TP Probabilités-Statistiques 2.
Illustration numérique de convergences en loi

Rappel : taper `krdc vnc://nom_du_serveur` dans un terminal.
Les sujets et les corrigés des TPs sont mis après les séances sur ma page :
http://www.math.u-bordeaux1.fr/~chabanol/TP_ agreg.html
Préalable technique pénible :
Créer dans votre répertoire une fonction "pbinom" ainsi qu'une fonction "pnorm" qui fournissent la fonction de répartition de la binomiale de paramètres n, p, et de la loi normale centrée réduite.

```
function y=pbinom(k,n,p)
m=length(k);
y=cdfbin("PQ",k,n*ones(1,m),p*ones(1,m),(1-p)*ones(1,m));
endfunction
```

```
function y=pnorm(x)
y=(1+erf(x/sqrt(2)))/2;
endfunction
```

Pour illustrer une convergence en loi $X_n \to X$, on doit illustrer le fait que X_n et X ont "presque" la même loi. Pour cela il peut y avoir plusieurs solutions :

- Si X_n et X sont discrètes et qu'on connaît explicitement les lois de X_n et de X, on peut comparer les diagrammes en bâton de la loi de X_n pour n "grand" et celui de la loi de X (utiliser `plot2d3`)
- Si on connaît explicitement les fonctions de répartition de X_n et X, on peut prendre n "grand" et tracer les deux fonctions (si X_n est discrète, c'est une fonction en escalier).
- **Ce qu'on fait très souvent** : si on sait simuler X_n, on peut aussi utiliser le théorème de Glivenko-Cantelli : on simule un assez gros échantillon de X_n avec n assez grand (fixé) et on trace sa fonction de répartition empirique, qu'on compare avec la fonction de répartition de X.
  Il y a alors deux sources de différences dans les deux tracés, puisqu'il y a deux convergences : le fait que n est fini, et que la loi de X_n est donc a priori différente de celle de X, et le fait que l'échantillon est fini, et qu'on a donc qu'une approximation de la loi de X.
  On peut aussi éventuellement passer par un histogramme.
- Enfin si on ne connaît pas la loi limite, on peut prendre 2 valeurs n_1 et n_2 grandes et illustrer le fait que X_{n_1} et X_{n_2} ont "à peu près" la même loi (par exemple à l'aide de fonctions de répartition empiriques).

1. Approximations de la loi binomiale

1.1. **Approximation d'une loi binomiale par une loi de Poisson.** Lorsqu'on a une binomiale, pour de grandes valeurs de n, le calcul de $\binom{n}{k} p^k q^{n-k}$ devient vite pratiquement impossible, sauf si l'on cherche à calculer le logarithme de cette expression au lieu de l'expression elle-même (et à condition d'utiliser l'approximation des factorielles par la formule de Stirling). Il y a en pratique deux cas limites . Le premier cas est le suivant.
Lorsque n tend vers l'infini et que p tend vers 0 avec $p \sim \frac{\lambda}{n}$, la loi binomiale converge vers une loi de Poisson de paramètre λ. En pratique, on remplace la loi binomiale par une loi de Poisson dès que $n > 30$ et $np < 5$ ou dès que $n > 50$ et $p < 0.1$.

1. *Comment démontre-t-on cette convergence en loi ?*

2. Illustrer cette convergence. En particulier illustrer les valeurs utilisées "en pratique" : examiner ce qu'il se passe si n est trop petit, ou si p est trop grand.

1.2. **Approximation d'une loi binomiale par une loi normale.** Le deuxième cas limite avec les binomiales est donné par le théorème de Moivre-Laplace : si $0 < p < 1$ est un réel fixé et si (S_n) est une suite de v.a. telle que S_n suit une loi binomiale de paramètres (n,p), $Z_n = \frac{S_n - np}{\sqrt{np(1-p)}}$ converge en loi vers une loi normale centrée réduite.
En pratique pour calculer $P(a \leq Z_n \leq b)$ on remplace Z_n par une v.a. normale centrée réduite dès que $n > 30$, $np > 5$ et $n(1-p) > 5$.
3. Illustrer cette convergence en loi, par exemple en simulant un échantillon d'une loi binomiale. En particulier illustrer les valeurs utilisées "en pratique" : examiner ce qu'il se passe si n est trop petit, ou si p est trop grand ou trop petit.

2. D'autres convergences en loi

(1) Si (X_n) sont une suite de v.a. indépendantes identiquement distribuées de loi exponentielle de paramètre α, $Y_n = \alpha \max_{1\leq i\leq n} X_i - \ln(n)$ converge en loi vers une loi de fonction de répartition $e^{-e^{-x}}$.
4. Démontrer ce résultat.
5. Illustrer numériquement ce résultat, en simulant un échantillon de Y_n

(2) Si (X_n) sont une suite de v.a. indépendantes identiquement distribuées de loi uniforme sur $[0,1]$, $T_n = n \min_{1\leq i\leq n} X_i$ converge en loi vers une loi exponentielle de paramètre 1.
6. Démontrer ce résultat
7. Illustrer numériquement ce résultat, en simulant un échantillon de T_n

3. Le théorème central limite

D'après le théorème central limite, si (X_n) est une suite de variables aléatoires indépendantes de même loi et d'espérance μ et de variance σ^2 finies, $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ converge en loi vers une loi normale centrée réduite (on a noté $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$) . Cette quantité reste donc aléatoire, mais on sait comment elle est distribuée asymptotiquement.
8. Illustrer le théorème central limite dans le cas où les X_i suivent une loi uniforme sur $[0,1]$.
9. Illustrer le théorème central limite dans le cas où chaque X_i sont le résultat d'un lancer de dé et suit donc une uniforme sur $\{1,2,3,4,5,6\}$.
10. (Question subsidiaire). Dans le cas où les X_i suivent une loi de Pareto de paramètre $\alpha \in]1,2[$, les X_i sont intégrables d'espérance $\frac{\alpha-1}{\alpha}$ mais pas L^2 (le vérifier).
Illustrer le fait que dans ce cas il n'y a pas convergence en loi de $\sqrt{n}(\bar{X}_n - E[X_1])$.

4. Convergence en loi d'une chaîne de Markov

On peut aussi observer la convergence en loi de bien d'autres processus. Voici ci-dessous un exemple de chaîne de Markov.
Soient d balles ($d > 1$) numérotées de 1 à d et réparties dans deux urnes A et B. A chaque instant, on tire un nombre i uniformément entre 1 et d, et avec probabilité $\frac{d}{d+1}$ la balle numéro i change d'urne. Avec probabilité $\frac{1}{d}$ rien ne se passe.
Soit X_n le nombre de balles dans l'urne A après n tirages indépendants. La fonction ci-dessous prend le nombre de balles X_{n-1} dans A avant l'étape n et simule le nombre de balles X_n dans A avant l'étape $n+1$.

```
 function Y = Ehrenfest(d, X )
U=rand(1);
if (U<d/(d+1))
V=rand(1);
Y=X +(V>(X/d))-(V<(X/d));
else Y=X;
end
endfunction
```

11. Comprendre pourquoi la fonction ci-dessus génère bien ce qu'on veut.
Pour $d = 5$, générer un 500 échantillon de la chaîne à l'instant 1000 et à l'instant 1100 et mettre en évidence une convergence en loi.
La théorie des chaînes de Markov prédit une convergence en loi vers la loi $Bin(d, \frac{1}{2})$. Est-ce cohérent avec vos observations ?