Leçon 261 Loi d'une variable aléatoire : caractérisations, exemples, applications

Dorian Cacitti-Holland

2020-2021

Références.

1. Probabilités tomes 1 et 2 de Jean-Yves Ouvrard
2. Probabilité de Barbe et Ledoux
3. Probabilités et statistiques de Chabanol et Ruch
4. Les contre-exemples en mathématiques de Bertrand Hauchecorne

Développements.

1. Théorème de Weierstrass par les polynômes de Bernstein
2. Théorème central limite et intervalle de confiance

Table des matières

1 Loi de variables aléatoires

1.1 Définition, cas discret et cas à densité

(Chapitres 8.2 de Probabilités tome 2 de Jean-Yves Ouvrard et II.3 et III.1 de Probabilité de Barbe et Ledoux)

On considère $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, $(E, \mathcal{B})$ un espace probabilisable et X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(E, \mathcal{B})$.

1. Définition : La loi de X est la mesure de probabilité $\mathbb{P}_X$ définie par $\forall B \in \mathcal{B}, \mathbb{P}_X(B) = \mathbb{P}(\{\omega \in \Omega, X(\omega) \in B\}) = \mathbb{P}(X \in B)$
2. Exemple : δ_x est la mesure de probabilité de la variable aléatoire constante x
3. Définition : On dit que la loi de X est discrète si $\mathbb{P}_X$ est une combinaison linéaire finie ou dénombrable de masse de Dirac, ie $\mathbb{P}_X = \sum\limits_{i \in I} p_i \delta_{x_i}$ avec $p_i = \mathbb{P}(X = x_i)$
4. Proposition : Les lois discrètes usuelles sont :
   — La loi uniforme discrète $\mathbb{P}_X = \sum\limits_{k=a}^{b} \frac{1}{b-a} \delta_k$
   — La loi de Bernoulli $\mathbb{P}_X = p\delta_1 + (1-p)\delta_0$
   — La loi binomiale est $\mathbb{P}_X = \sum\limits_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$
   — La loi géométrique est $\mathbb{P}_X = \sum\limits_{k=1}^{+\infty} (1-p)^{k-1} p \delta_k$
   — La loi de Poisson est $\mathbb{P}_X = \sum\limits_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} \delta_k$
5. Définition : On dit qu'une mesure ν est absolument continue par rapport à une mesure μ si $\mu(A) = 0 \Rightarrow \nu(A)$, on le note $\nu << \mu$
6. Théorème de Radon-Nikodyn : Soit μ et ν deux mesures σ-finies telles que $\nu << \mu$, alors il existe f mesurable positive telle que $\nu(A) = \int_A f d\mu$
7. Définition : On dit que la loi de X est à densité si $\mathbb{P}_X$ est absolument continue par rapport à la mesure Lebesgue et la fonction f du théorème précédent est appelée la densité de f
8. Proposition : Les lois à densité usuelles sont :
   — La loi uniforme de densité $\frac{1}{b-a} \mathbb{1}_{[a,b]}$
   — La loi exponentielle de densité $\lambda e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}$
   — La loi normale de densité $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$

1.2 Espérance d'une variable aléatoire

(Chapitre III.4 de Probabilité de Barbe et Ledoux)

1. Définition : Si X est intégrable pour la mesure $\mathbb{P}$ alors l'espérance de X est $\mathbb{E}(X) = \int_\Omega X d\mathbb{P} = \int_{X(\Omega)} x d\mathbb{P}_X(x)$

2. Théorème de transfert : Si X est à valeurs dans $\mathbb{R}$, soit $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ borélienne, alors $\phi(X) \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ si et seulement si $\phi \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$, dans ce cas $\mathbb{E}(\phi(X)) = \int_\Omega \phi \circ X(\omega) d\mathbb{P}(\omega) = \int_\mathbb{R} \phi(x) d\mathbb{P}_X(x)$
3. Remarque : Ce théorème permet de calculer $\mathbb{E}(\phi(X))$ sans connaître la loi de $\phi(X)$
4. Application : Théorème de Weierstrass : Les fonctions polynomiales sur $[0, 1]$ sont denses dans $C([0, 1])$ pour la norme uniforme
5. Corollaire : Soit $A \in \mathcal{B}(\mathbb{R})$, alors $\mathbb{E}(\mathbb{1}_A(X)) = \mathbb{P}(X \in A)$
6. Proposition : Les espérances des lois usuelles sont dans l'ordre $\frac{a+b}{2}, p, np, \frac{1}{p}, \lambda, \frac{a+b}{2}, \frac{1}{\lambda}, m$
7. Théorème : Si pour toute f borélienne bornée, $\mathbb{E}(f(X)) = \int_\Omega f d\mu$, alors $\mathbb{P}_X = \mu$
8. Exemple : Soit X, Y deux variables aléatoires indépendantes de densité $f(x) = \frac{1}{x^2}\mathbb{1}_{[1,+\infty[}(x)$, alors (X, Y) est de densité g sur $\mathbb{R}^2$ donnée par $g(x, y) = \frac{1}{2u^v}\mathbb{1}_\Delta(x, y)$ avec $\Delta = \{(x, y) \in \mathbb{R}^2, x > y > 0, xy > 1\}$

2 Caractérisations par des fonctions

2.1 Fonctions de répartition

(Chapitre III.2 de Probabilité de Barbe et Ledoux)
On considère X à valeurs réelles.

1. Définition : La fonction de répartition de X est $F_X(t) = \mathbb{P}(X \leq t)$
2. Proposition : $0 \leq F_X \leq 1$, F_X est CADLAG, $F_X(x) \underset{x\to+\infty}{\longrightarrow} 1, F_X(x) \underset{x\to-\infty}{\longrightarrow} 0$, réciproquement toute fonction F vérifiant ces conditions est une fonction de répartition d'une variable aléatoire réelle
3. Théorème : La fonction de répartition de X caractérise sa loi, ie si Y est une variable aléatoire réelle telle que $F_X = F_Y$ alors $\mathbb{P}_X = \mathbb{P}_Y$
4. Exemple : Si $F(t) = (1 - e^{-\theta t})\mathbb{1}_{\mathbb{R}_+}(t)$ alors F est la fonction de répartion de $\mathcal{E}(\theta)$
5. Exemple : Si $F(x) = \mathbb{1}_{[x,+\infty[}(x)$ alors F est la fonction de répartition de la loi δ_x
6. Proposition : Si X est à densité f alors $F(t) = \int_{-\infty}^t f(x) d\lambda(x)$, F est dérivable presque partout et $F'(t) = f(t)$
7. Exemple : Si $X \sim \mathcal{U}([0, 1])$ alors $F(x) = x\mathbb{1}_{[0,1]}(x) + \mathbb{1}_{]1,+\infty[}$

2.2 Fonctions caractéristiques

(Chapitre III.5 de Probabilité de Barbe et Ledoux)

1. Définition : La fonction caractéristique de X est la transformée de Fourier de la mesure $\mathbb{P}_X$, ie $\varphi_X(t) = \int_\mathbb{R} e^{itx} d\mathbb{P}_X(x) = \mathbb{E}(e^{itX})$
2. Exemple : Si X est de densité f alors $\varphi_X(t) = \int_\mathbb{R} e^{itx} f(x)dx = \mathcal{F}(f)(t)$
3. Théorème : La fonction caractéristique de X caractérise sa loi
4. Proposition : Si X, Y sont indépendantes alors $\varphi_{(X,Y)}(t, s) = \varphi_X(t)\varphi_Y(s)$ et $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$

5. Exemple : Si $X \sim \mathcal{B}(n,p)$ alors $\varphi_X(t) = (1-p+pe^{it})^n$
6. Exemple : Si $X \sim \mathcal{N}(0,1)$ alors $\varphi_X(t) = e^{-\frac{t^2}{2}}$
7. Théorème : Si $\mathbb{E}(|X|^n) < +\infty$ alors φ est n-fois dérivable et $\varphi^{(n)}(0) = i^k \mathbb{E}(X^k)$, réciproquement si n pair et φ n-fois dérivable en 0 alors $\mathbb{E}(|X|^n) < +\infty$

2.3 Fonctions génératrices

(Chapitre 5.3 de Probabilités tome 1 de Jean-Yves Ouvrard)
On considère X variable aléatoire à valeurs dans $\mathbb{N}$.

1. Définition : $G_X = \sum\limits_{n=0}^{+\infty} \mathbb{P}(X=n)s^n$
2. Exemple : Si $X \sim \mathcal{B}(n,p)$ alors $G_X = (sp+1-p)^n$
3. Théorème : G_X est continue sur $[-1,1]$ et de classe C^∞ sur $]-1,1[$ et $\forall n \in \mathbb{N}, \mathbb{P}(X = n) = \frac{G^{(n)}(0)}{n!}$
4. Corollaire : G_X caractérise la loi de X
5. Proposition : Si X, Y indépendantes alors $G_{X+Y}(t) = G_X(t)G_Y(t)$
6. Théorème : X admet un moment d'ordre k si et seulement si G_X est k-fois dérivable à gauche en 1, dans ce cas pour $k=1$, $\mathbb{E}(X) = G'_X(1)$
7. Application : Soit $(X_n)_{n\in\mathbb{N}}$ suite de variables aléatoires discrètes, T variable aléatoire dans $\mathbb{N}$ toutes indépendantes, $S_n = \sum\limits_{k=1}^{n} X_k$ et $S = S_T$, alors $G_S = G_T \circ G_{X_1}$

3 Lois des vecteurs aléatoires

3.1 Lois conjointes et marginales

(Chapitre III.3 de Probabilité de Barbe et Ledoux)
On considère X à valeurs dans $\mathbb{R}^d$ muni de sa tribu borélienne.

1. Définition : La fonction de répartition de X est $F_X(t) = \mathbb{P}(X_1 \leq t_1, ..., X_d \leq t_d)$
2. Proposition : La loi de X_i, appelée i-ième loi marginale de X, est donnée par $F_{X_i}(t_i) = \underset{t_1,...,t_d \to +\infty}{lim} F_X(t)$
3. Exemple : Si $X = (X_1, X_2)$ de loi discrète concentrée sur les points $(-1,0), (0,1), (0,-1), (1,0)$ tous de probabilité $\frac{1}{4}$ alors $\mathbb{P}_{X_1} = \mathbb{P}_{X_2} = \frac{1}{4}\delta_{-1} + \frac{1}{2}\delta_0 + \frac{1}{4}\delta_1$
4. Proposition : Si $X = (X_1, X_2)$ admet une densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$ alors X_1 et X_2 sont de densité $f_1(x) = \int_\mathbb{R} f(x,y)dy$ et $f_2(y) = \int_\mathbb{R} f(x,y)dx$
5. Exemple : Si $f(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{\|x\|^2}{2}}$ alors $X_i \sim \mathcal{N}(0,1)$
6. Théorème : La fonction de répartition de X caractérise sa loi

3.2 Cas particulier des vecteurs gaussiens

(Chapitre IX.2 de Probabilités et statistiques de Chabanol et Ruch)

On considère X un vecteur aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d$.

1. Définition : On dit que X est un vecteur gaussien si $\forall a \in \mathbb{R}^d, a^t X \sim \mathcal{N}(m, \sigma^2)$
2. Remarque : Si X est gaussien alors les composantes de X sont gaussiennes mais l'inverse est faux
3. Exemple : Soit $Y \sim \mathcal{N}(0,1)$ et $\varepsilon \sim \frac{1}{2}(\delta_{-1} + \delta_1)$ indépendantes alors $Y, \varepsilon Y \sim \mathcal{N}(0,1)$ mais $(Y, \varepsilon Y)$ n'est pas un vecteur gaussien
4. Théorème : Si $X \sim \mathcal{N}_d(m, \Gamma)$ alors $\varphi_X(u) = exp\left(iu^t m - \frac{1}{2} u^t \Gamma u\right)$
5. Théorème : Si $X \sim \mathcal{N}_d(m, \Gamma)$ avec Γ inversible alors $f_X(x) = \frac{1}{\sqrt{(2\pi)^d det(\Gamma)}} exp\left(-\frac{1}{2}(x-m)^t \Gamma^{-1}(x-m)\right)$
6. Corollaire : Théorème de Cochran : Si $X \sim \mathcal{N}_n(m, \sigma^2 I_n)$, $\mathbb{R}^n = E_1 \oplus ... \oplus E_p$, π_k la matrice du projeteur orthogonal sur E_k et $Y_k = \pi_k X$, alors les $Y_1, ..., Y_p$ sont indépendants de loi $\mathcal{N}_d(\pi_k m, \sigma^2 \pi_k)$ et les $\|Y_1 - \pi_1 m\|^2, ..., \|Y_p - \pi_p m\|^2$ sont indépendants et $\frac{1}{\sigma^2}\|Y_k - \pi_k m\|^2 \sim \chi^2(d_k)$
7. Application : Soit $(X_1, ..., X_n)$ un n-échantillon de loi $\mathcal{N}(m, \sigma^2)$, $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ et $\overline{S^2}_2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$, alors $\overline{X}_n$ et $\overline{S_n}^2$ sont indépendants, de plus $\sqrt{n}\overline{X}_n \sim \mathcal{N}(m, \sigma^2)$ et $\frac{n-1}{\sigma^2}\overline{S^2}_n \sim \chi^2(n-1)$

3.3 Différence entre indépendance et non corrélation

(Chapitres IV.1 de Probabilités de Barbe et Ledoux et IX.2 de Probabilités et statistiques de Chabanol et Ruch)

On considère X et Y deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$.

1. Définition : On dit que $(X, Y) \in L^2$ sont non corrélées si $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$
2. Remarque : Autrement dit $\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = 0$ et on dit que $X - \mathbb{E}(X)$ et $Y - \mathbb{E}(Y)$ sont orthogonales
3. Proposition : Si X et Y sont indépendantes alors X et Y sont non corrélées
4. Remarque : La réciproque est fausse en générale
5. Exemple : Si $X \sim \mathcal{N}(0,1)$ et $Y = X^2$ alors X et Y sont non corrélées mais ne sont pas indépendantes
6. Exemple : Si $\mathbb{P}_X = \frac{1}{3}(\delta_{-1} + \delta_0 + \delta_1)$ et $Y = X^2$ alors X et Y sont non corrélées mais ne sont pas indépendantes (Exemple 19 7 des Contre-exemples en mathématiques de Bertrand Hauchecorne)
7. Théorème : Si X gaussien alors les assertions suivantes sont équivalentes :
   — $(X_1, ..., X_d)$ sont mutuellement indépendantes
   — $(X_1, ..., X_d)$ sont deux à deux indépendantes
   — $(X_1, ..., X_d)$ sont deux à deux non corrélées
   — Γ est diagonale

4 Convergence en loi de suites de variables aléatoires

4.1 Définition et propriétés

(Chapitres V.4 de Probabilités de Barbe et Ledoux et IV.4 et IV.7 de Probabilités et statistiques de Chabanol et Ruch)

On considère $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires.

1. Défintion : On dit que $X_n \xrightarrow[n\to+\infty]{loi} X$ si $\forall\varphi \in C_b(\mathbb{R}^d), \mathbb{E}(\varphi(X_n)) \underset{n\to+\infty}{\longrightarrow} \mathbb{E}(\varphi(X_n))$
2. Remarque : La limite X n'est pas unique, seule sa loi l'est
3. Exemple : Si $S_n \sim \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ alors $S_n \xrightarrow[n\to+\infty]{loi} \mathcal{P}(\lambda)$ (Chapitre IV.7 de Probabilités et statistiques de Chabanol et Ruch)
4. Théorème de Lévy : $X_n \xrightarrow[n\to+\infty]{loi} X \iff \forall t \in \mathbb{R}^d, \phi_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} \phi_X(t)$
5. Théorème : Si $\forall t \in \mathbb{R}^d, \phi_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} \phi(t)$ avec ϕ continue en 0 alors ϕ est la fonction caractéristique d'une variable aléatoire X et $X_n \xrightarrow[n\to+\infty]{loi} X$
6. Théorème : $X_n \xrightarrow[n\to+\infty]{loi} X$ si et seulement si pour tout point t de continuité de F_X, $F_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} F_X(t)$
7. Exemple : Si $(X_n)_{n\in\mathbb{N}}$ de lois respectives $\delta_{\frac{1}{n}}$ alors $X_n \xrightarrow[n\to+\infty]{loi} X$ avec $\mathcal{L}(X) = \delta_0$
8. Proposition : Si $(X_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{N}$ et X à valeurs dans $\mathbb{N}$ alors $X_n \xrightarrow[n\to+\infty]{loi} X \iff \forall k \in \mathbb{N}, \mathbb{P}(X_n = k) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(X = k)$
9. Proposition : Si $X_n \xrightarrow[n\to+\infty]{loi} X$ et $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ continue alors $f(X_n) \xrightarrow[n\to+\infty]{loi} f(X)$
10. Lemme de Scheffé en loi : Si $(X_n)_{n\in\mathbb{N}}$ de densités respectives f_n et X de densité f tel que pour presque tout $x \in \mathbb{R}, f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x)$ alors $X_n \xrightarrow[n\to+\infty]{loi} X$
11. Théorème : $(G_{X_n})_{n\in\mathbb{N}}$ converge simplement vers G_X si et seulement si $X_n \xrightarrow[n\to+\infty]{loi} X$
12. Application : Si $X_n \sim \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ alors $X_n \xrightarrow[n\to+\infty]{loi} \mathcal{P}(\lambda)$

4.2 Lien avec les autres types de convergence

(Chapitres V.2, V.3 et V.4 de Probabilités de Barbe et Ledoux et IV.5 de Probabilités et statistiques de Chabanol et Ruch)

1. Proposition : Si $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ alors $X_n \xrightarrow[n\to+\infty]{loi} X$
2. Remarque : La réciproque est fausse
3. Exemple : Si $(X_n)_{n\in\mathbb{N}}$ IID de loi $\mathcal{B}(p)$ alors $X_n \xrightarrow[n\to+\infty]{loi} X$ avec $\mathcal{L}(X) = \mathcal{B}(p)$ mais ne converge pas en probabilité (Exemple 19.10 des Contre-exemles de Bertrand Hauchecorne)

4. Corollaire : La convergence presque sûre et les convergences L^p implique la convergence en loi
5. Proposition : Si $X_n \xrightarrow[n\to+\infty]{loi} X$ alors il existe $(\Omega', \mathcal{A}', \mathbb{P}'), X'_n, X'$ tel que $X_n \overset{loi}{=} X'_n, X \overset{loi}{=} X'$ et $X'_n \xrightarrow[n\to+\infty]{PS} X'$
6. Proposition : Si $X_n \xrightarrow[n\to+\infty]{loi} c$ constante alors $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} c$

4.3 Théorème central limite et lemme de Slutsky

(Chapitres V.5 de Probabilités de Barbe et Ledoux et IV.6 de Probabilités et statistiques)

1. Théorème central limite : Si $(X_n)_{n\in\mathbb{N}}$ IID de moyennes m et de variance σ^2 alors $\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n} X_k - nm\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0,\sigma^2)$
2. Corollaire : Si $\sigma^2 \neq 0$ alors $\frac{1}{\sqrt{n\sigma^2}}\left(\sum_{k=1}^{n} X_k - nm\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0,1)$
3. Application : La moyenne empirique $\hat{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ est un estimateur sans biais, fortement consistent et asymptotiquement normal de $m = \mathbb{E}(X_1)$, et si σ^2 connue alors un interalle de confiance asymptotique est $ICA_{1-\alpha}(m) = \left[\hat{X}_n - \sqrt{\frac{n}{\sigma^2}} q_{1-\frac{\alpha}{2}}, \hat{X}_n + \sqrt{\frac{n}{\sigma^2}} q_{1-\frac{\alpha}{2}}\right]$ avec $q_{1-\frac{\alpha}{2}}$ quantile d'ordre $1-\frac{\alpha}{2}$ de $\mathcal{N}(0,1)$
4. Proposition : δ-méthode : Si $\sqrt{n}(X_n - \theta) \xrightarrow[n\to+\infty]{} \mathcal{N}(0,\sigma^2)$ et $g : \mathbb{R} \longrightarrow \mathbb{R}$ dérivable en θ de dérivée non nulle alors $\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0, \sigma^2 g'(\theta)^2)$
5. Lemme de Slutsky : Si $X_n \xrightarrow[n\to+\infty]{loi} X, Y_n \xrightarrow[n\to+\infty]{loi} a$ alors $(X_n, Y_n) \xrightarrow[n\to+\infty]{loi} (X, a)$
6. Corollaire : Dans ce cas $X_n + Y_n \xrightarrow[n\to+\infty]{loi} X + a$ et $X_n Y_n \xrightarrow[n\to+\infty]{} Xa$
7. Application : Ce lemme est utile pour remplacer une donnée inconnue dans un intervalle de confiance par un estimateur de cette inconnue