B-L = Barbe Ledoux
Co = Cottrell
/ Garet + BL + Cottrell
O1 = Ouvrard 1
O2 = Ouvrard 2

260 — **Espérance, variance et moments d'une variable aléatoire.**

X v.a réelle définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$

I. DEFINITIONS ET PREMIÈRES PROPRIÉTÉS.

1. Espérance d'une v.a. B-L p52 ou G p143

Def 1 : Si X est intégrable (i.e. $\int |X(\omega)| d\mathbb{P}(\omega) < \infty$) ou positive, l'espérance de X est $\mathbb{E}[X] = \int_\Omega X(\omega) d\mathbb{P}(\omega) = \int_\Omega X d\mathbb{P} \in \mathbb{R}$. On note $X \in L^1(\Omega)$

Rq 2 : Si $\mathbb{E}[X] = 0$ on dit que X est centrée.
- L'espérance représente la valeur moyenne.
- Comme $\mathbb{E}[X]$ n'est rien d'autre que l'intégrale (au sens de Lebesgue) de la fonction mesurable X par rapport à la mesure de probabilité $\mathbb{P}$ on déduit les propriétés de l'espérance de celle de l'intégration. En particulier, l'espérance est linéaire.

Ex 3 : Si X est constante p.s. alors $\mathbb{E}[X] = X$ p.s.

Thm 4 : Thm de transfert. Soit $\Phi : \mathbb{R} \to \mathbb{R}$ une fonction borélienne.
- Si Φ est à valeurs positives,
$$\mathbb{E}[\Phi(X)] = \int_\Omega \Phi(X(\omega)) d\mathbb{P}(\omega) = \int_{\mathbb{R}} \Phi(x) d\mathbb{P}_X(x).$$
(où $\mathbb{P}_X$ est la probabilité définie par : $\forall A \in \mathcal{B}(\mathbb{R})\ \mathbb{P}_X(A) = \mathbb{P}(X \in A)$.
- Si Φ est à valeurs quelconques, $\Phi(X) \in L^1(\Omega)$ ssi $\Phi \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ et dans ce cas $\mathbb{E}[\Phi(X)] = \int_{\mathbb{R}} \Phi(x) d\mathbb{P}_X(x)$.

Rq 5 : Pour tout $A \in \mathcal{B}(\mathbb{R})$, $\mathbb{E}[\mathbb{1}_A(X)] = \mathbb{P}(X \in A)$

Thm 6 : Inégalité de Jensen G p152
Si $\Phi : \mathbb{R} \to \mathbb{R}$ est convexe et X et $\Phi(X)$ sont intégrables, alors $\Phi(\mathbb{E}[X]) \leq \mathbb{E}[\Phi(X)]$.

p58
Thm 7 : Inégalité de Markov G p144
Si $X \in L^1(\Omega)$ et $t > 0$, alors $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[|X|]}{t}$

p60
Def 8 : Si $X = (X_1, \dots, X_d) \in \mathbb{R}^d$ est un vecteur aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$ tq $\forall i \in [\![1,d]\!]$ X_i est intégrable. On définit son espérance par $\mathbb{E}[X] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d]) \in \mathbb{R}^d$

Thm 9 : Caractérisation de l'indépendance par l'espérance
Une famille de v.a $(X_i)_{i \in I}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$ est mutuellement indépendantes si pour toute famille finie $J \subset I$ et toute famille de fonctions boréliennes $(\Phi_i)_{i \in J}$ tq $\Phi_i(X_i)$ soit intégrable, $i \in J$, on a
$$\mathbb{E}\left[\prod_{i \in J} \Phi_i(X_i)\right] = \prod_{i \in J} \mathbb{E}[\Phi_i(X_i)]$$

Cor 10 : Si $(X_1, \dots, X_n)$ est mutuellement indépendantes, $\mathbb{E}[X_1 \cdots X_n] = \mathbb{E}[X_1] \cdots \mathbb{E}[X_n]$

Prop 4* : Si X v.a positive, $\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X > t)$

2. Espérances de lois usuelles

a. Cas discret B-L p55 ou O1 p115 ou appendice p743 G p148

Prop 11 : Si X est discrète et intégrable, $\mathbb{E}[X] = \sum_{x \in \mathbb{R}} x \mathbb{P}(X = x)$

Rq 12 : On peut indicer sur $\operatorname{supp}(X) = \{x \in \mathbb{R},\ \mathbb{P}(X = x) > 0\}$

Ex 13 :
- Loi de Bernoulli : si $X \sim \mathcal{B}(p)$, alors $\mathbb{E}[X] = p$.
- Loi binomiale : si $X \sim \mathcal{B}(n, p)$ alors $\mathbb{E}[X] = np$
- Loi de Poisson : si $X \sim \mathcal{P}(\lambda)$, alors $\mathbb{E}[X] = \lambda$.
- Loi géométrique : si $X \sim \mathcal{G}(p)$ alors $\mathbb{E}[X] = 1/p$.
- Loi uniforme : si $X \sim \mathcal{U}(\{x_1, \dots, x_n\})$, alors $\mathbb{E}[X] = \frac{1}{n} \sum_{i=1}^n x_i$.

b. Cas à densité. B-L p54 ou O1 p190 ou appendice

On dit que X est intégrable si $x \mapsto |x| f_X(x) \in L^1(\mathbb{R})$

Prop 14 : Si X admet une densité f_X et est intégrable, $\mathbb{E}[X] = \int_{\mathbb{R}} x f(x) dx$ G p149

Ex 15 :
- Loi uniforme : si $X \sim \mathcal{U}([a,b])$, $\mathbb{E}[X] = \frac{a+b}{2}$
- Loi normale : si $X \sim \mathcal{N}(m, \sigma^2)$, $\mathbb{E}[X] = m$
- Loi exponentielle : si $X \sim \mathcal{E}(\lambda)$, $\mathbb{E}[X] = 1/\lambda$.

C-ex 16 : La loi de Cauchy de densité $f_X : x \mapsto \frac{1}{\pi(1+x^2)}$ n'a pas d'espérance.

3. Espérance conditionnelle. B-L p156

Def 17 : Soit $\mathcal{G}$ une sous tribu de $\mathcal{F}$, X intégrable. Alors il existe une p.s unique v.a., appelée espérance conditionnelle de X sachant $\mathcal{G}$, notée $\mathbb{E}[X|\mathcal{G}]$ tq
- $\omega \mapsto \mathbb{E}[X|\mathcal{G}](\omega)$ est $\mathcal{G}$-mesurable
- $\forall B \in \mathcal{G}$ $\int_B \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_B X d\mathbb{P}$ i.e $\mathbb{E}[X \mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \mathbb{1}_B]$

Prop 18 : L'espérance conditionnelle sachant $\mathcal{G}$ est linéaire croissante
- si $\Phi : \mathbb{R} \to \mathbb{R}$ convexe et $\Phi(X) \in L^1(\Omega)$, $\Phi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\Phi(X)|\mathcal{G}]$ p.s.
- si $\mathcal{G} \perp\!\!\!\perp \sigma(X)$, $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ p.s.
- si Y est $\mathcal{G}$ mesurable et $XY \in L^1$, $\mathbb{E}[XY|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] Y$ p.s.

Ex 19.
- Si $X_1 \sim \mathcal{B}(n_1, p)$, $X_2 \sim \mathcal{B}(n_2, p)$ et $X_1 \perp\!\!\!\perp X_2$ alors $\mathbb{E}[X_1 | X_1 + X_2] = \frac{n_1(X_1 + X_2)}{n_1 + n_2}$ p.s. Co exo 6.1
- Si $X_1 \sim \mathcal{P}(\lambda_1)$, $X_2 \sim \mathcal{P}(\lambda_2)$ et $X_1 \perp\!\!\!\perp X_2$ alors $\mathbb{E}[X_1 | X_1 + X_2] = \frac{\lambda_1}{\lambda_1 + \lambda_2}(X_1 + X_2)$ p.s.

II - MOMENTS D'UNE VARIABLE ALEATOIRE

(B-L p53) **Def 20** = Si $\int_\Omega |X|^p d\mathbb{P} < \infty$, on définit le moment d'ordre p de X par ou X positive. On note $X \in \mathcal{L}^p(\Omega)$
$\mathbb{E}[X^p] = \int_\Omega X^p d\mathbb{P}$

G p155

1 - Moments d'ordre 2 = Variance et covariance BL p56 G

Def 21 : Si $X \in \mathcal{L}^2(\Omega)$, la variance de X est le réel
$Var(X) = \mathbb{E}[(X-\mathbb{E}[X])^2]$. mesure la dispersion
L'écart type de X est $\sigma(X) = \sqrt{Var X}$

Rq 22 : • Si $\sigma(X) = 1$, on dit que X est réduite
• $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
• $Var(X) = \|X - \mathbb{E}[X]\|_2^2$ norme de l'espace de Hilbert $\mathcal{L}^2$.

Ex 23 : • $Var(X) = 0 \Rightarrow X = \mathbb{E}[X]$ p-s.
• $X \sim \mathcal{B}(n,p)$, alors $Var(X) = np(1-p)$.
• $X \sim \mathcal{N}(m, \sigma^2)$ alors $Var(X) = \sigma^2$.

Prop 24 : Si $\alpha \in \mathbb{R}$, $Var(X+\alpha) = Var(X)$, $Var(\alpha X) = \alpha^2 Var(X)$

G p160 **Thm 25** Inégalité de Tchebychev
Si $X \in \mathcal{L}^2$ et $t > 0$ $\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{Var(X)}{t^2}$

Thm 26 Inégalité de Cauchy-Schwarz
Si $X \in \mathcal{L}^2$, $Y \in \mathcal{L}^2$ alors $XY \in \mathcal{L}^1$ et $\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$

G p155 **Def 27** : La covariance de $X, Y \in \mathcal{L}^2$ est définie par
$Cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$.

Cor 28 : $Cov(X,Y) \leq \sigma(X)\sigma(Y)$ pour $X, Y \in \mathcal{L}^2$.

Ex 29 : Si (X,Y) admet une densité $f_{X,Y} : (x,y) \mapsto \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right)$
alors $Cov(X,Y) = 0$.

2 - Lien avec l'indépendance BL p80 G

G p158 **Def 30** : Si $X, Y \in \mathcal{L}^2$, on dit que X et Y sont non corrélées si
$Cov(X,Y) = 0$. (i.e $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$).

G p159 **Rq 31** indépendantes $\Rightarrow$ non corrélées.

C-ex 32 : Soit $X \sim \mathcal{N}(0,1)$, $Y = X^2$. Alors X et Y sont non corrélées mais pas indépendantes.

Thm 33 : Soit $X = (X_1, \dots, X_n) \in \mathbb{R}^n$ vecteur gaussien. Si les composantes de X sont 2 à 2 non corrélées alors $(X_1, \dots, X_n)$ est mutuellement indépendante.

Prop 34 Identité de Bienaymé.
Si $X_1, \dots, X_n$ sont deux à deux non corrélées alors $Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i)$ G p159

App 35 : Si $X \sim \mathcal{B}(n,p)$, $Var(X) = np(1-p)$.

App 36 : Théorème de Weierstrass. ZQ
Soit $f : [0,1] \to \mathbb{C}$ continue et $\omega : h \mapsto \sup\{|f(u) - f(v)|, |u-v| \leq h\}$
Pour $n \in \mathbb{N}^*$, on pose $B_n = \sum_{k=0}^n \binom{n}{k} X^k (1-X)^{n-k} f\left(\frac{k}{n}\right)$, alors
* (B_n) converge uniformément vers f sur $[0,1]$, $\exists C > 0$ $\forall n \in \mathbb{N}^*$ $\|f - B_n\|_\infty \leq C\omega(1/\sqrt{n})$
* Cette majoration est optimale : il existe f continue, il existe $\delta > 0$ $\forall n \in \mathbb{N}^*$ $\|f - B_n\|_\infty \geq \delta\omega(1/\sqrt{n})$

DVLPT

3 - Moments d'ordre p. BL p56 G

G p187 **Thm 37** - Inégalité de Hölder
Si $X \in \mathcal{L}^p$, $Y \in \mathcal{L}^q$ avec $p, q \geq 1$ $1/p + 1/q = 1$. Alors $XY \in \mathcal{L}^1$ et
$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{1/p} \mathbb{E}[|Y|^q]^{1/q}$

G p190 **Thm 3-8** : L'application $p \mapsto \mathbb{E}[|X|^p]^{1/p}$ est croissante.

App 39 : Si $1 \leq p \leq q$, $\mathcal{L}^q(\Omega, \mathcal{F}, \mathbb{P}) \subset \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ et la convergence dans $\mathcal{L}^q$ implique la convergence dans $\mathcal{L}^p$.

III UTILISATION DES MOMENTS

G p209

1 - Fonction génératrice. Ici X est à valeurs dans $\mathbb{N}$

O1 138 **Def 40** : Pour $s \in \mathbb{R}$, on pose $G_X(s) = \mathbb{E}[X^s]$ lorsque $s^X \in \mathcal{L}^1$.
G_X est appelée fonction génératrice de X. Elle est bien définie sur $[-1,1]$.

Prop 41 : • $\forall s \in [-1,1]$ $|G_X(s)| \leq 1$ et $G_X(1) = 1$.
• $\forall s \in [-1,1]$ $G_X(s) = \sum_{n=0}^{+\infty} p_n s^n$
• G_X est continue sur $[-1,1]$ et $\mathcal{C}^\infty$ sur $]-1,1[$.
• G_X caractérise la loi de X : $\forall n \in \mathbb{N}$ $\mathbb{P}(X=n) = \frac{G_X^{(n)}(0)}{n!}$

Ex 42 : • $X \sim \mathcal{B}(n,p)$: $G_X(s) = (ps+q)^n$, $s \in [-1,1]$.
• $X \sim \mathcal{P}(\lambda)$, $G_X(s) = \exp(\lambda(s-1))$, $s \in [-1,1]$

Prop 43 : Si $X \perp\!\!\!\perp Y$, $\forall s \in [-1,1]$ $G_{X+Y}(s) = G_X(s) G_Y(s)$.

App 44 : On ne peut pas truquer deux dés indépendants de manière à ce que la somme des points obtenus en les lançant soit équirépartie.

Cot Bxo 3-2.

Prop 45 = X admet un moment d'ordre $r \in \mathbb{N}^*$ ssi G_X est r-fois dérivable à gauche en 1.
Dans ce cas, $G_X^{(r)}(1^-) = \sum_{k=r}^{+\infty} k(k-1)\dots(k-r+1)\,\mathbb{P}(X=k)$
$= \mathbb{E}[X(X-1)\cdots(X-r+1)]$

Rq 46 = $\mathbb{E}[X] = G'_X(1^-)$ si la prop 45 s'applique.

Thm 47 = Processus de Galton Watson.
Soit X une v.a à valeurs dans $\mathbb{N}$. On note $p_k = \mathbb{P}(X=k)$, pour $k \in \mathbb{N}$ et $m = \mathbb{E}[X]$. Soit $(X_{i,j})_{i,j\in\mathbb{N}}$ une famille de v.a iid de loi $\mathbb{P}_X$, on construit Z_n par $\begin{cases} Z_0 = 1 \\ Z_{n+1} = \sum_{i=1}^{Z_n} X_{i,n} \end{cases}$ Alors

* G_X est bien définie sur $[0,1]$, de classe $\mathcal{C}^1$, strictement croissante sur $]0,1[$, convexe sur $]0,1[$ et même strictement convexe sur $]0,1[$ ssi $p_0 + p_1 < 1$.
* $\forall n \in \mathbb{N}^*$, $G_{Z_n} = \underbrace{G_X \circ \dots \circ G_X}_{n \text{ fois}}$ sur $[0,1]$
* La probabilité d'extinction $\pi_\infty = \mathbb{P}(\exists n \in \mathbb{N}\ Z_n = 0)$ est le plus petit point fixe de G_X sur $[0,1]$.
* Si $m \leq 1$ alors $\pi_\infty = 1$; sinon π_∞ est l'unique point fixe de G_X sur $]0,1[$.

DVLPT

BL61 2- Fonction caractéristique (et transformée de Laplace)

Gp212

Def 48 = On appelle fonction caractéristique de X la fonction
$\varphi_X : \mathbb{R} \to \mathbb{C}$
$t \mapsto \mathbb{E}[e^{itX}]$.

Rq 49 = • $\forall t \in \mathbb{R}$ $|\varphi_X(t)| \leq 1$.
• $\varphi_X(0) = 1$.
(• Si X est de densité f, φ_X est la transformée de Fourier de f.)

Thm 50 = Si X et Y sont deux v.a. tq $\varphi_X = \varphi_Y$ alors $\mathbb{P}_X = \mathbb{P}_Y$.

Ex 51 = • Si $X \sim \mathcal{N}(0,1)$, $\varphi_X : t \mapsto e^{-t^2/2}$ + $\mathcal{C}(0,1)$] DVLPT
• Si $X \sim \mathcal{P}(\lambda)$, $\varphi_X : t \mapsto \exp(\lambda(e^{it} - 1))$

BL p64

Prop 52 =
• Si $\mathbb{E}[|X|^n] < \infty$ alors φ_X est n-fois dérivable et $\varphi_X^{(n)}(t) = i^k \mathbb{E}[X^k e^{itX}]$.
• Si n est pair et si φ est n-fois dérivable en 0, alors X admet un moment d'ordre n.

O2 (Prop 53 = Si X admet des moments de tout ordre et si $\limsup_n \frac{\mathbb{E}[|X|^n]^{1/n}}{n} < \infty$, alors φ_X est analytique.)

Rq 54 = En général, une loi n'est pas caractérisée par ses moments.
contre ex :

Prop 55 = Si φ_X est analytique, alors $\mathbb{P}_X$ est caractérisée par $(\mathbb{E}[X^k])_{k\in\mathbb{N}}$

Def 56 = On appelle transformée de Laplace de X la fonction
$L_X : t \mapsto \mathbb{E}[e^{tX}]$, définie lorsque $e^{tX} \in L^1$.

Prop 57 = Si X et Y sont deux v.a. tq $L_X = L_Y$ sur un voisinage de 0, alors $\mathbb{P}_X = \mathbb{P}_Y$.

Thm 58 = Si $e^{tX} \in L^1$ pour t dans un intervalle ouvert contenant 0, alors L_X est analytique sur un voisinage de 0 et
$L_X(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \mathbb{E}[X^k]$ pour t dans ce voisinage.

3- Convergence. BL p131

Gp237

Thm 59 = Loi faible des grands nombres.
Soit (X_i) iid et L^1, alors $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{\mathbb{P}} \mathbb{E}[X_1]$

Thm 60 = Loi forte des grands nombres. (admise)
Soit (X_i) iid. $X_i \in L^1 \Leftrightarrow \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to+\infty]{p.s} \mathbb{E}[X_1]$

Rq 61 = Cela signifie que la moyenne empirique est un estimateur fortement consistant de l'espérance.

Gp277

Thm 62 = Théorème central limite.
Soit (X_i) iid et L^2, si $\sigma^2 = \mathrm{Var}(X) > 0$
Alors $\frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^{n} X_i - n\mathbb{E}[X_1]\right) \xrightarrow[n\to+\infty]{} Z$ où $Z \sim \mathcal{N}(0,1)$

App 63 = Intervalle de confiance asymptotique pour le paramètre d'une loi de Bernoulli.