TD16

Convergence de variables aléatoires

Inégalités de concentration

Exercice 16.1 (★★)

En appliquant l'inégalité de Markov à $|X|$, où X est une variable aléatoire suivant la loi normale centrée réduite, montrer que pour tout réel x strictement positif, on a :

$$\int_0^x e^{-\frac{t^2}{2}}\mathrm{d}t \geq \sqrt{\frac{\pi}{2}} - \frac{1}{x}.$$

Exercice 16.2 (★★)

On note Φ la fonction de répartition d'une variable aléatoire suivant une loi $\mathscr{N}(0,1)$.

1. Montrer à l'aide de l'inégalité de Bienaymé-Tchebychev que : $\forall x \in \mathbb{R}_+^*,\ 1 - \Phi(x) \leq \dfrac{1}{2x^2}$.
2. En déduire que $\displaystyle\int_0^{+\infty} (1-\Phi(t))\mathrm{d}t$ converge et calculer sa valeur à l'aide d'une intégration par parties.

Convergence en probabilité

Exercice 16.3 (★)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires, définies sur un même espace probabilisé, indépendantes, et qui suivent toutes la loi de Poisson de paramètre $\lambda > 0$.

Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i$. Étudier la convergence en probabilité de la suite $(S_n)_{n\in\mathbb{N}^*}$.

Exercice 16.4 (★)

Soit $(\varepsilon_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires et X une variable aléatoire, toutes définies sur le même espace probabilisé. On suppose que pour tout $n \in \mathbb{N}^*$, $\varepsilon_n \hookrightarrow \mathscr{U}(\{1-\frac{1}{n}, 1+\frac{1}{n}\})$, et que X admet un moment d'ordre 1.

Pour tout $n \in \mathbb{N}^*$, on pose $X_n = \varepsilon_n X$. Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers X.

Exercice 16.5 (★★)

1. Soit (X_n) une suite de variables aléatoires à valeurs positives ayant toutes une espérance et une variance. On suppose que $\displaystyle\lim_{n\to+\infty} E(X_n) = m$ et $\displaystyle\lim_{n\to+\infty} V(X_n) = 0$.

   En appliquant l'inégalité de Markov à $(X_n - m)^2$, montrer que $X_n \xrightarrow{P} m$.
2. Soit $(Y_k)_{k\geq 1}$ une suite de variables aléatoires i.i.d. suivant toutes une loi exponentielle de paramètre $\lambda > 1$. Pour tout $n \geq 1$, on pose $X_n = \displaystyle\prod_{k=1}^n Y_k$. À l'aide de l'inégalité de Markov, montrer la convergence en probabilité de la suite (X_n) vers la variable aléatoire certaine égale à 0.

Exercice 16.6 (★★)

On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes suivant toutes la loi de Bernoulli de paramètre p. Pour tout $n \in \mathbb{N}^*$, on pose $Y_n = X_n + X_{n+1}$ et $\overline{Y_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} Y_k$.

1. Soit $n < n'$. Donner une condition nécessaire et suffisante sur n et n' pour que Y_n et $Y_{n'}$ soient indépendantes.
2. Montrer que la suite $\overline{Y_n}$ converge en probabilité vers la variable certaine égale à $2p$.

Exercice 16.7 (★★★ - ✍)

Soit $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$ deux suites de variables aléatoires et X et Y deux variables aléatoires, toutes définies sur un même espace probabilisé. On suppose que $X_n \xrightarrow{P} X$ et $Y_n \xrightarrow{P} Y$.

1. Soit $\varepsilon > 0$. Montrer que $[|X_n + Y_n - (X+Y)| \geq \varepsilon] \subset [|X_n - X| \geq \frac{\varepsilon}{2}] \cup [|Y_n - Y| \geq \frac{\varepsilon}{2}]$.
2. Montrer que $X_n + Y_n \xrightarrow{P} X + Y$.
3. On suppose de plus que toutes les variables aléatoires considérées sont à valeurs strictement positives. Montrer que $X_n \cdot Y_n \xrightarrow{P} X \cdot Y$.

Exercice 16.8 (★★★ - QSP HEC 2016)

Soit $X_1, \ldots, X_n$ des variables i.i.d. selon la loi uniforme $\mathscr{U}([0,1])$. Pour tout $n \in \mathbb{N}^*$, on pose :

$$Y_n = \left(\prod_{i=1}^{n} X_i\right)^{1/n}.$$

1. Expliquer pourquoi la figure ci-dessous, obtenue à partir du code `Python` suivant, suggère la convergence en probabilité de la suite $(Y_n)_{n\in\mathbb{N}}$.

```python
x = np.linspace(0.2,0.8,20)
n = np.array([50,200,800])
for i in range(3):
    N = 1000
    LY = np.sum(np.log(rd.random([n[i],N])),0) # sommation des colonnes
    C = np.exp(LY/n[i])
    plt.subplot(1,3,i+1) #place trois graphiques côte à côte
    plt.hist(C,x,density='True',edgecolor = 'k')
plt.show()
```

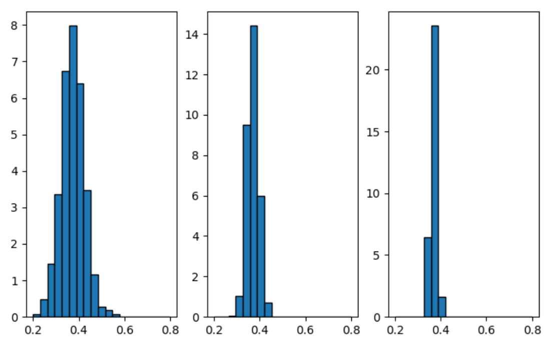

Figure 1 - *Histogrammes.*

2. Justifier la convergence en probabilité de la suite $(Y_n)_{n\in\mathbb{N}}$.

Convergence en loi

Exercice 16.9 (★)

Soit (X_n) une suite de variables aléatoires discrètes telles que $X_n \hookrightarrow \mathscr{U}([\![1, n]\!])$. Étudier la convergence en loi de la suite (X_n).

Exercice 16.10 (★★)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires telle que pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathscr{P}\left(\frac{1}{n}\right)$.
Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire X dont on précisera la loi.
Montrer qu'il s'agit également d'une convergence en probabilité.

Exercice 16.11 (★★)

Une urne contient des boules numérotées de 1 à n ($n \geq 2$). On tire avec remise une boule de l'urne jusqu'à obtenir un numéro supérieur ou égal au premier numéro tiré. Soit X_n la variable égale au nombre de tirages effectués.

1. Montrer que pour tout $(n,k) \in (\mathbb{N} \setminus \{0,1\})^2$, on a $P(X_n = k) = \dfrac{1}{n}\displaystyle\sum_{j=0}^{n-1}\left(\left(\frac{j}{n}\right)^{k-2} - \left(\frac{j}{n}\right)^{k-1}\right)$.
2. À l'aide des sommes de Riemann, déterminer $\displaystyle\lim_{n\to+\infty} P(X_n = k)$ pour tout $k \in \mathbb{N}$, $k \geq 2$.
3. En déduire que $(X_n)_{n\geq 2}$ converge en loi vers une variable aléatoire dont on précisera la loi.

Exercice 16.12 (★★★ - QSP HEC 2017)

Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires réelles, indépendantes, définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$ et soit $(p_i)_{i\in\mathbb{N}^*}$ une suite de réels de $]0,1[$ telle que pour tout $i \in \mathbb{N}^*$, la variable aléatoire X_i suit une loi géométrique de paramètre p_i. On pose pour tout $i \in \mathbb{N}^*$: $q_i = 1 - p_i$.
On pose pour tout $n \in \mathbb{N}^*$: $Z_n = \min(X_1, X_2, \ldots, X_n)$.

1. Calculer pour tout $k \in \mathbb{N}^*$, la probabilité $P(Z_n \geq k)$. Quelle est la loi de Z_n ?
2. On suppose que pour tout $i \in \mathbb{N}^*$, on a $p_i = \dfrac{1}{(i+1)^2}$.

   Montrer que la suite de variables aléatoires $(Z_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable aléatoire dont on précisera la loi.

Exercice 16.13 (★)

Soit (X_n) une suite de variables aléatoires telles que $X_n \hookrightarrow \mathscr{N}(m, \frac{\sigma^2}{n})$.

1. Déterminer la fonction de répartition de X_n.
2. Étudier la convergence en loi de la suite (X_n).
3. Étudier la convergence en probabilité de la suite (X_n).

Exercice 16.14 (★★)

Pour $n \in \mathbb{N}^*$, soit f_n la fonction définie sur $\mathbb{R}$ par : $\forall t \in \mathbb{R}, \quad f_n(t) = \dfrac{1}{\pi}\dfrac{n}{1+n^2t^2}$.

1. Montrer que f_n est une densité de probabilité.
2. Soit (X_n) une suite de variables aléatoires telles que X_n admet f_n pour densité.
   (a) Déterminer la fonction de répartition de X_n.
   (b) Montrer que (X_n) converge en loi vers la variable certaine égale à 0.
   (c) Montrer qu'il s'agit également d'une convergence en probabilité.

Exercice 16.15 (★★)

Soient (X_n) une suite de variables aléatoires indépendantes suivant la loi uniforme sur $[0,1]$. On pose pour tout $n \geq 1$, $Y_n = \min(X_1, \cdots, X_n)$.

1. Déterminer la fonction de répartition F_n de Y_n.
2. Montrer que (Y_n) converge en probabilité vers la variable constante égale à 0.
3. Montrer que la suite (nY_n) converge en loi vers une variable suivant une loi exponentielle.

Exercice 16.16 (★★)

Considérons la fonction f définie sur $\mathbb{R}$ par : $\quad \forall x \in \mathbb{R},\ f(x) = \dfrac{e^x}{(1+e^x)^2}$.

1. Montrer que f est une densité de probabilité.
2. Soit à présent $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires i.i.d. de densité f.
   (a) Déterminer la fonction de répartition des X_i.
   (b) En déduire la fonction de répartition de $M_n = \max(X_1, \ldots, X_n)$.
   (c) Étudier la convergence en loi de la suite $(M_n - \ln(n))_{n\in\mathbb{N}}$.

Exercice 16.17 (★★★)

Soit (X_n) une suite de variables aléatoires discrètes telles que $n \times X_n$ suit une loi uniforme sur $[\![0, n]\!]$.

1. Rappelons que si `x` et `y` sont des vecteurs de même dimension, la commande `plt.step(x,y)` trace la fonction en escalier qui vaut `y[k]` sur l'intervalle $[\mathtt{x}[\mathtt{k}-1], \mathtt{x}[\mathtt{k}][$.

   On considère le code `Python` suivant ainsi que la figure qu'il renvoie :

```python
n = np.array([4,10,30])
for i in range(3):
    x = np.linspace(0,1,n[i]+1)
    y = np.arange(0,n[i]+1)/(n[i]+1)
    plt.subplot(1,3,i+1)
    plt.step(x,y)
plt.show()
```

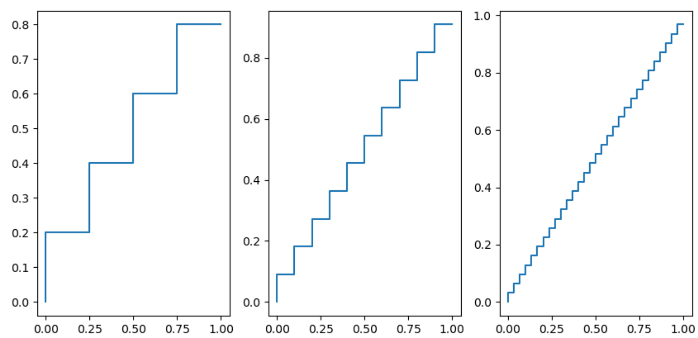

Figure 1 - *Fonctions en escaliers.*

   Que peut-on conjecturer ?

2. Étudier la convergence en loi de (X_n).

Exercice 16.18 (★★)

On considère une suite (X_n) de variables aléatoires indépendantes suivant la loi de Poisson de paramètre 1, et on pose $S_n = X_1 + \cdots + X_n$.

1. Exprimer $P(S_n \leq n)$ sous forme d'une somme.
2. En utilisant le théorème central limite, montrer que : $\quad \displaystyle\sum_{k=0}^{n} \frac{n^k}{k!} \underset{n\to+\infty}{\sim} \frac{e^n}{2}$.

Exercice 16.19 (★★★★ - Convergence en probabilité et convergence en loi (Oral HEC) - ✍)
Soient $(X_n)_{n\geq 1}$ et X des variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ telles que $X_n \xrightarrow{P} X$. On note F_n la fonction de répartition de X_n et F la fonction de répartition de X.
Soit x un point de continuité de F, et soit $\delta > 0$ fixé.

1. Montrer qu'on peut choisir $\varepsilon > 0$ tel que $F(x-\varepsilon) > F(x) - \delta$ et $F(x+\varepsilon) < F(x) + \delta$.
2. Montrer que :
$$[X_n \leq x] \subset [X \leq x+\varepsilon] \cup [|X_n - X| > \varepsilon].$$
Montrer de même que :
$$[X \leq x-\varepsilon] \subset [X_n \leq x] \cup [|X_n - X| > \varepsilon].$$
3. En déduire que :
$$F_n(x) \leq F(x) + \delta + P(|X_n - X| > \varepsilon) \quad \text{et} \quad F(x) - \delta \leq F_n(x) + P(|X_n - X| > \varepsilon).$$
4. Montrer que $X_n \xrightarrow{\mathscr{L}} X$.

Approximation

Exercice 16.20 (★)
Soit S une variable aléatoire de loi $\mathscr{B}(45\,;0.7)$. Donner une approximation de la probabilité $P(28 < S \leq 39)$ sans correction de continuité puis avec correction de continuité.
Un calcul détaillé à l'aide de loi binomiale donne $P(28 < S \leq 39) \approx 0,8332$.

Exercice 16.21 (★★)
On effectue des lancers successifs d'un dé équilibré. On cherche à déterminer le nombre n de lancers nécessaires pour garantir avec moins de 5% d'erreur que la fréquence d'apparition du 1 sera $\dfrac{1}{6} \pm 0.01$.

1. **Méthode 1.** À l'aide de l'inégalité de Bienaymé Tchebychev.
   (a) Soit (Y_n) une suite de variables aléatoires mutuellement indépendantes, de même loi et admettant un moment d'ordre 2. On pose $S_n = \displaystyle\sum_{k=1}^{n} Y_k$. Prouver que :
   $$\forall a > 0, \quad P\left(\left|\frac{S_n}{n} - E(Y_1)\right| \geq a\right) \leq \frac{V(Y_1)}{na^2}.$$
   (b) En déduire le nombre n de lancers nécessaires.
2. **Méthode 2.** À l'aide du théorème central limite, déterminer approximativement ce nombre n.
3. Un calcul détaillé à l'aide de la loi binomiale prouve qu'il faut $n \geq 5395$. Que dire des résultats obtenus précédemment ?

Exercice 16.22 (★★)
Dans le désert, une Renault 4L crève en moyenne tous les 4000 km. On considère donc qu'à chaque kilomètre, la probabilité de crever est de $\frac{1}{4000}$. Un équipage s'inscrit au 4L Trophy, rallye de 6000 km dans le désert. Il aimerait savoir combien de roues de secours emporter pour avoir moins de 10% de chances de manquer de roues de secours. Deux roues de secours sont-elles suffisantes ? Trois ? On donne $e^{-3/2} \approx 0.22$.

Exercice 16.23 (★★)

Une entreprise compte 300 employés. Chacun d'eux téléphone en moyenne 6 minutes par heure. Combien de lignes téléphonique doit faire installer l'entreprise afin que la probabilité que toutes les lignes soient occupées en même temps soit inférieure ou égale à 0.025 ?

Exercice 16.24 (★★★★ - QSP HEC 2018)

On considère la fonction `Python` suivante :

```python
def bn(N,n,p):
    X = rd.geometric(p,[N,n])
    k = 0
    for i in range(N):
        Y = np.sum(X[i,:])
        if Y >=n/p :
            k = k+1
    f = k/N
    return f
```

1. De quel nombre réel `bn(N,1,0.4)` fournit-il une valeur approchée lorsque N est grand ?

2. Donner une valeur approximative de `bn(10000,10000,0.4)`.
