Feuille de TP n^o 3-Statistiques d'ordre et processus de Poisson

Soit (X_n) une suite de variables aléatoires réelles et X une autre variable aléatoire réelle. On dit que (X_n) converge en loi vers X, on note

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}} X,$$

si la fonction de répartition de X_n converge simplement vers la fonction de répartition X en tout point où elle est continue. On pourra illustrer cette convergence en tracant sur le même graphique la fonction de répartition empirique d'un m-échantillon de X_n pour n assez grand et la fonction de répartition de X. Ou encore, on pourra tracer un histogramme d'un m-échantilonn de X_n pour n assez grand et la densité de X si elle existe sur le même graphique.

1 Statistiques d'ordre

Soit X une variable aléatoire réelle de fonction de répartition F. On suppose ici que F est continue. On note F^- son inverse généralisé. On pose, pour tout $p \in]0,1[$, $k_p = F^-(p)$ le quantile d'ordre p ($p = \frac{1}{2}$ la médiane, $p = \frac{1}{4}$ le quartile etc...) Soit $(X_1(\omega), \cdots, X_n(\omega))$ un n-échantillon de X.

▶ Montrer que pour tout $i, j \in \{1, \cdots, n\}$, $i \neq j$, $\mathbb{P}(X_i = X_j) = 0$. Ainsi, presque surement, il existe une permutation σ de $\{1, \cdots, n\}$ telle que

$$X_{\sigma(1)}(\omega) < \cdots < X_{\sigma(n)}(\omega).$$

On posera pour $i \in \{1, \cdots, n\}$, $X_{(i)} = X_{\sigma(i)}$. On appel $X_{(i)}$ la $i$ème statisque d'ordre, et $X_{([np])}$ le quantile empirique d'ordre p ([.] représente la partie entière.)

1.1 Convergence des quantiles empiriques

▶ Illustrer le fait que, pour tout $p \in]0,1[$, presque surement

$$X_{([np])} \xrightarrow[n\to\infty]{} k_p.$$

On pourra par exemple l'illustrer pour X de loi $\mathcal{N}(0,1), \mathcal{E}(1)$ et $\mathcal{U}([0,1])$, en prenant un $p \in]0,1[$ quelconque.

▶ Illustrer le fait que, pour tout $p \in]0,1[$ tel que F soit dérivable au point k_p, de dérivé $f(k_p)$,

$$\sqrt{n}(X_{([np])} - k_p) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left(0, \frac{p(1-p)}{f(k_p)^2}\right)$$

On pourra l'illustrer avec l'exemple des lois précédentes.

1.2 Théorème de Glinvenko-Cantelli

D'après le théorème de Glinvenko-Cantelli, si F_n^ω est la fonction de répartition empirique d'un n-échantillon de X, alors presque surement

$$\| F_n^\omega - F \|_\infty \underset{n\to\infty}{\longrightarrow} 0.$$

Dans le cas où F est continue on peut remarquer que

$$\| F_n^\omega - F \|_\infty = \max_{1\leq i\leq n} \left[\max \left(\left| \frac{i}{n} - F(X_{(i)}(\omega)) \right|, \left| \frac{i-1}{n} - F(X_{(i)}(\omega)) \right| \right) \right].$$

► En déduire que la loi de $\| F_n - F \|_\infty$ est indépendante de X dans le cas ou F est continue.

► Illustrer ce fait en simulant des échantillons de cette loi dans le cas où X est de loi $\mathcal{N}(0,1), \mathcal{E}(1)$ et $\mathcal{U}([0,1])$.

1.3 Théorème de Kolmogorov-Smirnov

D'après le théorème de Kolmogorov-Smirnov, dans le cas où F est continue,

$$\sqrt{n} \| F_n - F \|_\infty \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mu_{KS},$$

où μ_{KS} est une loi qui ne dépend pas de X.

► Illustrer ce fait en prenant pour X la loi uniforme sur $[0,1]$.

2 Processus de Poisson

2.1 Définition

Soient $\lambda > 0$ et (S_n) une suite de variables aléatoires indépendante de loi $\mathcal{E}(\lambda)$. On pose pour tout $n \in \mathbb{N}^*$, $T_n = S_1 + \cdots + S_n$, et pour tout $t \in \mathbb{R}^+$,

$$N_t = \sum_{i=1}^{\infty} \mathbb{1}_{\{T_n \leq t\}}.$$

On appel (N_t) un processus de Poisson d'intensité λ. C'est un processus défini sur $\mathbb{R}^+$ à valeurs dans $\mathbb{N}$ dont les trajectoires sont toutes presque surement croissantes, continues à droites et constantes par morceaux. T_n représente l'instant du $n$ème saut.

De plus pour tout $t \in \mathbb{R}^+$, N_t suit une loi $\mathcal{P}(\lambda t)$ et (N_t) est un processus à accroissements indépendants et stationaires. i.e, pour toute subdivision finie $(t_0, \cdots, t_n)$ de $\mathbb{R}^+$, les variables aléatoires $N_{t_{i+1}} - N_{t_i}$, $i \in \{0, \cdots, n-1\}$, sont indépendantes et leur loi ne dépend que des accroissements $(t_{i+1} - t_i)$, $i \in \{0, \cdots, n-1\}$.

2.2 Simulation et propriétés asymptotiques

▶ Simuler et afficher une trajectoire d'un processus de Poisson d'intensité 1 jusqu'au $n$ème saut.

▶ Simuler et afficher une trajectoire d'un processus de Poisson d'intensité 1 jusqu'au temps t.

▶ Simuler et afficher une trajectoire jusqu'au temps t du processus $(N_t - \lambda t)$ où (N_t) est un processus de Poisson d'intensité λ.

▶ Illustrer le fait que presque surement

$$\frac{N_t}{t} \underset{t\to\infty}{\longrightarrow} \lambda.$$

et que

$$\sqrt{\frac{t}{\lambda}}\left(\frac{N_t}{t} - \lambda\right) \xrightarrow[t\to\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

2.3 Propriétés

2.3.1 Somme de processus de Poisson indépendant

▶ Illustrer le fait que la somme de deux processus de Poisson indépendants d'intensités respectives λ et μ est un processus de Poisson d'intensité $\lambda + \mu$.

Par exemple on peut modéliser l'arrivée de clients dans deux files d'attente distinctes par deux processus de Poisson indépendants. Ainsi l'arrivée globale des clients peut être aussi modélisé par un processus de Poisson.

2.3.2 Décomposition d'un processus de Poisson

Soit (N_t) un processus de Poisson d'intensité λ. Soit $p \in [0,1]$. On construit les processus constants par morceaux, continues à droite, nuls en zero, (N_t^1) et (N_t^2), de la manière suivante. A chaque saut de (N_t), indépendament des autres, on choisit de faire sauter d'une unité (N_t^1) avec probabilité p ou (N_t^2) avec probabilité $1-p$. Alors ces deux processus sont des processus de Poisson indépendants d'intensité respective $p\lambda$ et $(1-p)\lambda$.

▶ Simuler et afficher sur le même graphique les processus (N_t), (N_t^1) et (N_t^2).

Dans une file d'attente on peut simuler l'arrivée de clients par un processus de Poisson. Si on peut considérer qu'un client est "sympatique" avec probabilité p et "non sympatique" avec probabilité $1-p$ (penser à p négligeable devant 0,) alors l'arrivée des clients "sympatiques" peut être modélisé par un processus de Poisson.

2.3.3 Processus de Poisson composé

Soit (X_n) une suite de variables aléatoire réelles indépentantes identiquement distribuées de loi ν et (T_n) la suite des temps de saut d'un processus de Poisson (N_t) d'intensité $\lambda > 0$. On pose pour $n \in \mathbb{N}^*$, $U_n = X_1 + \cdots + X_n$.

On définit alors pour tout $t \in \mathbb{R}^+$

$$\begin{aligned} Y_t &= \sum_{i=1}^{\infty} U_n \mathbb{1}_{\{T_n \leq t\}} \\ &= \sum_{i=1}^{N_t} U_i \end{aligned}$$

On dit que (Y_t) est un processus de Poisson composé d'intensité λ et de loi de saut ν.

► Simuler une trajectoire de processus de Poisson d'intensité 1 et de loi de saut $\mathcal{N}(0, 1)$.

Ces processus sont utiles par exemple pour modéliser l'arrivée aléatoire de groupes de personnes à des temps aléatoires.