Développement: Théorème des évènements rares de Poisson

Adrien Fontaine

7 septembre 2013

Référence : Ouvrard 2, théorème 14.20p321 pour la première version du théorème et théorème 14.11p313 pour le théorème de Lévy
Durett pour la deuxième version du théorème

1 Développement

Dans ce document, on va donner deux preuves du théorème des évènements rares de Poisson. La première est la plus classique. Cependant, elle présente (à mon sens) de nombreux inconvénients. Tout d'abord, elle "cache" quelques subtilités sur lesquelles il est facile de se faire piéger : utilisation du logarithme complexe, formule de Taylor avec reste intégral... Par ailleurs, elle utilise le théorème de Lévy qui est un résultat non trivial. Enfin, cette preuve ne donne aucune information sur la vitesse de convergence. C'est pourquoi, je donne une seconde preuve, qui a, elle, le mérite de donner des informations sur la vitesse de convergence et d'éviter les "pièges" de la première démonstration. Commençons par donner l'énoncé du théorème :

Théorème 1 (*Théorème des évènements rares de Poisson*)

Soit, pour tout $n \in \mathbb{N}^*$, une famille finie $(X_{n,j})_{1\leq j\leq M_n}$ de variables aléatoires de loi de Bernoulli de paramètre $p_{n,j}$.

$$\mathbb{P}(X_{n,j} = 1) \qquad \mathbb{P}(X_{n,j} = 0) = 1 - p_{n,j}$$

On suppose que

(i) M_n *tend en croissant vers* $+\infty$.

(ii) $\sum_{j=1}^{M_n} p_{n,j} \to \lambda > 0$ *quand* $n \to +\infty$

(iii) $\max_{1\leq j\leq M_n} p_{n,j} \to 0$

Alors, si $S_n = X_{n,1} + ... + X_{n,M_n}$, *alors* $S_n \xrightarrow[n\to\infty]{\mathcal{L}} Z$, *où* $Z \sim \mathcal{P}(\lambda)$.

1.1 Première démonstration

Démonstration : On va utiliser le théorème de Lévy. Par indépendance des $A_{n,j}$, $1 \le j \le M_n$, on a pour tout $t \in \mathbb{R}$,

$$\begin{aligned}\varphi_{S_n}(t) &= \prod_{j=1}^{M_n} \varphi_{X_{n,j}}(t) \\ &= \prod_{j=1}^{M_n} [p_{n,j}\exp(it) + (1-p_{n,j})] \\ &= \prod_{j=1}^{M_n} [1 + p_{n,j}(\exp(it)-1)]\end{aligned}$$

Si log est la détermination principale du logarithme complexe, il résulte de la formule de Taylor avec reste intégral que, pour tout z tel que $|z| < 1$, on a :

$$\log(1+z) = z - z^2 \int_0^1 (1-u)\frac{1}{(1+uz)^2}du$$

Notons $z = \exp(it) - 1$. Puisque $\max_{1\le j\le M_n} p_{n,j} \to 0$, il existe N tel que, pour tout $n \ge N$, on ait $\max_{1\le j\le M_n} |p_{n,j}z| < 1/2$. Pour tout $n \ge N$, on a alors

$$\log \varphi_{S_n}(t) = z\sum_{j=1}^{M_n} p_{n,j} - z^2 \sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{(1+up_{n,j}z)^2}du$$

D'après l'inégalité triangulaire, on a, pour tout $n \ge N$, et pour tout $u \in [0,1]$,

$$|1+up_{n,j}z| \ge 1 - p_{n,j}|z| \ge \frac{1}{2}$$

on a donc, pour tout $n \ge N$,

$$|\sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{(1+up_{n,j}z)^2}du| \le 2\left[\max_{1\le j\le M_n} p_{n,j}\right]\left[\sum_{j=1}^{M_n} p_{n,j}\right]$$

Il résulte alors des hypothèses que $\lim_{n\to+\infty} \log \varphi_{S_n}(t) = \lambda z$. Autrement dit, on a :

$$\forall t \in \mathbb{R}, \lim_{n\to\infty} \varphi_{S_n}(t) = \exp[\lambda(\exp(it)-1)]$$

ce qui, en vertu du théorème de Lévy, démontre le résultat. ∎

1.2 Deuxième démonstration

Démonstration : Pour μ et ν deux mesures sur un ensemble dénombrable S, on note [1]

$$\|\mu - \nu\| := \sum_{z\in S} |\mu(\{z\}) - \nu(\{z\})| = 2.\sup_{A\subset S} |\mu(A) - \nu(A)|$$

1. il s'agit de la distance en variation totale entre μ et ν. C'est une distance sur l'ensemble des mesures de probabilité sur un ensemble donné, qui est compatible avec la convergence en loi. Mais, on a pas besoin de savoir tout ça pour faire cette démonstration.

La première égalité est une définition. Pour obtenir la deuxième, on remarque que pour tout $A \subset S$,

$$\begin{aligned}\sum_{z\in S} |\mu(z) - \nu(z)| &\geq |\mu(A) - \nu(A)| + |\mu({}^cA) - \nu({}^cA)| \\ &= 2|\mu(A) - \nu(A)|\end{aligned}$$

et il y a égalité quand $A = \{z/\mu(\{z\}) \geq \nu(\{z\})\}$.
On démontre ensuite trois lemmes :

Lemme 1

Si $\mu_1 \times \mu_2$ désigne la mesure produit sur $\mathbb{Z} \times \mathbb{Z}$, i.e

$$(\mu_1 \times \mu_2)(x, y) = \mu_1(x)\mu_2(y)$$

alors,

$$\|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\| \leq \|\mu_1 - \nu_1\| + \|\mu_2 - \nu_2\|$$

Démonstration :

$$\begin{aligned}&\|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\| = \sum_{(x,y)\in\mathbb{Z}^2} |\mu_1(x)\mu_2(y) - \nu_1(x)\nu_2(y)| \\ &\leq \sum_{(x,y)\in\mathbb{Z}^2} |\mu_1(x)\mu_2(y) - \nu_1(x)\mu_2(y)| + \sum_{(x,y)\in\mathbb{Z}^2} |\nu_1(x)\mu_2(y) - \nu_1(x)\nu_2(y)| \\ &= \sum_{y\in\mathbb{Z}} \mu_2(y) \sum_{x\in\mathbb{Z}} |\mu_1(x) - \nu_1(x)| + \sum_{x\in\mathbb{Z}} \nu_1(x) \sum_{y\in\mathbb{Z}} |\mu_2(y) - \nu_2(y)| \\ &= \|\mu_1 - \nu_1\| + \|\mu_2 - \nu_2\|\end{aligned}$$

■

Lemme 2

Si $\mu_1 * \mu_2$ désigne le produit de convolution de μ_1 et μ_2, i.e

$$\mu_1 * \mu_2(x) \sum_{y\in\mathbb{Z}} \mu_1(x - y)\mu_2(y)$$

alors,

$$\|\mu_1 * \mu_2 - \nu_1 * \nu_2\| \leq \|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\|$$

Démonstration :

$$\begin{aligned}&\|\mu_1 * \mu_2 - \nu_1 * \nu_2\| \\ &= \sum_{x\in\mathbb{Z}} |\sum_{y\in\mathbb{Z}} \mu_1(x - y)\mu_2(y) - \sum_{y\in\mathbb{Z}} \nu_1(x - y)\nu_2(y)| \\ &\leq \sum_{x\in\mathbb{Z}} \sum_{y\in\mathbb{Z}} |\mu_1(x - y)\mu_2(y) - \nu_1(x - y)\nu_2(y)| \\ &= \|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\|\end{aligned}$$

■

Lemme 3

Soit μ la mesure définie par $\mu_(1) = p$ et $\mu(0) = 1 - p$. Soit ν une loi de Poisson de paramètre p. Alors, $\|\mu - \nu\| \leq 2p^2$.

Démonstration :

$$\begin{aligned}\|\mu-\nu\| &= |\mu(0)-\nu(0)|+|\mu(1)-\nu(1)|+\sum_{n\geq 2}\nu(n)\\ &= |1-p-e^{-p}|+|p-pe^{-p}|+1-e^{-p}(1+p)\end{aligned}$$

Or, $1-x\leq e^{-x}$ pour tout $x\geq 0$, donc

$$\begin{aligned}\|\mu-\nu\| &= e^{-p}-1+p+p(1-e^{-p})+1-e^{-p}-pe^{-p}\\ &= 2p(1-e^{-p})\\ &\leq 2p^2\end{aligned}$$

■

On peut maintenant passer à la preuve du théorème de Poisson à proprement dite.
Soit $\mu_{n,j}$ la loi de $X_{n,j}$. Soit μ_n la loi de S_n.
Soient $\nu_{n,j}$, ν_n et ν des lois de Poisson de paramètre $p_{n,j}$, $\lambda_n=\sum_{j\leq M_n}p_{n,j}$ et λ respectivement.
Alors, comme $\mu_n=\mu_{n,1}*...*\mu_{n,M_n}$ et $\nu_n=\nu_{n,1}*...*\nu_{n,M_n}$, les lemmes 1,2 et 3 montrent que

$$\|\mu_n-\nu_n\|\leq\sum_{j=1}^{M_n}\|\mu_{n,j}-\nu_{n,j}\|\leq 2\sum_{j=1}^{M_n}p_{n,j}^2$$

D'où, en utilisant la définition de distance en variation totale

$$\begin{aligned}\sup_A|\mu_n(A)-\nu_n(A)| &\leq \sum_{j=1}^{M_n}p_{n,j}^2\\ &\leq \underbrace{\left(\max_{1\leq j\leq M_n}p_{n,j}\right)}_{\to 0}\underbrace{\sum_{j=1}^{M_n}p_{n,j}}_{\to\lambda}\to 0\end{aligned}$$

Or, $\nu_n \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \nu$, donc $\forall x\in\mathbb{N},\nu_n(x)\to\nu(x)$.
Donc, si on fixe $x\in\mathbb{N}$, on a :

$$\begin{aligned}|\mu_n(x)-\nu_n(x)|\to 0\\ |\nu_n(x)-\nu(x)|\to 0\end{aligned}$$

Donc, $\mu_n(x)\to\nu(x)$. Donc, $\mu_n \underset{n\to\infty}{\overset{\mathcal{L}}{\to}} \nu$ ■

2 Annexe

2.1 Remarques

Remarque 1

Le théorème des évènements rares de Poisson tire son nom du fait qu'il montre qu'un phénomène aléatoire qui peut se représenter comme une superposition d'évènements rares (c'est à dire d'évènements de petites probabilité, au sens des conditions du théorème) et indépendants, suit approximativement une loi de Poisson.

Remarque 2

L'avantage de la deuxième preuve est qu'elle fournit des informations sur la vitesse de convergence.
En effet, si $p_{n,j} = \frac{1}{n}$ pour $1 \leq j \leq n$, la dernière majoration devient

$$\sup_A |\mu_n(A) - \nu_n(A)| \leq \frac{1}{n}$$

Cette majoration est plutôt bonne, car on peut en fait montrer que

$$\sup_{A \subset \mathbb{Z}} |\mu_n(A) - \nu_n(A)| \sim \frac{3}{4en}$$

Exemple 1

Dans une classe de 400 élèves, le nombre d'étudiants ayant leur anniversaire le jour de l'examen final a approximativement une loi de Poisson de paramètre $\frac{400}{365} = 1,096$.

2.2 Théorème de Lévy

Théorème 2 (*Théorème de Lévy*)

Soit pour tout $n \in \mathbb{N}$, une variable aléatoire X_n, définie sur un espace probabilisé $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$, à valeurs dans $\mathbb{R}^d$, de fonction caractéristique φ_{X_n}.

1. *Si la suite de variables aléatoires (X_n) converge en loi vers X, où X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{R}^d$, alors la suite (φ_{X_n}) des fonctions caractéristiques converge simplement (et même uniformément sur tout compact de $\mathbb{R}^d$) vers la fonction caractéristique φ_X de X.*
2. *Inversement, si la suite φ_{X_n} des fonctions caractéristiques converge simplement vers une fonction φ continue en 0, alors φ est la transformée de Fourier d'une probabilité μ sur $\mathbb{R}^d$, et la suite des variables aléatoires (X_n) converge en loi vers μ.*
   De plus, il existe une variable aléatoire (non unique) X définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{R}^d$, telle que la suite de variables aléatoires (X_n) converge en loi vers X.