

Résumé de cours :
Semaine 35, les 24 juin et 25 juin.

Première partie

Les probabilités (fin)

1 La variance

Définition. Soit $k \in \mathbb{N}^*$ et X une variable aléatoire réelle. Si X^k est d'espérance finie, on dit que $E(X^k)$ est le moment d'ordre k de X .

Notation. On note $L^2(\Omega, P)$ l'ensemble des variables aléatoires X discrètes à valeurs réelles possédant un moment d'ordre 2, définies sur l'espace probabilisé (Ω, \mathcal{F}, P) .

Lemme : Si $X_1, X_2 \in L^2(\Omega, P)$, alors $X_1 X_2 \in L^1(\Omega, P)$.

Corollaire. $L^2(\Omega, P)$ est un sous-espace vectoriel de $L^1(\Omega, P)$.

Définition. Si $X_1, X_2 \in L^2(\Omega, P)$, la covariance est $Cov(X_1, X_2) = E[(X_1 - E(X_1))(X_2 - E(X_2))]$.

Propriété. Cov est une forme bilinéaire symétrique positive sur $L^2(\Omega, P)$, mais ce n'est pas un produit scalaire.

Définition. Si $X \in L^2(\Omega, P)$, la variance de X est $Var(X) = E[(X - E(X))^2]$.
L'écart type de X est $\sigma(X) = \sqrt{Var(X)}$.

Remarque. $Var(X) = 0$ si et seulement si X est presque sûrement constante.

Définition. X est réduite si et seulement si $X \in L^2(\Omega, P)$ et $Var(X) = 1$.

Propriété. Formule de Koenig-Huygens : Si $X \in L^2(\Omega, P)$, $Var(X) = E(X^2) - E(X)^2$.
Si $X_1, X_2 \in L^2(\Omega, P)$, alors $Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$: donc, si deux variables aléatoires de $L^2(\Omega, P)$ sont indépendantes, elles sont orthogonales au sens de Cov (la réciproque est fautive).

Propriété. Pour $a, b \in \mathbb{R}$ et $X \in L^2(\Omega, P)$, $Var(aX + b) = a^2 Var(X)$.

Propriété. Si $X \in L^2(\Omega, P)$ avec $\sigma(X) \neq 0$, alors $\frac{X - E(X)}{\sigma(X)}$ est centrée et réduite.

Propriété.

◇ Si $X_1, X_2 \in L^2(\Omega, P)$, $Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1, X_2)$.

◇ Si $X_1, \dots, X_k \in L^2(\Omega, P)$, $Var(X_1 + \dots + X_k) = \sum_{i=1}^k Var(X_i) + 2 \sum_{1 \leq i < j \leq k} Cov(X_i, X_j)$.

◇ Si X_1, \dots, X_k sont k variables aléatoires de $L^2(\Omega, P)$ que l'on suppose **deux à deux indépendantes**, alors $Var(X_1 + \dots + X_k) = Var(X_1) + \dots + Var(X_k)$.

Il faut savoir le démontrer.

Propriété. Inégalité de Cauchy-Schwarz : pour tout $X, Y \in L^2(\Omega, P)$, $E(XY)^2 \leq E(X^2)E(Y^2)$, avec égalité ssi il existe $\alpha, \beta \in \mathbb{R}$ tel que $(\alpha, \beta) \neq (0, 0)$ et $\alpha X + \beta Y$ est presque sûrement nulle.
pour tout $X, Y \in L^2(\Omega, P)$, $Cov(X, Y)^2 \leq Var(X)Var(Y)$, avec égalité ssi il existe $\alpha, \beta \in \mathbb{R}$ tel que $(\alpha, \beta) \neq (0, 0)$ et $\alpha X + \beta Y$ est presque sûrement constante.

Définition. (hors programme) : Soient $X, Y \in L^2(\Omega, P)$ telles que $Var(X)Var(Y) > 0$. Le coefficient de corrélation linéaire entre X et Y est $Corr(X, Y) \triangleq \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}$.

Propriété. $Corr(X, Y) \in [-1, 1]$.

Propriété. $|Corr(X, Y)| = 1$ si et seulement si il existe $(a, b) \in \mathbb{R}^2$ tel que $P(Y = aX + b) = 1$.

Remarque. $Corr(X, Y)$ indique dans quelle mesure Y dépend **linéairement** de X , mais $Corr(X, Y)$ ne mesure pas les dépendances non linéaires (on peut avoir par exemple $Corr(X, X^2) = 0$).

Formule. Espérance et variance pour les lois au programme.

◇ **Loi de Bernoulli** de paramètre $p \in [0, 1]$: $P(X = 1) = p$ et $P(X = 0) = 1 - p$.

$$\boxed{E(X) = p \text{ et } Var(X) = p(1 - p)} .$$

◇ **Loi binomiale** de paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$: Pour tout $k \in \{0, \dots, n\}$, $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ (et $P(X = m) = 0$ pour $m \notin \{0, \dots, n\}$).

$$\boxed{E(X) = np \text{ et } Var(X) = np(1 - p)} .$$

◇ **Loi géométrique** de paramètre $p \in]0, 1[$: Pour tout $n \in \mathbb{N}^*$, $P(X = n) = (1 - p)^{n-1} p$ (et $P(X = 0) = 0$).

$$\boxed{E(X) = \frac{1}{p} \text{ et } Var(X) = \frac{1 - p}{p^2}} .$$

◇ **Loi de Poisson** de paramètre $\lambda \in \mathbb{R}_+^*$:

pour tout $n \in \mathbb{N}$, $P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$. $E(X) = \lambda = Var(X)$.

Il faut savoir le démontrer.

2 Propriétés de convergence

Formule. Inégalité de Bienaymé-Tchebychev : Soit X une variable aléatoire réelle. Alors, pour

$$\text{tout } \varepsilon > 0, \quad \boxed{P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}} .$$

Il faut savoir le démontrer.

Définition. (hors programme) Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires et soit X une variable aléatoire. X_n converge vers X en probabilité ssi pour tout $\varepsilon > 0$, $P(|X_n - X| \geq \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.

Théorème. Loi faible des grands nombres :

Soit (X_n) une suite de variables aléatoires dans $L^2(\Omega, P)$ que l'on suppose toutes de même loi et deux à deux indépendantes. Posons $\mu = E(X_n)$, qui est indépendante de n . Alors $\frac{X_1 + \dots + X_n}{n}$ converge en probabilité vers la variable aléatoire constante égale à μ .

Il faut savoir le démontrer.

Deuxième partie

Théorie de l'intégration

Notation. $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $a, b \in \mathbb{R}$ avec $a < b$, E est un Banach, i.e un \mathbb{K} -espace vectoriel normé complet, f est une application de $[a, b]$ dans E .

3 Intégration des applications en escalier

3.1 Les applications en escalier

Définition. On appelle subdivision de $[a, b]$ toute famille finie $(a_i)_{0 \leq i \leq n}$ de réels telle que $a = a_0 < a_1 < \dots < a_n = b$.

Notation. On notera \mathcal{S} l'ensemble des subdivisions de $[a, b]$.

Exemple. $\left(a + i \frac{b-a}{n}\right)_{0 \leq i \leq n} \in \mathcal{S}$. On dit que c'est une subdivision uniforme.

Définition. Le pas de $\sigma = (a_i)_{0 \leq i \leq n} \in \mathcal{S}$ est $\delta(\sigma) = \max_{1 \leq i \leq n} (a_i - a_{i-1})$.

Notation. Le support de la subdivision σ est l'ensemble $A(\sigma) \triangleq \{a_i / 0 \leq i \leq n\}$.

Propriété. Notons $\mathcal{P}_f([a, b])$ l'ensemble des parties finies de $[a, b]$ contenant a et b .

$$\begin{array}{lcl} \text{L'application } A : \mathcal{S} & \longrightarrow & \mathcal{P}_f([a, b]) \\ \sigma & \longmapsto & A(\sigma) \end{array} \quad \text{est bijective.}$$

Définition. $\sigma \in \mathcal{S}$ est plus fine que $\sigma' \in \mathcal{S}$ ssi $A(\sigma) \supseteq A(\sigma')$. Dans ce cas, on note $\sigma' \preceq \sigma$.

Propriété. \preceq est une relation d'ordre partiel.

Définition. Si $\sigma, \sigma' \in \mathcal{S}$, on pose $\sigma \cup \sigma' \triangleq A^{-1}(A(\sigma) \cup A(\sigma'))$: c'est l'unique subdivision de $[a, b]$ dont le support est la réunion des supports de σ et de σ' . C'est $\sup\{\sigma, \sigma'\}$.

Définition. f est une application en escalier sur $[a, b]$ si et seulement s'il existe une subdivision $(a_i)_{0 \leq i \leq n}$ de $[a, b]$ telle que, pour tout $i \in \mathbb{N}_n$, f est constante sur l'intervalle $]a_{i-1}, a_i[$.

Définition. Si f est en escalier et $\sigma = (a_i)_{0 \leq i \leq n} \in \mathcal{S}$, σ est une subdivision adaptée à f si et seulement si, pour tout $i \in \mathbb{N}_n$, f est constante sur l'intervalle $]a_{i-1}, a_i[$.

Propriété. Les applications en escalier de $[a, b]$ sont bornées.

Propriété. Soit f une application en escalier et σ une subdivision de $[a, b]$ adaptée à f . Alors toute subdivision plus fine que σ est aussi adaptée à f .

3.2 Intégrale d'une application en escalier

Définition. Soit f une application en escalier et $\sigma = (a_i)_{0 \leq i \leq n}$ une subdivision adaptée à f . Pour tout $i \in \mathbb{N}_n$, notons λ_i la valeur constante de f sur $]a_{i-1}, a_i[$. On pose

$$\int_a^b f(t) dt = \sum_{i=1}^n (a_i - a_{i-1}) \lambda_i.$$

Cette quantité est indépendante du choix de σ parmi les subdivisions adaptées à f .

Il faut savoir le démontrer.

Remarque. Lorsque $E = \mathbb{R}$, $\int_a^b f$ représente une somme d'aires de rectangles, affectées d'un signe négatif lorsque $\lambda_i < 0$, donc $\int_a^b f$ est l'aire algébrique de la surface située entre le graphe de f et l'axe des abscisses.

Propriété. Supposons que f est en escalier et soit g une application de $[a, b]$ dans E qui ne diffère de f qu'en un nombre fini de points de $[a, b]$. Alors g en escalier et $\int_a^b g = \int_a^b f$.

Théorème. Notons $\mathcal{E}([a, b], E)$ l'ensemble des applications en escalier de $[a, b]$ dans E . C'est un \mathbb{K} -espace vectoriel et l'application $\mathcal{E}([a, b], E) \rightarrow E$
 $f \mapsto \int_a^b f$ est linéaire.

Il faut savoir le démontrer.

Propriété. Soient F un second \mathbb{K} -espace vectoriel de dimension finie et $u \in L(E, F)$.

Si f est en escalier, $u \circ f$ est en escalier et $\int_a^b u \circ f = u \left(\int_a^b f \right)$.

Propriété. Si f est une application en escalier à valeurs dans \mathbb{R}_+ , $\int_a^b f \geq 0$.

Corollaire. Si $f, g \in \mathcal{E}([a, b], E)$, alors $[\forall t \in [a, b], f(t) \leq g(t)] \implies \int_a^b f \leq \int_a^b g$.

Inégalité triangulaire : Pour tout $f \in \mathcal{E}([a, b], E)$, $\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt$.

Relation de Chasles : Soit $f \in \mathcal{E}([a, b], E)$ et $c \in]a, b[$.

Alors $f|_{[a, c]}$ et $f|_{[c, b]}$ sont des applications en escalier et $\int_a^b f = \int_a^c f + \int_c^b f$.

4 Les applications réglées

4.1 Définition

Définition. On dit que $f : [a, b] \rightarrow E$ est réglée si et seulement si c'est la limite uniforme d'une suite d'applications en escalier, c'est-à-dire si et seulement si il existe une suite $(f_n) \in \mathcal{E}([a, b], E)^{\mathbb{N}}$ telle que $\sup_{x \in [a, b]} \|f_n(t) - f(t)\| \xrightarrow{n \rightarrow +\infty} 0$. On note $\mathcal{R}([a, b], E)$ l'ensemble des applications réglées.

Propriété. $\mathcal{R}([a, b], E)$ est l'adhérence de $\mathcal{E}([a, b], E)$ dans $(\mathcal{B}([a, b], E), \|\cdot\|_{\infty})$.

4.2 Les applications continues par morceaux

Propriété. $C([a, b], E) \subset \mathcal{R}([a, b], E)$: toute application continue est réglée.

Il faut savoir le démontrer.

Définition. $f : [a, b] \rightarrow E$ est continue par morceaux si et seulement si il existe une subdivision $\sigma = (a_i)_{0 \leq i \leq n}$ de $[a, b]$ telle que, pour tout $i \in \mathbb{N}_n$, $f|_{]a_{i-1}, a_i[}$ est prolongeable par continuité sur $[a_{i-1}, a_i]$, ce qui est équivalent à f est continue sur $[a, b] \setminus \{a_0, \dots, a_n\}$ et f admet en chaque a_i une limite à droite (sauf en b) et une limite à gauche (sauf en a). Dans ce cas, on dit que la subdivision σ est adaptée à f .

Définition. Si I est un intervalle quelconque de \mathbb{R} , $f : I \rightarrow E$ est continue par morceaux si et seulement si toutes ses restrictions aux segments inclus dans I sont continues par morceaux.

Propriété. Les applications continues par morceaux de $[a, b]$ dans E sont réglées.

Théorème. (Hors programme) Une application de $[a, b]$ dans E est réglée si et seulement si elle admet en tout point de $[a, b]$ une limite à droite (sauf en b) et une limite à gauche (sauf en a).

Corollaire. Les applications monotones de $[a, b]$ dans \mathbb{R} sont réglées.

Corollaire. Le produit de deux applications réglées est réglé.

5 Intégration des applications réglées

5.1 Construction

Définition. Soit $f : [a, b] \rightarrow E$ une application réglée.

Il existe $(f_n)_{n \in \mathbb{N}} \in \mathcal{E}([a, b], E)^{\mathbb{N}}$ telle que $f_n \xrightarrow[n \rightarrow +\infty]{\|\cdot\|_{\infty}} f$. On pose $\int_a^b f(t) dt = \lim_{n \rightarrow +\infty} \left(\int_a^b f_n(t) dt \right)$.

Il faut savoir le démontrer.

Remarque. Seule la construction de l'intégrale sur $[a, b]$ d'une application continue par morceaux est au programme.

5.2 Propriétés

Théorème. $\mathcal{R}([a, b], E)$ est un \mathbb{K} -espace vectoriel et $\begin{array}{ccc} \mathcal{R}([a, b], E) & \longrightarrow & E \\ f & \longmapsto & \int_a^b f \end{array}$ est linéaire.

Il faut savoir le démontrer.

Propriété. Soit F un second \mathbb{K} -espace vectoriel de Banach et $u \in L(E, F)$ que l'on suppose continue.

Si $f \in \mathcal{R}([a, b], E)$, alors $u \circ f \in \mathcal{R}([a, b], F)$ et $\int_a^b u \circ f = u \left(\int_a^b f \right)$.

Il faut savoir le démontrer.

Propriété. On suppose que E est de dimension finie et que $e = (e_1, \dots, e_p)$ est une base de E . Soit $f \in \mathcal{R}([a, b], E)$. Notons f_1, \dots, f_p les applications coordonnées de f , de sorte que, pour tout $t \in [a, b]$, $f(t) = \sum_{j=1}^n f_j(t)e_j$. Alors f_1, \dots, f_p sont réglées et $\int_a^b f(t) dt = \sum_{j=1}^n \left(\int_a^b f_j(t) dt \right) e_j$.

Remarque. Réciproquement, si f_1, \dots, f_p sont réglées, alors f est aussi réglée.

Propriété. Supposons que $E = \prod_{i=1}^p E_i$, où pour tout $i \in \mathbb{N}_p$, E_i est un espace de Banach. Soit $f \in \mathcal{R}([a, b], E)$. Notons f_1, \dots, f_p les applications composantes de f , de sorte que, pour tout $t \in [a, b]$, $f(t) = (f_1(t), \dots, f_p(t))$. Alors f_1, \dots, f_p sont réglées et $\int_a^b f = \left(\int_a^b f_i \right)_{1 \leq i \leq p}$.

Remarque. Réciproquement, si f_1, \dots, f_p sont réglées, alors f est aussi réglée.

Inégalité triangulaire : Pour tout $f \in \mathcal{R}([a, b], E)$, $\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt$.

Propriété. Si f est une application réglée à valeurs dans \mathbb{R}_+ , $\int_a^b f \geq 0$.

Corollaire. Si $f, g \in \mathcal{R}([a, b], E)$, alors $[\forall t \in [a, b], f(t) \leq g(t)] \implies \int_a^b f \leq \int_a^b g$: l'intégrale est croissante.

Exemple. Si f est réglée, $\left\| \int_a^b f(t) dt \right\| \leq (b-a) \sup_{t \in [a,b]} \|f(t)\|$.

Propriété. Soit f une application réglée (resp : continue par morceaux) de $[a, b]$ dans E . Si g est une application de $[a, b]$ dans E qui ne diffère de f qu'en un nombre fini de points de $[a, b]$, alors g est réglée (resp : continue par morceaux) et $\int_a^b f = \int_a^b g$.

Relation de Chasles : soit $f \in \mathcal{R}([a, b], E)$ et $c \in]a, b[$.

Alors $f|_{[a,c]}$ et $f|_{[c,b]}$ sont réglées et $\int_a^b f = \int_a^c f + \int_c^b f$.

Convention : Si f est une application définie en $\alpha \in \mathbb{R}$, on convient $\int_\alpha^\alpha f = 0$.

Convention : Si $f : [a, b] \rightarrow E$ est réglée, on convient que $\int_b^a f = - \int_a^b f$.

Propriété. La relation de Chasles se généralise au cas d'une application f réglée sur l'intervalle $[\min(a, b, c), \max(a, b, c)]$, les réels (a, b, c) étant quelconques.

Remarque. Avec ces conventions, les égalités établies dans ce paragraphe restent valables, mais ce n'est pas le cas des inégalités.

6 Sommes de Riemann

Notation. On fixe une application f de $[a, b]$ dans E .

Définition. On appelle subdivision pointée de $[a, b]$ tout couple (σ, ξ) , où $\sigma = (a_i)_{0 \leq i \leq n}$ est une subdivision de $[a, b]$ et où $\xi = (\xi_i)_{1 \leq i \leq n}$ vérifie $\forall i \in \mathbb{N}_n \quad \xi_i \in [a_{i-1}, a_i]$.

Notation. Notons \mathcal{S}' l'ensemble des subdivisions pointées de $[a, b]$. Si $(\sigma, \xi) = ((a_i), (\xi_i)) \in \mathcal{S}'$, on notera $f_{\sigma, \xi}$ l'application en escalier définie par $\forall i \in \mathbb{N}_n \quad \forall x \in]a_{i-1}, a_i[\quad f(x) = f(\xi_i)$,

Définition. Soit $(\sigma, \xi) = ((a_i)_{0 \leq i \leq n}, (\xi_i)_{1 \leq i \leq n}) \in \mathcal{S}'$. On appelle somme de Riemann associée à f et à (σ, ξ) la quantité $S(f, \sigma, \xi) = \int_a^b f_{\sigma, \xi} = \sum_{i=1}^n (a_i - a_{i-1}) f(\xi_i)$.

Théorème. Si f est une application réglée de $[a, b]$ dans E ,

$$\forall \varepsilon \in \mathbb{R}_+^* \quad \exists \alpha \in \mathbb{R}_+^* \quad \forall (\sigma, \xi) \in \mathcal{S}' \quad (\delta(\sigma) \leq \alpha \implies \|S(f, \sigma, \xi) - \int_a^b f\| \leq \varepsilon).$$

À savoir démontrer lorsque f est continue.

Corollaire. Soit $(\sigma_n, \xi_n)_{n \in \mathbb{N}} \in \mathcal{S}'^{\mathbb{N}}$ une suite de subdivisions pointées dont le pas tend vers 0. Alors, si f est réglée, la suite des sommes de Riemann associée à f et à (σ_n, ξ_n) converge vers $\int_a^b f$. Plus

précisément, en notant, pour tout $n \in \mathbb{N}$, $\sigma_n = (a_{i,n})_{0 \leq i \leq \varphi(n)}$ et $\xi_n = (\xi_{i,n})_{1 \leq i \leq \varphi(n)}$, si f est réglée et si $\max_{1 \leq i \leq \varphi(n)} (a_{i,n} - a_{i-1,n}) \xrightarrow{n \rightarrow +\infty} 0$,

$$\text{alors } \sum_{i=1}^{\varphi(n)} (a_{i,n} - a_{i-1,n}) f(\xi_{i,n}) \xrightarrow{n \rightarrow +\infty} \int_a^b f.$$

Cas particulier : si f est continue par morceaux, $\frac{b-a}{n} \sum_{i=1}^n f\left(a + i \frac{b-a}{n}\right) \xrightarrow{n \rightarrow +\infty} \int_a^b f$.

7 Primitives

Notation. Conformément au programme officiel, on se limite au cas où E est un \mathbb{K} -espace vectoriel de dimension finie. On sait alors qu'il est complet, donc c'est bien un espace de Banach. On fixe un intervalle I de \mathbb{R} d'intérieur non vide et une application $f : I \rightarrow E$.

Définition. $g : I \rightarrow E$ est une primitive de f si et seulement si g est dérivable sur I et $g' = f$.

Propriété. Si f admet une primitive g_0 sur I , alors g est une primitive de f si et seulement si il existe $k \in E$ tel que $\forall x \in I \quad g(x) = g_0(x) + k$.

Propriété. On suppose que f est réglée sur I (c'est-à-dire que les restrictions de f aux intervalles compacts inclus dans I sont réglées). Soit $a \in I$. Alors $x \mapsto \int_a^x f(t) dt$ est continue sur I .

Théorème fondamental de l'analyse : On suppose que f est continue sur I . Soit $a \in I$.

Alors
$$\boxed{\begin{array}{l} F : I \rightarrow E \\ x \mapsto \int_a^x f(t) dt \end{array}} \text{ est l'unique primitive de } f \text{ s'annulant en } a.$$

Il faut savoir le démontrer.

Corollaire. Soient $(a, b) \in \mathbb{R}^2$ avec $a \neq b$ et f une application continue de $[a, b]$ dans E . Si F est une primitive de f , alors $\int_a^b f(t) dt = F(b) - F(a) \stackrel{\text{notation}}{=} [F(t)]_a^b$.

Corollaire. Si f est une application de classe C^1 sur $[a, b]$, $\int_a^b f'(t) dt = f(b) - f(a)$.

Théorème. Soit f une application de $[a, b]$ dans \mathbb{R} .

Si f est continue, positive et si $\int_a^b f = 0$, alors f est identiquement nulle sur $[a, b]$.

Il faut savoir le démontrer.

Définition. Si $f : [a, b] \rightarrow E$ est réglée, la valeur moyenne de f est $\frac{1}{b-a} \int_a^b f(t) dt$.

Propriété. Si $f : [a, b] \rightarrow \mathbb{R}$ est une application continue, f atteint sa valeur moyenne : il existe $c \in]a, b[$ tel que $f(c) = \frac{1}{b-a} \int_a^b f(t) dt$.