Rappels de probabilités

Olivier FAUGERAS

12 Novembre 2008

Plan

Plan

Plan

Définitions

Moyennes

Densités

Indépendance

Le lemme de Borel-Cantelli

Fonctions caractéristiques

Plan

Plan

Plan

Espace de probabilités

Soit Ω un ensemble non vide.

Définition

Une tribu ou σ-algèbre est un ensemble $\mathcal{A}$ de sous-ensembles de Ω avec les propriétés suivantes

1. $\emptyset, \Omega \in \mathcal{A}$,
2. *Si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$,*
3. *Si $A_1, A_2, \cdots \in \mathcal{A}$, alors $\bigcup_{k=1}^{\infty} A_k, \bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$.*

Espace de probabilités

Définition

Soit $\mathcal{A}$ une σ-algèbre de Ω. Une application $P : \mathcal{A} \to [0, 1]$ est une mesure de probabilité sur $(\Omega, \mathcal{A})$ si

1. $P(\emptyset) = 0$, $P(\Omega) = 1$,
2. *Si $A_1, A_2, \dots \in \mathcal{A}$, alors*

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} P(A_k)$$

3. *Si $A_1, A_2, \dots \in \mathcal{A}$ sont disjoints alors*

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

Espace de probabilités

- Un élément A de $\mathcal{A}$ est un *événement*.
- $P(A)$ est la probabilité de l'événement A.
- Une propriété vraie sauf pour un événement de probabilité nulle est dite vraie presque sûrement (p.s.).
- Si $\Omega = \mathbb{R}^n$, la plus petite σ-algèbre, notée $\mathbb{B}$, contenant tous les ouverts de $\mathbb{R}^n$ s'appelle la tribu des boréliens.

Variables aléatoires

Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace de probabilité. Une application

$$\mathbf{X} : \Omega \to \mathbb{R}^n$$

est une variable aléatoire (v.a.) de dimension n si pour tout borélien $B \in \mathbb{B}$, $\mathbf{X}^{-1}(B) \in \mathcal{A}$. On dit aussi que $\mathbf{X}$ est $\mathcal{A}$-mesurable.

Lemme

Soit $\mathbf{X} : \Omega \to \mathbb{R}^n$ une variable aléatoire. Alors

$$\mathcal{A}(\mathbf{X}) = \{\mathbf{X}^{-1}(B) \mid B \in \mathbb{B}\}$$

est une σ-algèbre, dite engendrée par $\mathbf{X}$.

Variables aléatoires

- Soit $A \in \mathcal{A}$, la fonction indicatrice de A

$$\chi_A(\omega) = \begin{cases} 1 & \text{si} \quad \omega \in A \\ 0 & \text{sinon} \end{cases}$$

  est une v.a.
- Si $A_1, \ldots, A_m \in \mathcal{A}$ satisfont $\bigcup\limits_{k=1}^{m} A_k = \Omega$ et $a_1, \ldots, a_m$ sont des réels, $X = \sum_{k=1}^{m} a_i \chi_{A_i}$ est une v.a. appelée simple.

Processus stochastiques

Définition

1. *Un processus stochastique est une famille* $\{\mathbf{X}(t) \mid t \geq 0\}$ *de variables aléatoires indexée par le réel positif* t.
2. *Pour chaque point* ω *de* Ω*, la fonction* $t \to \mathbf{X}(t, \omega)$ *s'appelle une trajectoire.*

Processus stochastiques

Définition

Un processus $Y(\cdot)$ est une modification du processus $X(\cdot)$ si $P(Y_t = X_t) = 1$ pour tout $t \geq 0$.

Les processus X et Y sont aussi dits indiscernables.

Intégration

- Pour toute v.a. simple $X \sum_{k=1}^{m} a_k \chi_{A_k}$ on définit
$$\int_\Omega X\,dP = \sum_{k=1}^{m} a_k P(A_k)$$
- Si X est une v.a. réelle non négative on définit
$$\int_\Omega X\,dP = sup_{Y \leq X,\, Y\,\text{simple}} \int_\Omega Y\,dP$$
- Si X est une v.a. réelle on définit
$$\int_\Omega X\,dP = \int_\Omega X^+\,dP - \int_\Omega X^-\,dP,\; X^+ = \max(X,\,0)\; X^- = \max(-X$$
- Si $\mathbf{X}$ est une variable aléatoire de dimension n on définit $\int_\Omega \mathbf{X}\,dP = (\int_\Omega X_1\,dP, \cdots, \int_\Omega X_n\,dP)$

Intégration

Définition

L'espérance ou la valeur moyenne $\mathbb{E}(\mathbf{X})$ *d'une v.a. est l'intégrale* $\int_\Omega \mathbf{X}\, dP$*.*

Définition

La variance $V(\mathbf{X})$ *d'une v.a. est l'intégrale* $\int_\Omega |\,\mathbf{X} - \mathbb{E}(\mathbf{X})\,|^2\, dP$*.*

On remarque que $V(\mathbf{X}) = \mathbb{E}(|\,\mathbf{X}\,|^2) - |\,\mathbb{E}(\mathbf{X})\,|^2$.

Lemme (inégalité de Chebyshev)

Soit $\mathbf{X}$ *une v.a. réelle et p un réel tel que* $1 \leq p < \infty$*, alors*

$$P(|\,\mathbf{X}\,| \geq \lambda) \leq \frac{1}{\lambda^p}\mathbb{E}(|\,\mathbf{X}\,|^p)\ \forall \lambda > 0$$

Fonctions de distribution

Définition

1. *Soit* $\mathbf{X}$ *une v.a. de dimension n. La fonction de distribution de* $\mathbf{X}$ *est la fonction* $F_{\mathbf{X}} : \mathbb{R}^n \to [0,\ 1]$

$$F_{\mathbf{X}}(x) = P(\mathbf{X} \leq x) \, \forall x \in \mathbb{R}^n$$

2. *Soient* $X_1, \ldots, X_m$ *m v.a. réelles de dimension n. Leur fonction de répartition jointe est la fonction* $F_{\mathbf{X}_1,\cdots,\mathbf{X}_m} : (\mathbb{R}^n)^m \to [0,\ 1]$

$$F_{\mathbf{X}_1,\cdots,\mathbf{X}_m}(x_1, \cdots, x_m) = P(\mathbf{X}_1 \leq x_1, \cdots, \mathbf{X}_m \leq x_m)$$

Fonctions de densité

Définition

Soit $\mathbf{X}$ *une v.a. de dimension* n *et* F *sa fonction de distribution. S'il existe une fonction intégrable positive* $f : \mathbb{R}^n \to \mathbb{R}$ *telle que*

$$F(x) = F(x_1, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(y_1, \cdots, y_m)\, dy_n \cdots dy_1$$

on dit que f *est la fonction de densité de* $\mathbf{X}$ *(par rapport à la mesure de Lebesgue sur* $\mathbb{R}^n$*).*

Indépendance

Définition

Soient $A_1, \ldots, A_n, \ldots$ des événements. Ils sont dits indépendants si pour tous $1 \leq k_1 \leq \cdots \leq k_m$ on a

$$P(A_{k_1} \cap \cdots \cap A_{k_m}) = P(A_{k_1}) \cdots P(A_{k_m})$$

Indépendance

Définition

Soient $\mathcal{A}_i \subset \mathcal{A}$, $i = 1, \cdots$ des σ-algèbres. On dit qu'elles sont indépendantes si pour tous $1 \leq k_1 \leq \cdots \leq k_m$ pour tous événements $A_{k_i} \in \mathcal{A}_{k_i}$

$$P(A_{k_1} \cap \cdots \cap A_{k_m}) = P(A_{k_1}) \cdots P(A_{k_m})$$

Indépendance

Définition

Soient $\mathbf{X}_i$, $i = 1, \cdots$ des v.a. de dimension n. Elles sont indépendantes si pour tout entier $k \geq 2$ et quels que soient les boréliens $B_1, \cdots, B_k$ de $\mathbb{R}^n$ on a

$$P(\mathbf{X}_1 \in B_1, \cdots, \mathbf{X}_k \in B_k) = P(\mathbf{X}_1 \in B_1) \cdots P(\mathbf{X}_k \in B_k)$$

Indépendance

Les variables aléatoires $\mathbf{X}_1, \ldots, \mathbf{X}_m$ de dimension n sont indépendantes si et seulement si

$$F_{\mathbf{X}_1,\cdots,\mathbf{X}_m}(x_1,\cdots,x_m) = F_{\mathbf{X}_1}(x_1)\cdots F_{\mathbf{X}_m}(x_m)\,\forall x_i \in \mathbb{R}^n,\ i = 1,\cdots,m$$

Si elles ont des densités on a l'équivalence

$$f_{\mathbf{X}_1,\cdots,\mathbf{X}_m}(x_1,\cdots,x_m) = f_{\mathbf{X}_1}(x_1)\cdots f_{\mathbf{X}_m}(x_m)\,\forall x_i \in \mathbb{R}^n,\ i = 1,\cdots,m$$

Indépendance

Théorème

Si $X_1, \dots, X_m$ sont des v.a. réelles indépendantes telles que $\mathbb{E}(|X_i|) < \infty$, $i = 1, \cdots, m$ alors $\mathbb{E}(|X_1 \cdots X_m|) < \infty$ et

$$\mathbb{E}(X_1 \cdots X_m) = \mathbb{E}(X_1) \cdots \mathbb{E}(X_m)$$

Théorème

Si $X_1, \dots, X_m$ sont des v.a. réelles indépendantes telles que $V(|X_i|) < \infty$, $i = 1, \cdots, m$ alors

$$V(X_1 + \cdots + X_m) = V(X_1) + \cdots + V(X_m)$$

Le lemme de Borel-Cantelli

Définition

Soit A_n, $n \geq 1$ une suite d'événements d'un espace de probabilité. L'événement

$$\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \{\omega \in \Omega \,|\, \omega \in \text{ un nombre infini de } A_n\}$$

s'appelle "A_n infiniment souvent", en abrégé "A_n i.s.".

Lemme

Si $\sum_{n=1}^{\infty} P(A_n) < \infty$ alors $P(A_n \text{ i.s.}) = 0$.

Le lemme de Borel-Cantelli

Application : une suite de v.a. $\{X_k\}_{k=1}^{\infty}$ converge en probabilité vers une v.a. X si

$$\lim_{k\to\infty} P(|\, X_k - X \,| > \varepsilon) = 0$$

Théorème

Si $X_k \to X$ en probabilité, alors il existe une sous-suite $\{X_{k_j}\}_{j=1}^{\infty} \subset \{X_k\}_{k=1}^{\infty}$ telle que

$$X_{k_j}(\omega) \to X(\omega) \textit{ presque partout}$$

Le lemme de Borel-Cantelli

Preuve :
Pour chaque j on choisit $k_j > k_{j-1}$ suffisamment grand pour que $P(|X_{k_j} - X| > \frac{1}{j}) \leq \frac{1}{j^2}$. On considère la suite d'événements $A_j = \{|X_{k_j} - X| > \frac{1}{j}\}$ et $k_{j-1} < k_j < \cdots$. Puisque $\sum \frac{1}{j^2} < \infty$, le lemme de Borel-Cantelli implique $P(A_j \text{i.s.}) = 0$. Pour presque tous les échantillons ω on a donc $\left|X_{k_j}(\omega) - X(\omega)\right| < 1/j$ si $j \geq J$ pour un certain $J(\omega)$.

Fonctions caractéristiques

Définition

Soit X une v.a. à valeurs dans $\mathbb{R}^n$

$$\Phi_X(\lambda) = \mathbb{E}\left(e^{i\lambda \cdot X}\right) \quad \lambda \in \mathbb{R}^n$$

est la fonction caractéristique de X.

Fonctions caractéristiques

Lemme

1. *Si* $X_1, \cdots, X_m$ *sont des v.a. indépendantes à valeurs dans* $\mathbb{R}^n$ *alors*

$$\Phi_{X_1+\cdots+X_m}(\lambda) = \Phi_{X_1}(\lambda)\cdots\Phi_{X_m}(\lambda)$$

2. *Soit* X *une v.a.r.*

$$\Phi^{(k)}(0) = i^k\mathbb{E}(X^k) \quad k = 0, 1, \cdots$$

3. *Si* X *et* Y *sont des v.a. telles que*

$$\Phi_X(\lambda) = \Phi_Y(\lambda) \quad \forall\lambda$$

Alors