Loi de proba

Définition : On appelle probabilité P sur un espace mesurable $(\Omega, \mathcal{F})$ toute application de $\mathcal{F}$ dans $[0, 1]$ vérifiant

1/ $P(A) \geq 0$, $\forall A \in \mathcal{F}$

2/ $P(\Omega) = 1$

3/ $\forall A, B \in \mathcal{F}$, $A \cap B = \phi$ alors $P(A \cap B) = P(A) P(B)$

4/ Soit $\{A_n\}$ une suite d'évènement décroissante (au sens inclusion) vers ϕ, $A_n \in \mathcal{F}$, $\forall n \in \mathbb{N}$, $A_n \searrow \phi$

$A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ et $\bigcap_{n=1}^{\infty} A_n = \phi$ alors $\lim_{n \to \infty} P(A_n) = 0$

(continuité en 0 de P).

- Le Triplet $(\Omega, \mathcal{F}, P)$ on appelle espace probabilisé.

Définition : On dit qu'une propriété est vérifiée presque-sûrement et on note PP sûre si elle est vraie sur un évènement $C \in \mathcal{F}$, telle que $P(C) = 1$

Théorème : La fct $P(\cdot)$ est σ-additive sur $\mathcal{F}$ ie : pour chaque suite d'évènements $\{A_n\}$ $(\forall n \in \mathbb{N}^*)$, $(A_n \in \mathcal{F})$ disjoints deux à deux on a :

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Variable aléatoire :

Définition : une application $X : (\mathcal{F}, P) \longrightarrow (E, \mathcal{E})$ s'appelle une variable aléatoire sur $(\Omega, \mathcal{F}, P)$ si X est $(\mathcal{F}, \mathcal{E})$ mesurable $(\forall B \in \mathcal{E}, X^{-1}(B) \in \mathcal{F})$

Remarque : Si $E = \mathbb{R}$ (Resp $E = \mathbb{R}^d$ ou $d \in \mathbb{N}^*$) on dit que X est une variable aléatoire (Resp de dimension d)

Proposition : une application $X : \Omega \longrightarrow \mathbb{R}^d$ est v.a si et seulement si ses composantes sont de v.a réelle.

Définition : La σ-algèbre minimale $B(\mathbb{R})$ de $\mathbb{R}$ engendrée par l'intervalle comme $]a, b[$ est appelée la σ-algèbre borélienne si $B \in B(\mathbb{R})$ on dit que B est un ensemble borélien et par conséquent on dit que $(\mathbb{R}, B(\mathbb{R}))$ est l'espace Borélien.

- Soit $B \in B(\mathbb{R})$, note $X^{-1}(B) = \{w ; X(w) \in B\}$. Notons que $\mathcal{A}_X = \{A : A = X^{-1}(B), B \in B(\mathbb{R})\}$ est une σ-algèbre

Définition : On dit que A_X est la σ-algèbre engendrée par la v.a X, Il est évident que $A_X \subseteq \mathcal{F}$.

La Loi de Probabilité (d'un v.a) :

Définition :

Soit $P_X(B) = P(\{w : X(w) \in B\}) = P(X \in B)$

$= P(X^{-1}(B)) = P \circ X^{-1}$.

avec $B \in \mathcal{B}(\mathbb{R})$. On dit que $P_X(.)$ est la loi de X par rapport à P.

$Q : (E, \mathcal{E}) \longrightarrow [0,1]$

Proposition :

Soit $(E, \mathcal{E})$ un espace mesuré et Q une probabilité $(E, \mathcal{E})$. Alors il existe un espace de probabilité $(\Omega, \mathcal{F}, P)$ et une v.a sur $(\Omega, \mathcal{F}, P)$ à valeurs dans E et de loi de probabilité $P_X = Q$

Proposition :

Soit $(E, \mathcal{E})$ un espace mesuré et Q une probabilité sur $(E, \mathcal{E})$. une variable aléatoire X à valeurs E sur un espace de probabilité $(\Omega, \mathcal{F}, P)$ admet Q pour loi de proba. si et seulement si pour tout application $\mathcal{E}$-mesurable et positive (ou intégrable)

Q_X loi de $X \Leftrightarrow h : E \longrightarrow [0,+\infty[$

$h : E \longrightarrow [0, +\infty[$

$$\int_{\Omega} h(X(\omega)) \, dP(x) = \int_E h(x) \, dQ(x)$$

Définition :

Si Q une proba sur $\mathbb{R}$, On appelle fonction de répartition de Q et on note F_Q l'application de $\mathbb{R}$ dans $[0,1]$ définit pour tout $x \in \mathbb{R}$ par $F_Q(x) = Q(]-\infty, x])$

si X est un variable aléatoire réel. La fct de répartition de la loi X est appellé la fct de répartition de la variable X. On la noté F_X ($Q = P_X$)

• La fonction de répartition de Q d'une probabilité a les propriétés suivantes :

Prop :

La fct de répartition de Q est croissante continuité à droite sur $\mathbb{R}$, $\lim_{x \to +\infty} F_Q(x) = 1$ et $\lim_{x \to -\infty} F_Q(x) = 0$ / $\forall x \in \mathbb{R}$.

Variables aléatoires discrètes

Définition :

Une mesure Q sur $\mathbb{R}$ est dite discrètes si elle peut s'écrire :

$$Q = \sum_{n \in \mathbb{N}} P_n \delta_{e_n}.$$

$(P_n)_{n \in \mathbb{N}}$: est une suite réelle positive ou nulle.

(e_n) : est une suite d'éléments de $\mathbb{R}$.

δ_a : est la masse de dirac au points a $\delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{sinon.} \end{cases}$

• Une var. a. r X est dite discrète si sa loi est discrète. Dans ce cas : $\forall n \in \mathbb{N}$, $P_n = P(X = e_n)$

Proposition :

Une v.a est discrète si et seulement si il existe partie D de $\mathbb{R}$ dénombrable telle que : $P(X \in D) = 1$

La loi de X est $P_X = \sum_{x \in D} P(X = x)\, \delta_x$.

$X \sim \mathcal{B}in(n,p)$

$P(X=k) = C_n^k p^k q^{n-k}$

$P_X(X=k) = \sum_{k=0}^{n} P(X=k)\delta_k = \sum_{k=0}^{n} C_n^k p^k q^{n-k} \delta_k$

Poisson

$Y \sim \mathcal{P}(\lambda)$

$P(X=n) = e^{-\lambda} \frac{\lambda^n}{n!}$

$P(Y=n) = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$, $P_Y = \sum_{n \geq 0} P(Y=n)\delta_n = \sum_{n \geq 0} e^{-\lambda} \frac{\lambda^n}{n!} \delta_n$

. Variables aléatoires absolument continues :

Définition :

On dit qu'une mesure Q sur $\mathbb{R}$ est absolument continue si elle admet une densité f par rapport à la mesure de Lebesgue $\lambda \in \mathbb{R}$ ie : $f : \mathbb{R} \longrightarrow [0, +\infty[$ intégrable au sens de Lebesgue. et tq : $\forall A \in \mathcal{B}(\mathbb{R})$, $Q(A) = \int_A f \, d\lambda$. On note $Q = f\lambda$ ou $dQ = f d\lambda$

On dit qu'une v.a.r est absolument continue sur $\mathbb{R}$. si sa loi est absolument continue sur $\mathbb{R}$.

. Moments d'une variable aléatoire :

Comme la mesure de proba P est une mesure finie sur $(\Omega, \mathcal{F})$ on a $\forall p, q \in [1, +\infty[$ avec $p < q$

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

Définition :

Si X est une v.a.r sur $(\Omega, \mathcal{F}, P)$ positive (intégrable) sur Ω. On appelle espérance mathématique de X, et on note $E(X)$ la quantité : $E(X) = \int_{\Omega} X(\omega)\, dP(\omega)$

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

$$m_p = \int_{\Omega} X^p(\omega)\, dP(\omega) \qquad m_p = \int_{\Omega} (X(\omega) - m_1)^p\, dP(\omega)$$

. Le moments centré d'ordre 2 s'appelle la variance de X

. Sa racine carré positive s'appelle l'écart-type de X et sa note σ_X.

. Si X et Y sont deux v.a.r de carré intégrable on appelle covariance de X et Y, et on note :

$$\mathrm{cov}(X, Y) = E(XY) - E(X)E(Y)$$

Loi d'un Vecteur aléatoire :

Définition :

Une application $X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

$\omega \longmapsto X(\omega)$

tq $X(\omega) = (X_1(\omega), \dots, X_n(\omega))^T = (x_1, \dots, x_n)^T$

st appelé un vecteur aléatoire si $X^{-1}(B) \in \mathcal{F}$, $\forall B \in \mathcal{B}(\mathbb{R}^n)$
où $\mathcal{B}(\mathbb{R}^n)$ st la σ. algèbre borélienne de $\mathbb{R}^n$.

$B \in \mathcal{B}(\mathbb{R}^n) \Leftrightarrow B = B_1 \times B_2 \times \dots \times B_n$ avec $B_i \in \mathcal{B}(\mathbb{R}), i = \overline{1,n}$

Définition :

Soit P_X st une mesure sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ déterminer par :

$P_X(B) = P(w \in \Omega, X(w) \in B) = P(X \in B) = P(X^{-1}(B))$, $B \in \mathcal{B}(\mathbb{R}^n)$

P_X s'appelle la loi (distribution de X dans $\mathbb{R}^n$,

Loi d'un vecteur aléatoire

Remarque :

On définit la fonction de répartition $F(x) = F(x_1, \dots, x_n)$ $= P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$
d'un vecteur $X = (x_1, x_2, \dots, x_n)^T$

Définition :

Soit $y = g(x)$, une application de $\mathbb{R}^n$ dans $\mathbb{R}$ $g : \mathbb{R}^n \to \mathbb{R}$
nous disons que g st une fct Borélienne, pour tout ensemble borélien $B \in \mathcal{B}(\mathbb{R}^n)$

. $g^{-1}(B)$ st un ensemble dans $\mathcal{B}(\mathbb{R}^n)$ ie :

$\{ x : g(x) \in B \} = g^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$. ($g$ st $\mathcal{B}(\mathbb{R}^n)$ st mesu

Théorème :

Soit X un vecteur aléatoire sur $(\Omega, \mathcal{F}, P)$:

$X : \Omega \to \mathbb{R}^n$ et $g : \mathbb{R}^n \to \mathbb{R}$ une application borélienne dans ce cas la statistique $g(X) = g(X(w))$

est une v.a sur $(\Omega, \mathcal{F}, P)$

$$\begin{tikzcd}(\Omega, \mathcal{F}, P) \arrow[r, "X"] \arrow[rr, bend right, "g \circ X"'] & (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \arrow[r, "g"] & (\mathbb{R}, \mathcal{B}(\mathbb{R}))\end{tikzcd}$$

Loi de proba discrète

Soit $X = (X_1, X_2, \dots, X_n)$ un vecteur aléatoire sur $(\Omega, \mathcal{F}, P)$

Définition :

L'ensemble $\Lambda = X(\omega)$ qui représente l'image de ω ds $\mathbb{R}^n$, s'appelle l'espace des réalisations de X dans $\mathbb{R}^n$.

Donc $\Omega = X^{-1}(\Lambda)$ est l'image réciproque complète de Λ

Soit $B \in \mathcal{B}(\mathbb{R}^n)$, dans ce cas $P_X(B) = P\{\omega : X(\omega) \in B\}$
$= P(X^{-1}(B))$

qui l'on appelle loi de X dans $\mathbb{R}^n$ comme :

$P(X \in \Lambda) = 1$, on peut dire aussi que :

$P_X(\Lambda^c) = P_X(\mathbb{R}^n - \Lambda) = 0$

Si X est telque son espace Λ est fini ou infinie dénombrable ie : $\Lambda = \{x_1, x_2, \dots, x_n, \dots\}$; $x_i \in \mathbb{R}^n$

Alors, on est dans ce cas discret et la mesure P_X est déterminée par les mesures des points

$x_i \in \Omega$, $P_X = \sum_{x \in \omega} P(X=x)\,\delta_x$

Loi absolument continues

Définition :
On dit que un vecteur X a une distribution absolument continue par rapport à la mesure de Lebesgue λ^n dans $\mathbb{R}^n$.
Si sa distribution P_X est donnée par :

$$P_X(B) = \int_B f_X(x)\, d\lambda^n(x) = \int_B dP_X(x)$$

$\forall B \in \mathcal{B}(\mathbb{R}^n)$ où $f_X(x)$ est une fct $\mathcal{B}(\mathbb{R}^n)$ mes la fct $f_X(\cdot)$ s'appelle la densité de X par rapport à la mesure de Lebesgue.
On peut montrer que $f_X(x) \geq 0$ et $\int_{\mathbb{R}^n} f_X(x)\, dx = 1$.
Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire dont la densité est $f_X(x)$, $x = (x_1, x_2, \dots, x_n)^T$
Considérons une statistique $Y = f(x)$ où $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ est une application dérivable.
Notons : $y = f(x)$ i.e : $y = (y_1, \dots, y_n)^T$ où $y_i = f_i(x)$
$x \in \mathbb{R}$, $f_j(x_1, \dots, x_n)$

$X : \omega \longrightarrow \mathbb{R}^n \longrightarrow \mathbb{R}^n$
$\omega \longrightarrow X(\omega) = (X_1(\omega), \dots, X_n(\omega)) \longrightarrow f(\cdot)$
$(f(x_1, \dots, x_n)$
$(f_1(x) \dots f_n(x)$

Jaccobien d'une fonction :

. Le jaccobien de f est une application tq :

$Df = \mathbb{R}^n \longrightarrow \mathbb{R}$

donnée par la formule $Df(x) = \det \left\| \dfrac{\partial f_i(x)}{\partial x_j} \right\|$

i.e. $Df(x)$ est le déterminant de la matrice Jacobien.

- Si $Df(x) \neq 0$ au voisinage d'un points $x \in \mathbb{R}^n$ dans ce cas $f^{-1}(y)$ est au voisinage du points $y = f(x)$ avec $Df^{-1}(f(x))\, Df(x) = 1$

$$\text{Jaccobien de } f(x) = \det \begin{vmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \\ \vdots & & & \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{vmatrix}$$

Alors f^{-1} existe alors pour tout φ $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ intégrable

$$\int_A \varphi(y)\,dy = \int_{f^{-1}(A)} \varphi(f(x))\,|\Delta f(x)|\,dx. \quad \text{où } A \in \mathcal{B}(\mathbb{R}^n)$$

Lemme :

Soit $y = g(x)$ et $f_X(x)$ la densité de $X : \Omega \longrightarrow \mathbb{R}^n$

On g est tq g^{-1} existe Dans ce cas la densité $f_Y(y)$ de Y est donnée par : $f_Y(y) = f_X(g^{-1}(y)) \, |Dg^{-1}(y)|$

$y_1 = x_1 + x_2$

$y_2 = x_1 - x_2$ $\qquad g^{-1}(y) = g_1^{-1}(y)\, g_2^{-1}(y)$

$x_1 = ?$

$x_2 = ?$

$X : (X_1, X_2)$

$g(X) = g(X_1, X_2)$

$= g_1(x_1), g_2(x_2))$

$= (X_1 + X_2, X_1 - X_2)$

Théorème :

Soit $X = (X_1, X_2, \dots, X_n)^T : \omega \longrightarrow \mathbb{R}^n$ Dans ce cas la densité de X_1 est : $f_{X_1}(x_1) = \int_{\mathbb{R}^{n-1}} f_X(x_1, x_2, \dots, x_n)\, dx_2, \dots, dx_n$.

Lemme :

Soient $Y = g(x)$ et $g_X(x)$ la densité de $X \in \mathbb{R}^n$ où g est tq g^{-1} existe Dans ce cas la densité $f_Y(y)$ de la v.a Y est donnée par :

$$f_Y(y) = f_X(g^{-1}(y)) \, |Dg^{-1}(y)| \quad (*)$$

Exemple :

Soit $X = (X_1, X_2)^T$, $Y_1 = X_1 + X_2$

On considère, $Y = (Y_1, Y_2)^T = g(X)$, $Y_1 = X_1 + X_2$, $Y_1 = g_1(X)$

et $Y_2 = x_2 = g_2(x)$

c.à.d. $g(x) = (y_1, y_2) = (g_1(x), g_2(x))^T$

avec $g_1(x) = x_1 + x_2$, $g_2(x) = x_2$

Dans ce cas : $\dfrac{\partial g_1}{\partial x_1}(x) = \dfrac{\partial g_1(x)}{\partial x_2} = 1$

$\dfrac{\partial g_1(x)}{\partial x_1} = 0$, $\dfrac{\partial g_2(x)}{\partial x_2} = 1$

$\Rightarrow \det \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 = Dg(x)$

Notons que $x = (x_1, x_2)^T = g^{-1}(y)$ est donnée par les formules : $x_1 = g_1^{-1}(y) = y_1 - y_2$, $x_2 = g_2^{-1}(y) = y_2$

$\dfrac{\partial g_1^{-1}(y)}{\partial y_1} = 1$, $\dfrac{\partial g_1^{-1}(y)}{\partial y_2} = -1$

$\dfrac{\partial g_2^{-1}(y)}{\partial y_1} = 0$, $\dfrac{\partial g_2^{-1}(y)}{\partial y_2} = 1$

$Dg^{-1}(y) = \det \left| \dfrac{\partial g_i^{-1}(y)}{\partial y_j} \right| = 1$

Par (*)

$f_Y(y) = f_X(g^{-1}(y)) \, |Dg^{-1}(y)|$ et par conséquant

- $f_{Y_1}(y_1) = \int_{\mathbb{R}} f_Y(y) \, dy_2 = \int_{\mathbb{R}} f_X(y_1 - y_2, y_2) \, dy_2$
- $f_{Y_2}(y_2) = \int_{\mathbb{R}} f_Y(y) \, dy_1 = \int_{\mathbb{R}} f_X(y_1 - y_2, y_2) \, dy_1$.

Théorème :

Si la densité $f_X(x)$ du vecteur $X \in \mathbb{R}^n$ est présente par la formule $f_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i)$, où $f_{X_i}(x_i)$ est la densité de X_i. Dans ce cas les variables $X_1, X_2, \dots, X_n$ sont indépendantes.

Remarque :

Soit $X = (X_1, X_2)^T$ un vecteur aléatoire, dont les composantes sont indépendantes, Dans ce cas $f_X(x) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$ et donc la densité de la stat $Y_1 = X_1 + X_2$ est donnée par la formule : $f_{Y_1}(y_1) = \int_{\mathbb{R}} f_{X_1}(y_1 - y_2) f_{X_2}(y_2)\, dy_2$

$$= \int_{\mathbb{R}} f_{X_2}(y_2) f_{X_1}(y_1 - y_2)\, dy_2 .$$

Inégalités importantes

1) L'inégalité de Markov :

Soit X une v a tq $(E(X)^k)$ existe $k>0$, Dans ce cas $\forall \varepsilon > 0$:

$$P(\omega : |X(\omega)| \geq \varepsilon) \leq \frac{E(|X|^k)}{\varepsilon^k}$$

Dém : $A = \{\omega : |X(\omega)| \geq \varepsilon\}$ alors :

$$E(|X|^k) = \int_{\Omega} |X(\omega)|^k \, dP(\omega) = \int_{A} |X(\omega)|^k \, dP(\omega)$$

$$+ \underbrace{\int_{A^c} |X(\omega)|^k \, dP(\omega)}_{\geq 0}$$

$$\geq \int_A |X(\omega)|^k \, dP(\omega) = \int_\Omega |X(\omega)|^k \, \mathbb{1}_{\|x\| \geq \varepsilon}(x) \, dP$$

$$\geq \varepsilon^k \int_A dP(\omega) = \varepsilon^k P(A)$$

$$P(A) \leq \frac{E(|X|^k)}{\varepsilon^k}.$$

2) L'inégalité de Tchebychev :

Soit Y une v.a tq $E(Y^2)$ existe ie Y^2 est intégrable.

Dans ce cas : $P(|Y(\omega) - E(Y(\omega))| \geq \varepsilon) \leq \frac{Var(Y)}{\varepsilon^2}$, $\forall \varepsilon > 0$

Dém : Posons $X = Y - E(Y)$ et $k = 2$ dans l'inégalité de Markov.

Définition :

Soit $f : \mathbb{R} \to \mathbb{R}$ un application, on dit que f est une fct convexe si :

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$$

$\forall (x, y) \in \mathbb{R}^2$ et $\alpha > 0$, $\beta \geq 0$, $\alpha + \beta = 1$

On va montrer qu'une fct f est convexe si et seulement si $\forall x_0$ fixé on a :

$$f(x) \geq f(x_0) + k(x_0)(x - x_0), \quad \forall x \in \mathbb{R}.$$

où $k(x_0)$ est un const dépendante de x_0.

$$f(x) \geq f(\blacksquare(x_0)) + k(x_0)(x - x_0)$$

$$\geq f(E(x)) + k(E(x))(x - E(x)) \qquad x_0 = E(X)$$

3) L'inégalité de Jenson :

. Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fct convexe de X une v.a sur $(\Omega, \mathcal{F}, P)$ tq $E(X)$ et $E(f(x))$ existe ① Dans ce cas $f(E(X)) \leq E(f(x))$

On pose $x_0 = E(X)$ dans ⊛ alors :

$f(x) \geq f(E(X)) + k(E(X))(x - E(X))$ alors $E(f(x)) \geq f(E(X))$

4) L'inégalité de Kolimogov :

. Soit $X_1, X_2, \ldots, X_n$ de v.a indépendantes sur $(\Omega, \mathcal{F}, P)$ tq $E(X_i) = 0$, $E(x_i^2) < \infty$, $i = \overline{1,n}$ alors $\forall \varepsilon > 0$,

$P(\max |X_1, \ldots, X_n| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \sum_{i=1}^{\infty} E(X_i^2)$.

Scanné avec CamScanner

Probabilités Conditionnelles

Soient $(\Omega, \mathcal{F}, P)$ un espace probabilité et $\mathcal{G}$ un algèbre finie sur Ω, $\mathcal{G} \in \mathcal{F}$ engendrée par une partie finie $\mathcal{B} = (B_1, B_2, \dots, B_n)$ de Ω. $\bigcup_{j=1}^{n} B_j = \Omega$, $B_j \in \mathcal{F}$.
$P(B_j) > 0$, $j = \overline{1,n}$, $B_i \cap B_j = \phi \ \forall i \neq j$
Il est connu que $\mathcal{G} = \mathcal{E}(\mathcal{B}) = \{B.\ B = \bigcup_{i,j} B_{ij},\ B_{ij} \in \mathcal{B}\}$
sous indices

. Soit $A \in \mathcal{F}$ considérons les probabilités conditionnelles

$$P(A/B_j) = \frac{P(A \cap B_j)}{P(B_j)} \quad ; \ j = 1, 2, \dots, n.$$

. En utilisant ces probabilités conditionnelles, nous pouvons construire une v.a discrète :

$$X(w) = \sum_{j=1}^{n} P(A/B_j)\, \mathbb{1}_{B_j}(w) \ / \ w \in \Omega$$

- Il est évidant que : $X(w) = P(A/B_j)$, $w \in B_j$, $j = \overline{1,n}$

Définition :

. La variable aléatoire $P(A/\mathcal{G})$ est appelée la probabilité conditionnelles de l'évènement A par rapport à l'algèbre finie $\mathcal{G} = \mathcal{E}(\mathcal{B})$

(ou par rapport à la partition $\mathcal{B}$)

• Il est évident que $\forall A \in \mathcal{F}$; $0 \leq P(A/\sigma)(\omega) \leq 1$

$\omega \in \Omega$

En particulier : $P(\cdot\ /\sigma)(\omega) \leftarrow$ est une v a / $\sigma \subseteq \mathcal{F}$

$P(A/\sigma)$ $\sigma \subseteq \mathcal{F}$

↑ v a ↑ sous tribu de $\mathcal{F}$

$\underline{P(A|\sigma)}$, $\sigma \in \mathcal{F}$

une vecteur ← un elements de $\mathcal{F}$. $P(A/\sigma(X))$; $\sigma(X)$ soustribu

$P(\phi\ /\sigma)(\omega) = 0$ et $(P(\omega/\sigma))(\omega) = 1$

Si $A, C \in \mathcal{F}$, et $A \cap C = \phi$ Alors : $P(A \cup C/\sigma)(\omega) = P(A/\sigma)(\omega)$

$+ P(C/\sigma)(\omega)$.

Remarque :

De la déf $P(A/\sigma)$, il suit que si σ est triviale $\sigma = \{\phi, \Omega\}$

alors : $P(A/\sigma) = P(A/\Omega) = P(A)$, $A \in \mathcal{F}$

• De la déf de l'esperence mathématique d'une v.a

Il suit que : $E(P(A|\sigma)) = P(A)$ ie :

$P(A|\sigma)$ est un estimateur sans biais par $P(A)$

Dém : $P(A\backslash\sigma)(\omega) = \sum_{j=1}^{n} P(A \backslash B_j) \mathbb{1}_{B_j}(\omega)$

$\Rightarrow$

$$E(P(A|\sigma)) = \sum_{j=1}^{n} P(A \backslash B_j) E(\mathbb{1}_{B_j}(\omega))$$

$$E[P(A/\sigma)] = P(A)\ ?$$

On a : $P(A/\sigma) = \sum P(A/B_j)\, \mathbb{1}_{B_j}(\omega)$

$$E(P(A/\sigma)) = \sum P(A/B_j)\, E[\mathbb{1}_{B_j}(\omega)]$$

$$= \sum \frac{P(A \cap B_j)}{\cancel{P(B_j)}} \cdot \cancel{P(B_j)}$$

$$E[P(A/\sigma)] = \sum P(A \cap B_j)$$

$$= P[\cup (A \cap B_j)]$$

$$= P[A \cap \underbrace{\cup B_j}_{\omega}]$$

$$= P(A \cap \omega) = P(A)$$

Scanné avec CamScanner

$$= \sum_{j=1}^{n} P(A \cap B_j) \frac{1}{\cancel{P(B_j)}} \cdot \cancel{P(B_j)}$$

$$= \sum_{j=1}^{n} P(A \cap B_j) = P(A \cap \bigcup_{j=1}^{n} B_j) = P(A)$$

$\cap$ disjoints

Remarque :

Soient X et Y deux v.a indép alors :

$P(X+t = z / Y = y)$ égale à $P(X + y = z)$ en effet :

$$P(X+Y=z / Y=y) = \frac{P(X+Y=z, Y=y)}{P(Y=y)}$$

$$= \frac{P(X+Y=z, t=y)}{P(Y=y)}$$

et comme X et Y sont indép, on a :

$$P(X+Y=z, Y=y) = P(X+Y=z) P(Y=y)$$

Soit X une v.a discrète sur $(\Omega, \mathcal{F}, P)$

$X \in \Lambda = \{x_1, \dots, x_n, \dots\}$. $(X : \Omega \longrightarrow \Lambda)$ et

$P(X = x_i) \neq 0$, $i \in \mathbb{N}^*$

Dans ce cas $\forall A \in \mathcal{F}$ on a :

$$P(A / X = x_n) = \frac{P(A \cap \{X = x_n\})}{P(X = x_n)}$$

Considérons la v.a :

$P(A\backslash X)(w) = P(A / X = x_n)(w)$ si $w \in \{w : X(w) = x_n\}$

Donc : $P(A \backslash X)(w)$ est une v.a discrète déterminée sur $(\Omega, \mathcal{F}, P)$

. La proba conditionnelle $P(A\backslash X)$ de A par rapport à la v.a X est appellée aussi proba conditionnelle de A par rapport de la σ-algèbre $G = \sigma(X)$

$$P(A\backslash X) = P(A / \sigma(X))$$

. Si $A \in G$, il existe une suite d'indices $\{j_k\}$ telle que

$$A = \bigcup_k \{X = x_{j_k}\}$$

$$P(A \backslash X = x_n) = \frac{P(\cup\{X = x_{j_n}\} \cap \{X = x_n\})}{P(X = x_n)} = \begin{cases} 1 & \text{si } n \in \{j_k\} \\ 0 & \text{sinon} \end{cases}$$

. Dans $\forall A \in \mathcal{F}$, $P(A\backslash G)(w) = \mathbb{1}_A(w) = \begin{cases} 1 & \text{si } w \in A \\ 0 & \text{sinon.} \end{cases}$

. Il est évident que $P(A/G)$ est une v.a G mesurable

Maintenant nous pouvons définir $P(A/G)$ dans le cas général

Remarque :

Soit $A \in \mathcal{F}$, si $P(A) = 1$, on dit que l'évenement A a lieu avec la probabilité 1 ou presque sûrement.

Définition :

On appelle proba conditionnelle $P(A \backslash G)$ d'un évènement $A \in \mathcal{F}$ par rapport à une σ-algèbre σ la v.a qui vérifie les conditions suivantes :

1) $P(A/G)$ est G mes

2) $\forall B \in G$, $P(A \cap B) = \int_B P(A \backslash G)(\omega) \, dP(\omega)$

vérifions que cette déf est correcte ($\Leftrightarrow$ déf 1 : triviale)

. Soit A fixe ds $\mathcal{F}$. Considérons sur (Ω, G) deux mesures $P_1(\cdot)$ et $P_2(\cdot)$ tq : $P_1(B) = P(B)$, $B \in G$

$P_2(B) = P(A \cap B)$, $B \in G$

. Puisque $P(A \cap B) \leq P(B)$ on a la mesure $P_2(\cdot)$ est absolument continue par rapport à $P_1(\cdot)$

Définition de abs continue de mesure :

On dit que μ est abs continue par rapport à λ si $\forall A \in \mathcal{F}$, $\lambda(A) = 0 \Rightarrow \mu(A) = 0$

$$\mu(A) = \int_A f \, d\lambda.$$

donc par Théorème de Radon-Nikodym il existe une fct G.mesurable $f(\omega)$ tq :

$$P_2(B) = \int_B f(\omega) \cdot dP_1(\omega), \quad \forall B \in G.$$

où $P(A\cap B) = \int_B f(\omega) \cdot dP(\omega)$

$P(A\cap B) = \int_B f(\omega) \cdot dP(\omega)$, $\forall B \in G$.

. On remarque que cette fct déf de façon unique P.P sur car si $\forall A \in \mathcal{F}$. $P(A\cap B) = \int_B f_1(\omega)\, dP(\omega) =$
$= \int_B f_2(\omega)\, dP(\omega)$

on a : $\int_B (f_1(\omega) - f_2(\omega))\, dP(\omega) = 0$, $\forall B \in G$.

donc : $f_1(\omega) = f_2(\omega)$ P.P sur

il reste à poser :
$P(A/G)(\omega) = f(\omega)$ pour avoir
$P(A\cap B) = \int_B P(A/G)\, dP$

On peut montrer que (presque sûrement).
1) $0 \leq P(A/G) \leq 1$, $\forall A \in \mathcal{F}$
2) $P(\Omega/G) = 1$
3) $P(\cup A_n / G) = \sum^{\infty} P(A_n/G)$.
Si la suite $\{A_n\}_{i=1}^{\infty}$ est tq $A_n \in G$, $A_i \cap A_j \neq \emptyset$, $i \neq j$

Démontrons la propriété (1).

Supposons qu'il existe $\varepsilon > 0$ tq : $P(B) > 0$ avec :

$B = \{\omega : P(A/G)(\omega) > 1 + \varepsilon\}$.

Notons que $P(A\backslash G)$ est G. mesurable et donc $B \in \mathcal{F}$

on a : $P(A \cap B) = \int_B P(A/G)(\omega)\,dP(\omega) \geq (1+\varepsilon)\int_B dP(\omega)$

$$= (1+\varepsilon)P(B)$$

et donc $P(B) \geq P(A \cap B) \geq (1+\varepsilon)P(B)$

Ce qui est possible si et seulement si $P(B) = 0$ donc $P\{(\omega), 0 \leq P(A/G) \leq 1)\} = 1$.

L'Espérance Conditionnelle.

Soit X une v.a sur $(\Omega, \mathcal{F}, P)$ tq $E(X)$ existe $(E(X) < \infty)$, soit G une σ-algèbre $G \subset \mathcal{F}$

Nous allons définir $E(X/G)$

Définition :

On appelle espérance conditionnelle $E(X/G)$ de la v.a qui véfie σ-algèbre G un v.a qui vérifie les conditions :

(1) $E(X/G)$ est G. mes

(2) $\int_B E(X/G)\,dP = \int_B X\,dP$ (*) $, \forall B \in G$.

Théorème :

Si $E(X)$ existe dans ce cas $E(X|G)$ existe et est unique P.P sûre

Dém :

Soit $Q(B) = \int_B X \, dP$, $B \in G$, il est évident que si $P(B) = 0$

Alors :

$Q(B) = 0$ et d'après le théorème de Radon Nikodym il existe une fct $Z(w)$ G.mesurable $Q(B) = \int_B Z(w) \, dP(w)$

Il nous reste à poser $E(X/G)(w) = Z(w)$ pour prouver l'unicité (P.P sûre) de $E(X/G)$ on remarque s'il existe 2 v.a $Z_1(w)$ et $Z_2(w)$ vérifiant

Définition :

EC ⊛ dans ce cas $\int (Z_1(w) - Z_2(w)) \, dP(w)$, $\forall B \in G$.

Ce qui est possible si B et seulement si $Z_1(w) = Z_2(w)$ P.P sûre

Remarque :

La proba conditionnelle peut être considérée comme un cas particulier de l'espérance conditionnelle

Soit

$$\mathbb{1}_A(w) = \begin{cases} 1 & \text{si } w \in A \\ 0 & \text{sinon} \end{cases}$$

Dans ce cas : $\boxed{P(A/G) = E(\mathbb{1}_A(w) / G)}$

$$\int_B X\,dP = \int_B E(X|G)\,dP$$

$$\int_B \mathbb{1}_A(\omega)\,dP(\omega) = \int_B E(\mathbb{1}_A|G)\,dP$$

$$)\,dP = P(A\cap B) = \int_B E(\mathbb{1}_A|G)\,dP.$$

Les Types de Convergence des suite de V.a

. Soient $(\Omega, \mathcal{F}, P)$ un espace probabilité

$X_n, X : (\Omega, \mathcal{F}, P) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $n \in \mathbb{N}^*$

des v.a sur $(\Omega, \mathcal{F}, P)$

On note v.a montrer que $\{\omega : X_n(\omega) \xrightarrow{n \to \infty} X(\omega)\} \in \mathcal{F}$

En effet : $X_n \longrightarrow X$, $\forall k \in \mathbb{N}^*, \exists m \in \mathbb{N}^* : |X_n - X| < \frac{1}{k}$ $\forall n \geq m$.

$$\{\omega : X_n(\omega) \longrightarrow X(\omega)\} = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m} \underbrace{\{\omega \; |X_n(\omega) - X(\omega)| < \frac{1}{k}\}}_{\in \mathcal{F}}.$$

comme $\{\omega : |X_n(\omega) - X(\omega)| < \frac{1}{k}\} \in \mathcal{F}$.

On remarque que :

$$\{\omega : X_n(\omega) \longrightarrow X(\omega)\} = \bigcap_{k=1}^{\infty} A_k$$

où :

$$A_k = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\omega : |X_n(\omega) - X(\omega)| < \frac{1}{k}\} = \liminf_n A_{nk}.$$

où $A_{nk} = \{\omega : |X_n(\omega) - X(\omega)| < \frac{1}{k}\}$.

Définition : (Convergence presque-sûre).

. On dit que la suite $\{X_n\}$ converge presque-sûre vers X lorsque $P\{\omega : X_n(\omega) \longrightarrow X(\omega)\} = 1$.

. On dit aussi que $\{X_n\}$ cv vers X avec une probabilité égale à 1.

et on note $X_n \xrightarrow{P.S} X$

. On peut montrer que $X_n \xrightarrow{P.S} X$, si et seulement si

$$P\left(\bigcup_n^{\infty} \{|X_n - X| \geq \varepsilon\}\right) \xrightarrow[n\to\infty]{} 0 \quad , \forall \varepsilon > 0$$

Comme $\{|X_n - X|\} = \bigcap_{k=1}^{\infty} \underbrace{\bigcup_{n=1}^{\infty} \bigcap_{m \geq n} \{|X_n - X| < \frac{1}{k}\}}_{A_k} = \bigcap_{k=1} A_k$

On sait :

$A_1 \supset A_2 \supset \dots$

et donc $P(\{X_n - X\}) = 1 \Leftrightarrow P(\bigcap_{k=1} A_k) = 1$

$\Leftrightarrow P(A_k) = 1, \forall k \in \mathbb{N}^* \Leftrightarrow \forall k \in \mathbb{N}^*, P(\bigcup_{n \geq 1}^{\infty} \bigcap_{m \geq n} \{|X_n - X| < \frac{1}{k}\})$

Notons :

$C_n = \bigcap \{|X_n - X| < \frac{1}{k}\}$ alors : $C_1 \subset C_2 \subset C_3 \dots$ et

$P\{|X_n - X|\} = 1 \Leftrightarrow \forall k \in \mathbb{N}^*, P(\bigcup_{n=1} C_n) = 1$

$\Leftrightarrow \forall k \in \mathbb{N}^*, \lim_{n\to\infty} P(C_n) = 1$

$\Leftrightarrow \forall k \in \mathbb{N}^*, \lim_{n\to\infty} P\left(\bigcap_{m \geq n} \{|X_n - X| < \frac{1}{k}\}\right)$

$\Leftrightarrow \forall k \in \mathbb{N}^*, \lim_{n\to\infty} P\left(\bigcup_{m \geq n} \{|X_m - X| \geq \frac{1}{k}\}\right) = 0$

Définition : (convergence en probabilité)

Soit $\{X_n\}$ une suite d'une v.a sur $(\Omega, \mathcal{F}, P)$

On dit que $\{X_n\}$ convergence en probabilité vers une v.a X et on note : $X_n \xrightarrow{P} X$ si :

$\forall \varepsilon > 0$ on a :

$$\lim_{n \to \infty} P(\{|X_n - X| > \varepsilon\}) = 0$$

Remarque :

On peut montrer que : $\sum_{n \geq 1} P\{|X_n - X| > \varepsilon\} < \infty$

alors $X_n \xrightarrow{P.S} X$ (Ex ϕ1)

supposons maintenant que les v.a $X_1, X_2, \dots$ sont indépendants

et que : $\sum_{n=1}^{\infty} P\{|X_n - X| > \varepsilon\} = +\infty$

Théorème :

. Dans l'espace probabilisé $(\Omega, \mathcal{F}, P)$ la c.v presque sûre implique la cv en proba.

$$X_n \xrightarrow{P.S} X \Rightarrow X_n \xrightarrow{P} X.$$

Preuve :

$$X_n \xrightarrow{P.S} X \Leftrightarrow P\left(\bigcup_{m \geq n} \{|X_m - X| > \varepsilon\}\right) \longrightarrow 0$$

$$\Rightarrow \lim P(\{|X_n - X| > \varepsilon\}) = 0$$

$$\Leftrightarrow X_n \xrightarrow{P} X$$

Remarque :

Soit $X_n \xrightarrow{P} X$ (C.V en proba)

Dans ce cas on peut extraire une sous suite $\{X_{n_k}\}$ telle que $X_{n_k} \longrightarrow X$

Définition : (Convergence en L_2)

Soit $\{X_n\}$ une suite de v.a sur $(\Omega, \mathcal{F}, P)$ tq :

$E(|X_n|^2) < +\infty$, $\forall n \geq 1$, et X une v.a, tq $E(|X|^2) < \infty$

On dit que $\{X_n\}$ C.V vers X en moyenne quadratique et on note $X_n \xrightarrow{m.q} X$ si : $E[|X_n - X|^2] \xrightarrow[n\to\infty]{} 0$

On peut montrer que $X_n \xrightarrow{m.q} X$

Si et seulement si : $\lim P\{|X_n - X_m| \geq \varepsilon\} = 0$ ie :

$\{X_n\}$ est une suite de Cauchy pour la C.V en proba

Théorème :

La C.V quadratique implique la C.V en proba

Preuve :

$$E(|X_n - X|^2) \geq \int_{\{|X_n - X| > \varepsilon\}} (x_n - x)^2 \, dP(x_n, x)$$

$$\geq \varepsilon^2 P(|X_n - X| \geq \varepsilon)$$

et donc $P(|X_n - X| \geq \varepsilon) = \frac{1}{\varepsilon^2} E(X_n - X)^2$
alors :

$$E(|X_n - X|^2) \longrightarrow 0 \text{ alors } P(|X_n - X| > \varepsilon) \xrightarrow[n \to \infty]{} 0$$

Définition : (Convergence en loi ou C.V faible).
On dit que la suite $\{x_n\}$ converge en loi vers la v.a X
et on note $X_n \xrightarrow{loi} X$ ou $X_n \xrightarrow{\mathcal{D}} X$
Distribution

Si $F_n(x) \longrightarrow F(x)$ en tout point
x de continuité de $F(x) = P(X \leq x)$ où $F_n(x) = P(X_n \leq x)$