E, \mathcal{E})$ et Y v.a. dans $(F, \mathcal{F})$.
Les assertions suivantes sont équivalentes :

- *X et Y sont indépendantes*
- $\forall A \in \mathcal{E} \, \forall B \in \mathcal{F} \quad \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$
- *$\forall f : E \to \mathbb{R}$ et $g : F \to \mathbb{R}$ mesurables bornées $E(f(X)g(Y)) = E(f(X))E(g(Y))$.*
- $\mathbb{P}_{(X,Y)} = \mathbb{P}_X\mathbb{P}_Y$
- $F_{(X,Y)} = F_X F_Y$
- *$f_{(X,Y)} = f_X f_Y$ presque partout*
- $\varphi_{(X,Y)} = \varphi_X\varphi_Y$

Propriété 3 (Conséquences de l'indépendance)
Soit X et Y v.a. indépendantes. Alors :

- $E(XY) = E(X)E(Y)$
- $Var(X + Y) = Var(X) + Var(Y)$
- *La loi de $X + Y$ est le produit de convolution $\mathbb{P}_X * \mathbb{P}_Y$ et a pour fonction caractéristique $\varphi_{X+Y} = \varphi_X\varphi_Y$*

4 Vecteurs gaussiens

4.1 Définition et exemples

Définition 10 (Vecteur gaussien)
Un vecteur aléatoire $X = (X_1, ..., X_N)$ est gaussien si toute combinaison linéaire de ses coordonnées suit une loi normale, i.e. $\forall \lambda_1, ..., \lambda_N \in \mathbb{R}$ la v.a. $\sum_{i=1}^N \lambda_i X_i$ suit une loi normale.

Propriété 4 (Si condition d'indépendance en plus)
Si $X_1, ..., X_N$ suivent des lois normales et sont indépendantes Alors $X = (X_1, ..., X_N)$ est gaussien.

Remarque : Sans la condition d'indépendance, la propriété précédente est fausse et il faut avoir un contre exemple en tête !
Par exemple en prenant $X \hookrightarrow \mathcal{N}(0,1), \epsilon$ tel que $\mathbb{P}(\epsilon = -1) = \mathbb{P}(\epsilon = 1) = \frac{1}{2}$ et X indépendant de ϵ. On considère $Y = \epsilon X$ et on montre que (X,Y) n'est pas gaussien alors que $X \hookrightarrow \mathcal{N}(0,1)$ et $Y \hookrightarrow \mathcal{N}(0,1)$.

Application : Ex 5.3, ex 5.5.

4.2 Caractérisation de la loi d'un vecteur gaussien

Théorème 10 (Fonction caractéristique)
La fonction caractéristique d'un vecteur gaussien X est définie par :

$$\forall t \in \mathbb{R}^N \quad \varphi_X(t) = exp(i\sum_{j=1}^{N} \mu_j t_j - \frac{1}{2} \sum_{1 \leq j,k \leq N} t_j D_{j,k} t_k)$$

$\mu = (\mu_1, ..., \mu_N)$ *vecteur moyenne et* $D = (D_{j,k})$ *matrice de covariance de* X.

Remarque : Il faut savoir reconnaître cette fonction caractéristique !

Théorème 11 (Indépendance pour un vecteur gaussien)
$X = (X_1, X_2, ..., X_n)$ *vecteur gaussien. Les assertions suivantes sont équivalentes :*

- *Les* (X_k) *sont indépendants*
- *La matrice de covariances de* X *est diagonale*
- *Pour tout* $i \neq j$ $Cov(X_i, X_j) = 0$

Remarque : Ceci est vrai pour un vecteur gaussien mais pas en général, même pas pour 2 variables aléatoires suivant une loi normale !

Théorème 12 (Densité)
Soit $\mu \in \mathbb{R}^N$ *et* D *matrice* $N \times N$ *symétrique positive. Les assertions suivantes sont équivalentes :*

- $det(D) \neq 0$
- *La loi* $\mathcal{N}(\mu, D)$ *admet une densité*

Dans ce cas, la densité est

$$x \to \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{det(D)}} exp(-\frac{1}{2} < x - \mu, D^{-1}(x - \mu) >)$$

5 Convergence de variables aléatoires

5.1 Différents types de convergence

Définition 11 (Convergence presque sûre)
(X_n) *converge presque sûrement vers* X *s'il existe un événement* Ω^* *tel que* $P(\Omega^*) = 1$ *et* $\forall \omega \in \Omega^*$ $\lim_{n \to \infty} X_n(\omega) = X(\omega)$, *i.e.* $\mathbb{P}(\omega \in \Omega / \lim_{n \to \infty} X_n(\omega) = X(\omega)) = 1$.

Définition 12 (Convergence dans L^p)
(X_n) *converge dans* L^p *vers* X *si :*

- X *et* X_n *sont dans* L^p
- $\lim_{n \to \infty} E(|X_n - X|^p) = 0$

Définition 13 (Convergence en probabilité)
(X_n) *converge en probabilité vers* X *si pour tout* $\epsilon > 0$ $\lim_{n \to \infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$.

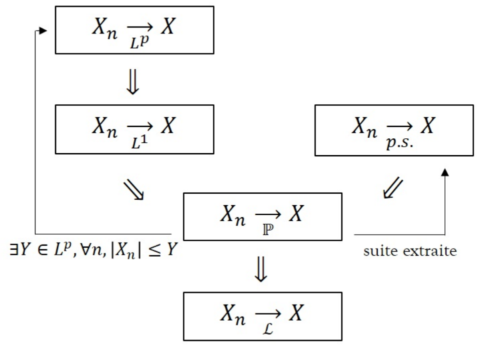
Propriété 5
Soit $f : \mathbb{R} \to \mathbb{R}$ *continue.*

- *Si* $X_n \to X$ *p.s. alors* $f(X_n) \to f(X)$ *p.s.*
- *Si* $X_n \xrightarrow{P} X$ *alors* $f(X_n) \xrightarrow{P} f(X)$
- $X_n \xrightarrow{P} X \Leftrightarrow \lim_{n \to \infty} E(\frac{|X_n - X|}{1 + |X_n - X|}) = 0$

5.2 Lien entre les convergences

Propriété 6

- *Si* $X_n \xrightarrow{L^p} X$ *alors* $X_n \xrightarrow{P} X$
- *Si* $X_n \xrightarrow{L^p} X$ *alors* $X_n \xrightarrow{L^1} X$
- *Si* $X_n \to X$ *p.s. alors* $X_n \xrightarrow{P} X$
- *Si* $X_n \xrightarrow{P} X$ *alors la suite admet une suite extraite qui converge presque sûrement vers* X
- *Si* $X_n \xrightarrow{P} X$ *et s'il existe* $Y \in L^p$ *telle que* $|X_n| \leq Y$ *alors* $X \in L^p$ *et* $X_n \xrightarrow{L^p} X$



Application : Ex 6.2

5.3 Loi forte des grands nombres

Théorème 13 (Loi forte des grands nombres)
Soit (X_n) *suite de v.a.r.i.i.d définies sur le même espace de probabilité et* $X_n \in L^2$.
Alors, en notant $S_n = \sum_{i=1}^{n} X_i$, *on a* $\lim_{n \to \infty} \frac{S_n}{n} = E(X_1)$ *p.s.*

6 Convergence de mesures, en loi

6.1 Définitions et propriétés

Définition 14 (Convergence faible)
(μ_n) *converge faiblement vers* μ *si pour toute application* $f : \mathbb{R}^N \to \mathbb{R}$ *continue et bornée on a* $\int_{\mathbb{R}^N} f(x)\mu_n(dx) \underset{n \to \infty}{\to} \int_{\mathbb{R}^N} f(x)\mu(dx)$.

Définition 15 (Convergence en loi)
(X_n) *converge en loi vers* X *si la mesure de probabilité* $\mathbb{P}_{X_n}$ *converge faiblement vers* $\mathbb{P}_X$, *i.e. pour toute application* $f : \mathbb{R}^N \to \mathbb{R}$ *continue et bornée* $\lim_{n \to \infty} E(f(X_n)) = E(f(X))$.

Propriété 7
Soit (X_n) *suite de v.a. définies sur le même espace.*
Si $X_n \xrightarrow{P} X$ *alors* $X_n \xrightarrow{\mathcal{L}} X$.

6.2 Lien avec les fonctions de répartition et caractéristiques

Théorème 14
Les assertions suivantes sont équivalentes :

- X_n *converge en loi vers* X
- *Pour tout* x *point de continuité de* F_X *on a* $\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$.
- *Pour tout* $x \in \mathbb{R}$ *on a* $\lim_{n \to \infty} \varphi_{X_n}(x) = \varphi_X(x)$

Application : Ex 6.3

6.3 Théorème Central Limite

Théorème 15
Soit (X_n) suite de v.a.i.i.d telle que $X_n \in L^2$. $\mu = E(X_i)$ et $\sigma = \sqrt{Var(X_i)}$. Alors :

$$\sqrt{n}\left(\frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\sigma}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

7 Espérance conditionnelle

Remarque : En proba, $L^2(\Omega, \mathcal{F}, \mathbb{P}) \subset L^1(\Omega, \mathcal{F}, \mathbb{P})$.

7.1 Définition de l'espérance conditionnelle dans L^2 et L^1 sachant une sous-tribu

Définition 16 (Espérance conditionnelle dans L^2 et L^1)
X v.a.r. de $L^2(\Omega, \mathcal{F}, \mathbb{P})$ (resp. $L^1(\Omega, \mathcal{F}, \mathbb{P})$) et $\mathcal{G}$ sous tribu de $\mathcal{F}$.
L'espérance conditionnelle de X sachant $\mathcal{G}$ est la variable aléatoire $Y \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ (resp. $L^1(\Omega, \mathcal{G}, \mathbb{P})$) telle que :

$$\forall Z \in L^2(\Omega, \mathcal{G}, \mathbb{P})\,(resp. L^1(\Omega, \mathcal{G}, \mathbb{P})) \quad E(XZ) = E(YZ)$$

Théorème 16 (Espérance conditionnelle dans L^2 = projection orthogonale)
L'espérance conditionnelle de $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ est la projection orthogonale de X sur $L^2(\Omega, \mathcal{G}, \mathbb{P})$.

7.2 Propriétés

Propriété 8
X v.a.r. de L^1 (ou L^2) et $\mathcal{G}$ sous-tribu de $\mathcal{F}$.

- *$X \to E(X|\mathcal{G})$ est linéaire*
- *$X \geq 0$ p.s. $\Longrightarrow E(X|\mathcal{G}) \geq 0$ p.s.*
- *$E(E(X|\mathcal{G})) = E(X)$*
- *Si X, Y et XY sont intégrables et X $\mathcal{G}$-mesurable*
  Alors $E(XY|\mathcal{G}) = XE(Y|\mathcal{G})$ p.s.
- *$\mathcal{G}$ et $\mathcal{H}$ sous-tribus de $\mathcal{F}$ avec $\mathcal{G} \subset \mathcal{H}$.*
  Alors $\forall X \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \quad E(E(X|\mathcal{H})|\mathcal{G}) = E(X|\mathcal{G})$ p.s.

Application : Ex 7.1

Théorème 17
- *Convergence monotone : Si $X_n \geq 0$ et (X_n) croît vers X p.s. Alors $\lim_{n\to\infty} E(X_n|\mathcal{G}) = E(X|\mathcal{G})$ p.s.*
- *Convergence dominée : Si $X_n \to X$ p.s. et s'il existe $Z \in L^1$ telle que $|X_n| \leq Z$ Alors $\lim_{n\to\infty} E(X_n|\mathcal{G}) = E(X|\mathcal{G})$ p.s.*

7.3 Espérance conditionnelle sachant une variable aléatoire

Définition 17
X v.a.r. On définit $E(X|Y) = E(X|\sigma(Y))$.

Théorème 18 (Espérance conditionnelle et densité)
(X, Y) vecteur aléatoire admettant une densité $f_{(X,Y)}$.
On suppose $X \in L^1$ et $f_Y(y) = \int f_{(X,Y)}(x, y)dx > 0$.
On pose $f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)}$ et $h(y) = \int_{\mathbb{R}} x f_{X|Y=y}(x)dx$.
Alors $E(X|Y) = h(Y)$ p.s.

Application : Ex 7.2

8 Processus stochastiques (optionnel)

Définition 18 (Processus stochastique)
Un processus stochastique est une collection de variables aléatoires $(X_t)_{t\in\mathcal{T}}$ définies sur le même espace et à valeurs dans le même espace.

Définition 19 (Filtration)
Une filtration est une suite croissante (pour l'inclusion) de sous-tribus de $\mathcal{F}$.
Le processus $(X_n)_{n\in\mathbb{N}}$ est adapté à la filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ si pour tout $n \in \mathbb{N}, X_n$ est $\mathcal{F}_n$-mesurable.

Définition 20 (Marche aléatoire)
Une marche aléatoire est un processus $(S_n)_{n\in\mathbb{N}}$ tel que $(S_n - S_{n-1})_{n\in\mathbb{N}}$ sont indépendants et stationnaires.

Définition 21 (Martingale)
Un processus $(X_n)_{n\in\mathbb{N}}$ muni d'une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une martingale si :

- *$(X_n)_{n\in\mathbb{N}}$ est adapté à $(\mathcal{F}_n)_{n\in\mathbb{N}}$*
- *$\forall n \in \mathbb{N}, \quad X_n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$*
- *$\forall n \in \mathbb{N}, \quad X_n = E(X_{n+1}|\mathcal{F}_n)$ p.s.*

9 Compléments (ECP+R)

Théorème 19 (Lemme de Borel-Cantelli)
$(A_n)_{n\in\mathbb{N}}$ suite de $\mathcal{F}$. On pose $\underset{n\to\infty}{limsup}(A_n) = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k$.

- *Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$*
  Alors $\mathbb{P}(limsup(A_n)) = 0$
- *Si $\mathbb{P}(limsup(A_n)) = 0$ et les A_n indépendants*
  Alors $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$

Définition 22 (Tribu de queue)
$(X_n)_{n\in\mathbb{N}}$ suite de v.a. $\mathcal{T}_n = \sigma(X_k, k \geq n)$.
La tribu de queue des X_n est la tribu $\mathcal{T}_\infty = \bigcap_{n\in\mathbb{N}} \mathcal{T}_n$

Théorème 20 (Loi du 0-1)
$(X_n)_{n\in\mathbb{N}}$ suite de v.a. ***indépendantes****.*
Si $A \in \mathcal{T}_\infty$ alors $\mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$.

Définition 23 (Temps d'arrêt)
Un temps d'arrêt est une v.a. $T : \Omega \to \mathbb{N} \cup \{\infty\}$ telle que $\forall n \in \mathbb{N} \quad \{T = n\} \in \mathcal{F}_n$

Théorème 21 (Théorème d'arrêt)
Soit (X_n) une $(\mathcal{F}_n)$-martingale et T temps d'arrêt ***borné*** *par rapport à $(\mathcal{F}_n)$.*
Alors $E(X_T) = E(X_0)$.