

## Probabilités TP 3

### Illustration numérique des théorèmes de convergence

Rappel : taper `krdc vnc://nom_du_serveur` dans un terminal.

Les sujets et les corrigés des TP sont mis le lendemain des séances sur ma page :

<http://www.math.u-bordeaux1.fr/~chabanol/stat.html>

#### 1. MISE EN ÉVIDENCE D'UNE CONVERGENCE PRESQUE SÛRE

**1.1. Illustration de la loi des grands nombres.** Rappel : La loi des grands nombres est le théorème qui dit que si  $(X_n)_{n \in \mathbb{N}}$  est une suite de variables aléatoires indépendantes intégrables, de même loi, d'espérance  $m$ , alors pour presque tout  $\omega$ ,  $\frac{\sum_{i=1}^n X_i(\omega)}{n}$  converge vers  $m$  : la moyenne arithmétique converge vers l'espérance. La loi forte des grands nombres entraîne également la convergence presque sûre des fonctions de répartition empiriques d'un échantillon de variables aléatoires indépendantes de même loi vers la fonction de répartition théorique.

On va par exemple travailler avec une variable aléatoire  $X$  de loi uniforme sur  $\{1, 2, 3, 4, 5, 6\}$  (on va simuler un lancer de dés). On va donc se servir de notre fonction `runifd`.

- (1) *Quelle est l'espérance  $m$  de  $X$  ? Quel est son écart-type  $\sigma$  ?*
- (2) *Générer un échantillon de  $X$  de taille 5000; utiliser `cumsum` pour obtenir un tableau avec les valeurs successives des  $\sum_{i=1}^n X_i$  pour  $n$  allant de 1 à 5000, puis les valeurs successives des  $\frac{\sum_{i=1}^n X_i}{n}$  pour  $n$  allant de 1 à 5000.*
- (3) *Représenter sur un même graphe les  $\frac{\sum_{i=1}^n X_i}{n}$  en fonction de  $n$ , ainsi qu'une droite horizontale correspondant à l'espérance : on peut constater la convergence.*
- (4) *Constater en traçant les courbes que les  $\sqrt{n}(\frac{\sum_{i=1}^n X_i}{n} - m)$  restent d'ordre 1 (mais ne convergent pas p.s.), alors que les  $n(\frac{\sum_{i=1}^n X_i}{n} - m)$  divergent et que les  $n^{0.1}(\frac{\sum_{i=1}^n X_i}{n} - m)$  tendent vers 0 p.s. : on voit le rôle du  $\frac{1}{\sqrt{n}}$  présent dans le théorème central limite.*
- (5) *Question subsidiaire. Représenter les fonctions de répartition empiriques de l'échantillon des  $X_i$  dont on a gardé respectivement que les 100, 500, 1000 et 5000 premiers termes, ainsi que la fonction de répartition théorique (vous pouvez utiliser votre fonction `plotrep`). La convergence de ces fonctions de répartition empirique est une conséquence de la loi forte des grands nombres.*
- (6) *Question subsidiaire. Refaire les questions 2 et 5 avec une variable aléatoire qui suit une loi uniforme sur  $[0, 1]$ .*

**1.2. Convergence presque sûre mais pas vers une constante.** La loi forte des grands nombres garantit la convergence presque sûre vers une constante (l'espérance). Mais on peut aussi dans d'autres situations converger vers une variable aléatoire non constante.

- (7) *Générer un 100-échantillon  $E$  de loi de Bernoulli de paramètre 0.5.*
- (8) *Obtenir un vecteur contenant les  $\sum_{i=1}^n \frac{E_i}{2^i}$ , pour  $n$  allant de 1 à 100.*
- (9) *La suite  $u_n = \sum_{i=1}^n \frac{E_i}{2^i}$  vous semble-t-elle converger ? Vers quelle valeur ? (faites une figure avec matlab)*
- (10) *Refaites l'expérience avec un autre échantillon : on observe encore une convergence (il y a encore convergence p.s.), mais la limite est différente : elle est aléatoire.*

- (11) Pour observer la loi de cette valeur limite, on va en fait illustrer la convergence en loi. Pour cela, on a besoin de beaucoup de tels échantillons de  $E$  (par exemple 2000), pour lesquels on va retenir uniquement la valeur "limite"  $\sum_{i=1}^{100} \frac{E_i}{2^i}$ , et on va observer comment sont réparties ces valeurs limites.

*Mettre dans un vecteur les 2000 valeurs successives obtenues pour  $\sum_{i=1}^{100} \frac{E_i}{2^i}$  lorsqu'on effectue 2000 fois les questions 7 et 8. Tracer la fonction de répartition empirique de ces valeurs. A votre avis, quelle est la loi limite ?*

- (12) *Question subsidiaire, que vous pouvez regarder chez vous, qu'on fera peut-être en TD. Prouver les convergences que vous avez observées.*
- (13) *Question subsidiaire. Mêmes questions si  $E$  ne suit pas une loi de Bernoulli mais une loi uniforme sur  $\{-1, 1\}$ .*

Pour les curieux : la série  $\sum_n \frac{A_n}{3^n}$  où les  $A_n$  sont i.i.d de loi uniforme sur  $\{0, 2\}$  converge en loi vers une loi uniforme sur l'ensemble de Cantor.

## 2. THÉORÈME CENTRAL LIMITE

Le théorème central limite affirme que si  $(X_n)$  est une suite de variables aléatoires indépendantes de même loi et d'espérance  $\mu$  et de variance  $\sigma^2$  finies,  $\sqrt{n}(\bar{x} - \mu)/\sigma$  converge en loi vers une loi normale centrée réduite (on a noté  $\bar{x} = \frac{\sum_{i=1}^n X_i}{n}$ ). Cette quantité reste donc aléatoire, mais on sait comment elle est distribuée asymptotiquement.

- (14) On reprend le lancer de dé de tout à l'heure, et on va effectuer plusieurs fois ( $N_{obs}$  fois) l'expérience (virtuelle) de le lancer  $n$  fois.

$X$  suit donc comme à la question 1 une loi uniforme sur  $\{1, 2, 3, 4, 5, 6\}$ , d'espérance  $m$  et d'écart-type  $\sigma$ .

*Construire une fonction **TCLde** qui prend en entrée deux entiers  $n$  et  $N_{obs}$ , génère successivement  $N_{obs}$   $n$ -échantillons de  $X$ , calcule pour chacun de ces échantillons la valeur "finale"  $\frac{\sqrt{n}}{\sigma}(\bar{x} - m)$ , stocke ces valeurs dans un vecteur, et enfin trace sur un même graphique la fonction de répartition empirique de ces grandeurs ainsi que le graphe de la fonction de répartition empirique de la loi normale centrée réduite.*

*Utiliser cette fonction avec  $N_{obs} = 1000$  et  $n = 10, 50, 300, 1000$ .*

- (15) *Question subsidiaire.* Un autre théorème de convergence en loi dit que si  $\lambda$  est un réel positif, une suite de variables aléatoires de loi binomiale de paramètre  $(n, \frac{\lambda}{n})$  converge en loi vers un Poisson de paramètre  $\lambda$ .

(a) *Soit  $Z$  une variable suivant une loi de Poisson de paramètre 2. Créer un vecteur contenant les  $P(Z = k)$  pour  $k$  variant de 0 à 8 (l'instruction pour la factorielle est **factorial**). Tracer la fonction de répartition d'une loi de Poisson de paramètre 2 (utiliser ce vecteur, et l'instruction **cumsum**).*

(b) *Générer un 5000 échantillon de loi binomiale de paramètre  $(10, \frac{2}{10})$  (on prend donc  $n = 10$ ). Tracer la fonction de répartition empirique et comparer avec la fonction de répartition d'une loi de Poisson de paramètre 2.*

(c) *Faire de même avec  $n = 100$  et  $n = 1000$  et observer la convergence.*

- (16) On peut aussi observer la convergence en loi de chaînes de Markov. La fonction ci-dessous génère un pas d'une chaîne de Markov sur  $\{1, 2, 3\}$ .

```

function Y = Markov( X )
U=rand(1);
if (X==1) Y=2*(U<1/3) + 3*(U>1/3); end
if (X==2) Y=1*(U<1/3) + 3*(U>1/3); end
if (X==3) Y=1*(U<2/3) + 2*(U>2/3); end
end

```

*Écrire la matrice de transition  $A$  de la chaîne. Vérifier (à l'aide de matlab) que  $[0.35 \ 0.25 \ 0.4]$  est une probabilité invariante. Est-ce la seule ?*

*Prendre un vecteur  $\pi_0$  de votre choix pour la loi initiale (la somme des probabilités doit quand même faire 1 !) et utiliser matlab pour calculer  $\pi_n$  pour  $n = 5, 10, 50, 100$  et observer la convergence en loi.*

*Autre mise en évidence par simulation de la chaîne :*

*Générer un 5000 échantillon de la chaîne à l'instant 1000 (en partant de l'état initial que vous voulez), et mettre en évidence la convergence en loi vers la loi invariante.*