Informatique : programmation en Python

En particulier :
Simulation de lois usuelles discrètes, à densité.

Chapitre 11 - Convergence de variables aléatoires (tout)

1. Inégalité de Markov et inégalité de Bienaymé-Tchebychev (rappel)
2. Convergence en probabilité
   (a) Définition : On dit que la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ **converge en probabilité vers la variable aléatoire** X lorsque :
   $$\forall \epsilon > 0, \quad \lim_{n\to+\infty} P\Big(|X_n - X| \geqslant \varepsilon\Big) = 0$$
   On note alors $X_n \xrightarrow{P} X$.
   (b) Méthode : on utilise souvent (mais pas toujours) l'inégalité de Bienaymé-Tchebychev.
   (c) Si la suite $(X_n)_{n\in\mathbb{N}}$ converge en probabilité vers X, et si la fonction f est **continue** sur $\mathbb{R}$, alors la suite $(f(X_n))_{n\in\mathbb{N}}$ converge en probabilité vers $f(X)$.
   (d) Si $X_n \xrightarrow{P} X$ et $Y_n \xrightarrow{P} Y$ alors $X_n + Y_n \xrightarrow{P} X + Y$.
   (e) **Loi faible des grands nombres** (*)
   Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles.
   Notons
   $$\forall n \in \mathbb{N}^*, \quad \overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{X_1 + \cdots + X_n}{n}$$
   Si :
   - les variables X_n sont **indépendantes**,
   - elles admettent toutes la **même espérance** notée m et la **même variance** notée σ^2,

   alors la suite $(\overline{X_n})_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable aléatoire certaine égale à m : $X_n \xrightarrow{P} m$
3. Convergence en loi
   (a) Définition **via la fonction de répartition**. Cas particulier des VARD à valeurs dans $\mathbb{Z}$.
   (b) Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de VARD de loi $\mathcal{B}\left(n, \frac{\lambda}{n}\right)$. Alors $X_n \xrightarrow{\mathcal{L}} Y$ où $Y \hookrightarrow \mathcal{P}(\lambda)$.
   Preuve guidée à savoir refaire :
   i. Montrer que $\forall k \in \mathbb{N}$, $\displaystyle\binom{n}{k} \underset{n\to+\infty}{\sim} \frac{n^k}{k!}$.
   ii. En déduire que $\forall k \in \mathbb{N}$, $\lim_{n\to+\infty} P(X_n = k) = \dfrac{\lambda^k}{k!}e^{-\lambda}$.
   (c) Si la suite $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X, et si la fonction f est **continue** sur $\mathbb{R}$, alors la suite $(f(X_n))_{n\in\mathbb{N}}$ converge en loi vers $f(X)$.

4. Théorème Central Limite
   (a) Le T.C.L
   Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires.
   Soit m un réel et σ un réel strictement positif.
   On suppose que :
   - les variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ sont indépendantes,
   - les variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ sont identiquement distribuées (elles suivent la même loi),
   - toutes ces variables aléatoires admettent une même espérance m et une même variance notée σ^2.

   On note : $\forall n \in \mathbb{N}^*$, $\overline{X_n} = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$. La variable aléatoire centrée réduite associée à $\overline{X_n}$ est :
   $$\overline{X}_n^* = \frac{\overline{X}_n - E(\overline{X}_n)}{\sqrt{V\left(\overline{X}_n\right)}} = \sqrt{n}\left(\frac{\overline{X}_n - m}{\sigma}\right)$$
   Alors

   la suite $(\overline{X_n}^{\,*})_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire qui suit une loi normale centrée réduite.
   $$\overline{X_n}^{\,*} \xrightarrow{\mathcal{L}} Y \quad \text{où } Y \hookrightarrow \mathcal{N}(0,1)$$
   (b) Convergence de la loi binômiale vers la loi normale : à savoir retrouver en étant un peu guidé.
   (c) Convergence de la loi de Poisson vers la loi normale : à savoir retrouver en étant un peu guidé.

Chapitre 12 - Estimation (tout)

1. **Vocabulaire :** notion d'échantillon i.i.d. (indépendant, identiquement distribué) $(X_1, \cdots, X_n)$ de même loi qu'une variable X donné.

   On considère une variable X dont la loi dépend d'un réel θ et $g : \mathbb{R} \to \mathbb{R}$.
2. Définition d'un estimateur $T_n = \varphi(X_1, \cdots, X_n)$ de $g(\theta)$.
3. Un estimateur T_n de $g(\theta)$ est **sans biais** si $E(T_n) = g(\theta)$.
4. Un estimateur T_n de $g(\theta)$ est **convergent** si $T_n \xrightarrow{P} g(\theta)$.
5. Si $T_n \xrightarrow{P} g(\theta)$ et f est une fonction continue alors $f(T_n) \xrightarrow{P} f(g(\theta))$.
6. La moyenne empirique $\overline{X_n} = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$ est un estimateur sans biais et convergent de $m = E(X)$.
7. **condition suffisante de convergence :** si $\lim_{n\to+\infty} E(T_n) = g(\theta)$ et si $\lim_{n\to+\infty} V(T_n) = 0$ alors $T_n \xrightarrow{P} g(\theta)$ (T_n est un estimateur convergent de $g(\theta)$.
8. comparaison de deux estimateurs : si l'on dispose de plusieurs estimateurs de $g(\theta)$, le plus intéressant est celui pour lequel la quantité $V(T_n) + (E(T_n) - g(\theta))^2$ est la plus faible.

9. **Intervalle de confiance**

Définitions : **intervalle de confiance, intervalle de confiance asymptotique**.
Nous avons vu différents exemples.
Retenir pour la construction d'un intervalle de confiance pour la **loi de Bernoulli** :
pour tout $p \in [0,1]$, $p(1-p) \leq \frac{1}{4}$ (via une étude de variations rapide).
Ingrédients pour la construction d'un intervalle de confiance : Markov, IBT, lois usuelles (dont loi normale !!), TCL...

10. **Méthode de Monte Carlo :** méthode pour estimer l'espérance $m = E(Y)$ d'une variable aléatoire Y ou une certaine probabilité.
Application à différents calculs approchés d'intégrales, de séries, de sommes finies.
Le plus souvent $Y = g(X)$ où X suit une loi usuelle discrète (estimation de somme ou de série) ou à densité (estimation d'une intégrale) que l'on peut simuler sur Python et g est une fonction bien choisie.
On estime alors l'espérance m par la moyenne empirique $\overline{Y_n} = \frac{1}{n}\sum_{k=1}^{n} Y_k$.
Exercice de cours à savoir refaire
Soit $M = \displaystyle\int_0^1 \frac{4}{1+t^2}dt$.

    (a) Ecrire M comme l'espérance d'une variable aléatoire Y.

    (b) Ecrire un programme Python, qui définit la fonction g où $g(t) = \frac{4}{1+t^2}$ puis qui affiche une estimation de M par la méthode de Monte-Carlo pour un échantillon de taille n.

    (c) Calculer la valeur exacte de M. Objectif du programme précédent ?

. (*) : **preuve exigible**