Points importants du cours 1

Vincent Guillemot

5 janvier 2009

1 Variables aléatoires réelles à densité

1.1 Notations, définitions

Définition 1 - *Variable aléatoire à densité*

Soit X une variable aléatoire réelle (v.a.r.) définie sur un espace probabilisé $(\Omega, \mathcal{T}, \Pr)$. On dit que X est une v.a.r. continue à densité si et seulement s'il existe une fonction $f : \mathbb{R} \to \mathbb{R}$ telle que

(i) f est positive,

(ii) f est continue sauf en un nombre fini de points,

(iii) en ces points, f admet une limite à gauche et à droite positive,

(iv) l'intégrale de f sur $\mathbb{R}$ existe et vaut 1,

(v) pour x dans $\mathbb{R}$,

$$\Pr(X \leq x) = F_X(x) = \int_{-\infty}^{x} f(t)dt.$$

f est appelée une densité de X.

Proposition 1

On peut définir une densité f d'une variable aléatoire X en dérivant une bonne fonction F. Les conditions pour une telle fonction $F : \mathbb{R} \to \mathbb{R}$ sont les suivantes :

(i) F croissante

(ii) $\lim_{t \to -\infty} F(t) = 0$ et $\lim_{t \to +\infty} F(t) = 1$

(iii) F est continue, de classe $\mathcal{C}_1$ sur $\mathbb{R}$ privé d'un nombre fini de points.

Proposition 2

Soit X une v.a.r. de densité f. Une v.a.r. à densité permet de mesurer la probabilité d'un intervalle :

$$\forall a < b \in \mathbb{R}, \Pr(a \leq X \leq b) = \int_{a}^{b} f(t)dt.$$

Remarque : Il faut bien retenir que pour x dans $\mathbb{R}$, $\Pr(X = \{x\}) = 0$: la probabilité d'un point est nulle.

Définition 2 - *Moment d'ordre k*

Soit X une v.a.r. définie sur un espace probabilisé $(\Omega, \mathcal{T}, \Pr)$. Le moment d'ordre k de X est le réel

$$m_k = \int_{-\infty}^{+\infty} t^k f(t)dt$$

sous réserve de convergence absolue de l'intégrale.

Définition 3 - *Espérance*

S'il existe, le moment d'ordre 1 m_1 d'une v.a.r. X est appelé son espérance et est noté $E(X)$ le plus souvent, parfois μ.

Définition 4 - *Moment centré d'ordre k*

Soit X une v.a.r. de densité f définie sur un espace probabilisé $(\Omega, \mathcal{T}, \Pr)$ d'espérance $E(X)$. Le moment centré d'ordre k de X est le réel

$$\mu_k = \int_{-\infty}^{+\infty} (t - E(X))^k f(t)dt$$

sous réserve de convergence absolue de l'intégrale

1.2 Variables usuelles

1.2.1 Variable aléatoire uniforme sur $[a;b]$

Définition 5

Une v.a.r. X à densité est dite uniforme de paramètres a et b réels (avec $a < b$) si et seulement si sa densité f est la suivante :

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \left\{ \begin{array}{l} \frac{1}{b-a} \text{ si } \in [a;b] \\ 0 \text{ sinon} \end{array} \right. \end{array}.$$

On note $X \sim \mathcal{U}(a, b)$.

1.2.2 Variable exponentielle de paramètre λ (non vu en cours)

Définition 6

Une v.a.r. X à densité est dite exponentielle de paramètre λ réel si et seulement si sa densité f est la suivante :

$$f : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \begin{cases} \lambda e^{-\lambda x} \text{ si } x > 0 \\ 0 \text{ sinon} \end{cases} \end{array}.$$

On note $X \sim \mathcal{E}(\lambda)$.

1.2.3 Variable gaussienne de paramètres μ et σ^2

Définition 7

Une v.a.r. X à densité est dite gaussienne de paramètres μ et σ réels ($\sigma > 0$) si et seulement si sa densité f est la suivante :

$$f : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \end{array}$$

On note $X \sim \mathcal{N}(\mu, \sigma^2)$.

1.2.4 Variable du Khi-deux à ν degrés de liberté

Définition 8

Soit ν un entier naturel non nul, et X_1, ..., X_ν des variables $\mathcal{N}(0,1)$ indépendantes. Alors la variable

$$X = \sum_{i=1}^{\nu} X_i^2$$

est une variable du Khi-deux à ν degrés de liberté. On note $X \sim \chi^2(\nu)$

1.2.5 Variable de Fisher à ν_1 et ν_2 degrés de liberté

Définition 9

Soit ν_1 et ν_2 deux entiers naturels non nuls, X_1 et X_2 deux variables indépendantes suivant une loi du Khi-deux à respectivement ν_1 et ν_2 degré de liberté. Alors la variable

$$X = \frac{X_1/\nu_1}{X_2/\nu_2}$$

est une variable de Fisher à ν_1 et ν_2 degrés de liberté. On note $X \sim \mathcal{F}(\nu_1, \nu_2)$.

1.2.6 Variable de Student à ν degrés de liberté

Définition 10

Soit ν un entier naturel non nul, deux variables Z et X indépendantes telles que $Z \sim \mathcal{N}(0,1)$ et $X \sim \chi^2(\nu)$. Alors la variable

$$T = \frac{Z}{\sqrt{\frac{1}{\nu}X}}$$

est une variable de Student à ν degrés de liberté. On note $T \sim \mathcal{T}(\nu)$.

1.3 Théorèmes limites

Théorème 1 - *Loi faible des grand nombres*

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. définies sur un espace probabilisé $(\Omega, \mathcal{T}, \Pr)$. Si on suppose qu'elles ont même espérance m et écart-type $\sigma > 0$, et qu'elles sont de plus deux à deux non corrélées, alors la variable $\frac{1}{n}\sum_{i=1}^{n} X_i$ converge en probabilité vers m.

Théorème 2 - *Théorème de la limite centrée*

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a.r. i.i.d. définies sur un espace probabilisé $(\Omega, \mathcal{T}, \Pr)$ et admettant une variance σ^2. Si les X_i sont mutuellement indépendantes, alors

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[loi]{} \mathcal{N}(m, \frac{1}{n}\sigma^2).$$

ATTENTION AUX APPROXIMATIONS, cela dépend des contraintes que l'on veut respecter. Selon les livres, l'approximation d'une v.a.r. de loi binomiale $\mathcal{B}(n,p)$ par une v.a.r. suivant une loi de Poisson $\mathcal{P}(\lambda = np)$ se fera pour $n > 30$ et $p < 1/10$, ou quand $n \geq 30$ et $p \approx 0.5$.

En cas de doutes ou de litiges, se référer à [1] qui produit le tableau d'approximations suivant :

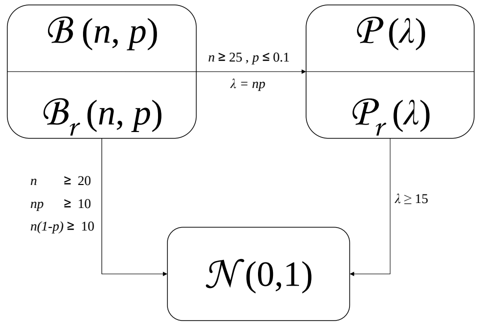

FIGURE 1 – Convergences en loi. Inspiré de la page 262 de [1].

2 Estimation ponctuelle

2.1 Notations, définitions et théorèmes préliminaires

Définition 11 - *Échantillon aléatoire*

Soit X une v.a.r. définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathrm{Pr})$. On appellera échantillon aléatoire de taille n de X (ou n-échantillon aléatoire de X) toute famille $(X_i)_{1 \leq i \leq n}$ de n v.a.r. i.i.d. de même loi que X.

Définition 12 - *Estimateur*

Un estimateur sur un n-échantillon $(X_i)_{1 \leq i \leq n}$ est une variable aléatoire Y fonction de toutes les v.a.r. du n-échantillon (sous la contrainte d'ordre très général que le résultat soit mesurable). Une estimation est la réalisation de cette variable Y.

Remarque : Un *estimateur* est une variable aléatoire, on peut donc déterminer son espérance ou sa variance. En revanche, une *estimation* est un réel fixe, déterministe, sans aucun caractère aléatoire.

Définition 13 - *Estimateur sans biais*

Un estimateur sans biais pour un paramètre θ est tel que $E(Y) = \theta$. Une réalisation d'un estimateur sans biais est appelée une estimation non biaisée pour θ.

Définition 14 - *Estimateur convergent*

Un estimateur est convergent si et seulement si il converge en probabilité vers la valeur du paramètre θ à estimer.

Théorème 3 - *Condition Suffisante de convergence*

Un estimateur sans biais dont la variance tend vers 0 quand la taille de l'échantillon tend vers $+\infty$ est un estimateur convergent.

Références

[1] Jean-Claude Droesbeke, *éléments de statistique*, Ellipses, 2002.