# Feuille d'exercices n°14 : Convergence approximation

## Inégalité de Markov

**Exercice 1.** ($\bigstar$)
On se propose de démontrer l'inégalité de Markov dans le cas particulier où la v.a.r. considérée $T$ est une v.a.r. à densité.
On suppose que $T$ est à valeur dans $\mathbb{R}_+$ et que $T$ admet une espérance.

*1.* Soit $a > 0$. Démontrer que : $\mathbb{E}(T) = \displaystyle\int_0^a t\, f(t)\, dt + \int_a^{+\infty} t\, f(t)\, dt$.

*2.* En déduire que : $\mathbb{E}(T) \geqslant a \displaystyle\int_a^{+\infty} f(t)\, dt$.

*3.* Conclure.

## Inégalité de Bienaymé-Tchebychev

**Exercice 2.** ($\bigstar$) *(extrait de* `EDHEC 2007`*)*
Soit $X$ une variable discrète à valeurs dans $\mathbb{N}^*$, de loi définie par :

$$\forall i \in \mathbb{N}^*,\ \mathbb{P}([X = i]) = \sum_{k=i}^{+\infty} \frac{1}{k} \left(\frac{1}{2}\right)^k$$

On admet que $X$ a une espérance et une variance respectivement égale à $\mathbb{E}(X) = \dfrac{3}{2}$ et $\mathbb{V}(X) = \dfrac{11}{12}$.

*1.* Écrire l'inégalité de Bienaymé-Tchebychev, pour la variable $X$.

*2.* En déduire que $\mathbb{P}([X \geqslant 3]) \leqslant \dfrac{11}{27}$.

**Exercice 3.** ($\bigstar\bigstar$)
En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que :

$$\forall x > 0,\ \int_{-\infty}^{x} \mathrm{e}^{-\frac{t^2}{2}}\, dt \geqslant \sqrt{2\pi}\left(1 - \frac{1}{2x^2}\right)$$

## Convergence en loi

**Exercice 4.** ($\bigstar\bigstar$)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur un même espace probabilisé, indépendantes, suivant la même loi uniforme sur $[0, 1]$.
Pour tout $n \geqslant 1$, on pose :

$$M_n = \max(X_1, X_2, \ldots, X_n) \quad \text{et} \quad Y_n = n(1 - M_n)$$

On admet que $M_n$ et $Y_n$ sont des variables aléatoires à densité.

*1.* Déterminer la fonction de répartition de $M_n$, puis celle de $Y_n$.

*2.* Montrer que la suite $(Y_n)_{n\in\mathbb{N}^*}$ converge en loi vers une variable remarquable.

**Exercice 5.** ($\bigstar\bigstar$)

*1.* Soit $X$ une v.a.r. suivant une loi exponentielle de paramètre 1.
Soit $Z$ la variable aléatoire définie par $Z = -\ln(X)$.
Déterminer la loi de $Z$.

*2.* On considère une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes suivant toutes la loi exponentielle de paramètre 1. On pose pour tout $n \in \mathbb{N}^*$, $Y_n = \max\limits_{1\leqslant k\leqslant n}(X_k) - \ln(n)$.

*a)* Déterminer la fonction de répartition de $Y_n$.

*b)* Montrer que la suite de variables aléatoires $(Y_n)_{n\in\mathbb{N}^*}$ converge en loi vers $Z$ lorsque $n$ tend vers $+\infty$.

**Exercice 6.** (★★)
Soit $n \in \mathbb{N}^*$. On considère une variable aléatoire réelle $X_n$ de loi exponentielle de paramètre $\frac{1}{n}$, et on définit la variable aléatoire $Y_n$, par $Y_n = \lfloor X_n \rfloor$, où $\lfloor x \rfloor$ désigne la partie entière du réel $x$. On rappelle que $\lfloor x \rfloor$ est égal à l'unique entier $k$ vérifiant $k \leqslant x < k+1$.

***1.*** ***a)*** Préciser les valeurs prises par $Y_n$ et déterminer la loi de $Y_n$.

***b)*** Calculer l'espérance $\mathbb{E}(Y_n)$ et la variance $\mathbb{V}(Y_n)$.

***2.*** Pour tout $n$ de $\mathbb{N}^*$, on pose $Z_n = X_n - Y_n$.
On définit ainsi une suite $(Z_n)_{n \geqslant 1}$ de variables aléatoires réelles.

***a)*** Montrer que pour tout $n \in \mathbb{N}^*$, $Z_n$ est une variable aléatoire à densité et déterminer une densité de $Z_n$.

***b)*** Montrer que la suite $(Z_n)_{n \geqslant 1}$ converge en loi vers une variable aléatoire $U$ dont on donnera la loi.

***c)*** Calculer l'espérance $\mathbb{E}(Z_n)$. Montrer que la suite $(\mathbb{E}(Z_n))_{n \geqslant 1}$ admet une limite que l'on déterminera. A-t-on $\lim\limits_{n \to +\infty} \mathbb{E}(Z_n) = \mathbb{E}(U)$ ?

**Exercice 7.** (★★)
Soit $n$ un entier supérieur ou égal à 2.
Un groupe de $n$ personnes portant des numéros de 1 à $n$ tire à tour de rôle, avec remise, un numéro compris entre 1 et $n$ dans une urne. Si l'une des personnes a tiré son propre numéro, on recommence la série des $n$ tirages, sinon on arrête.
On note $X_n$ le nombre de séries de $n$ tirages qu'il faut effectuer pour que chaque personne ait tiré un numéro différent du sien.

***1.*** Montrer que $X_n$ suit une loi géométrique dont on précisera le paramètre.

***2.*** Pour tout $k \in \mathbb{N}^*$, calculer $\lim\limits_{n \to +\infty} \mathbb{P}([X_n = k])$.
En déduire que la suite $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire $X$ dont on précisera la loi.

**Exercice 8.** (★★)
Soit $n$ un entier naturel strictement positif. On définit la fonction $f_n$ par :

$$\forall x \in \mathbb{R},\ f_n(x) = \begin{cases} a_n \left(1 - \dfrac{x}{n}\right)^n & \text{si } 0 \leqslant x \leqslant n \\ 0 & \text{sinon} \end{cases}$$

***1.*** Déterminer $a_n$ pour que $f_n$ soit une densité de probabilité.

***2.*** Soit alors $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires à densité telle que $X_n$ admette $f_n$ pour densité.

***a)*** Déterminer la fonction de répartition $F_n$ de $X_n$.

***b)*** Pour $x \in \mathbb{R}$, déterminer la limite de $F_n(x)$ lorsque $n$ tend vers $+\infty$.

***c)*** Montrer que la suite de variables aléatoires $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire $Z$ que l'on déterminera.

***3.*** On considère la fonction et le programme suivant :

```
function X = simulX(n)
    Y = n ⋆ rand()
    for k = 1:n
        U = n ⋆ rand()
        if U < Y then
            Y = U
        end
        X = Y
    end
endfunction
```

```
m = 20 000
n = 100
S = 0
for k = 1:m
    if simulX(n) <= 1 then
        S = S+1
    end
end
disp(S / m)
```

***a)*** Montrer que la fonction `X` de paramètre `n` simule une variable aléatoire de même loi que $X_n$.

***b)*** On donne $\mathrm{e}^{-1} \approx 0,37$.
En utilisant la question 2, ainsi que la loi faible des grands nombres, expliquer pourquoi le résultat affiché, après exécution du programme, est proche de $0,63$.

## Théorème Central Limite (TCL)

**Exercice 9.** (★★)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la même loi exponentielle de paramètre 1.

On pose, pour tout $n \in \mathbb{N}^*$, $\overline{X_n} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ et $X_n^* = \sqrt{n}(\overline{X_n}-1)$. On note $F_n$ la fonction de répartition de $X_n^*$ .

*1.* *a)* Donner l'espérance et la variance de $\overline{X_n}$.

*b)* Montrer que $\overline{X_n}$ converge en probabilité vers 1, c'est-à-dire que : $\forall \varepsilon > 0$, $\lim\limits_{n\to+\infty} P\left(\left|\overline{X_n}-1\right| > \varepsilon\right) = 0$.

*2.* *a)* Quelle est, pour $x \in \mathbb{R}$, la limite de $F_n(x)$ lorsque $n$ tend vers $+\infty$ ?

*b)* Calculer une valeur approchée de $\mathbb{P}\left(\left[1-\dfrac{2}{\sqrt{n}} \leqslant \overline{X_n} \leqslant 1+\dfrac{2}{\sqrt{n}}\right]\right)$, pour $n$ assez grand.

Si $\Phi$ la fonction de répartition de la loi normale réduite, on donne $\Phi(2) \approx 0{,}977$.

**Exercice 10.** (★★)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, suivant toutes la loi uniforme sur $[0,1]$.

On note $T_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$, et $T_n^*$ la variable centrée-réduite associée à $T_n$.

*1.* Étudier la convergence en loi de la suite de variables $(T_n^*)_{n\geqslant 1}$.

*2.* En déduire une fonction qui simule la loi normale centrée-réduite.
Puis, écrire une fonction de paramètres $m$ et $s$ qui simule une variable aléatoire de loi normale d'espérance $m$ et d'écart-type $s$.

**Exercice 11.** (★★) *(d'après* `EML 2006`*)*

*1.* Soit $Z$ une variable aléatoire suivant une loi géométrique de paramètre $p$.
Ainsi, pour tout $k \in \mathbb{N}^*$, $\mathbb{P}([Z=k]) = p\,(1-p)^{k-1}$.

*a)* Rappeler la valeur de l'espérance $E(Z)$ et celle de la variance $\mathbb{V}(Z)$ de la variable aléatoire $Z$

*b)* Écrire une fonction `Z` en langage **Scilab**, de paramètre `p`, qui simule la variable aléatoire $Z$.

*2.* Soient un entier $n$ supérieur ou égal à 2, et $n$ variables aléatoires indépendantes $Z_1$, $Z_2$, ..., $Z_n$, suivant toutes la loi géométrique de paramètre $p$.
*On considère donc un $n$-échantillon $(Z_1, Z_2, \dots, Z_n)$ de variables indépendantes et de même loi que $Z$.*

On considère la variable aléatoire $M_n = \dfrac{1}{n}\left(Z_1 + Z_2 + \cdots + Z_n\right)$.

*a)* Déterminer l'espérance $m$ et l'écart-type $\sigma_n$ de $M_n$.

*b)* Montrer que $\lim\limits_{n\to+\infty} \mathbb{P}([0 \leqslant M_n - m \leqslant \sigma_n])$ existe et exprimer sa valeur à l'aide de $\displaystyle\int_0^1 \mathrm{e}^{-\frac{x^2}{2}}\,dx$.

*c)* Écrire une fonction `M` en **Scilab**, de paramètres `n` et `p`, utilisant la fonction `Z` définie à la question 1.b), qui simule la variable $M_n$.

*d)* En se référant à la loi faible des grands nombres, et en utilisant la question 2.b), écrire un programme en **Scilab**, utilisant les fonctions `Z` et `M`, qui calcule et affiche une valeur approchée de $\displaystyle\int_0^1 \mathrm{e}^{-\frac{x^2}{2}}\,dx$.
On rappelle que `sqrt(x)` calcule $\sqrt{x}$ en **Scilab**, et que $\pi \approx 3,14$.

*e)* Justifier le résultat suivant : $\displaystyle\int_0^1 \mathrm{e}^{-\frac{x^2}{2}}\,dx = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mathrm{e}^{-\frac{\left(\frac{k}{n}\right)^2}{2}}$.
En déduire un second programme en **Scilab** qui calcule et affiche une valeur approchée de l'intégrale $\displaystyle\int_0^1 \mathrm{e}^{-\frac{x^2}{2}}\,dx$.

**Exercice 12.** (★★)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, suivant la même loi de Poisson de paramètre 1.

*1.* Pour tout $n \in \mathbb{N}^*$, donner la loi de $S_n = \sum\limits_{k=1}^{n} X_k$, son espérance et sa variance.

*2.* À l'aide du théorème de la limite centrée, déterminer $\lim\limits_{n\to+\infty} \mathrm{e}^{-n} \sum\limits_{k=0}^{n} \dfrac{n^k}{k!}$.

**Exercice 13.** (★★) *(d'après* `HEC 2001 - Maths III`*)*
On réalise une suite de lancers indépendants d'une pièce de monnaie équilibrée. On associe à cette expérience une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires indépendantes, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et suivant toutes la loi de Bernoulli de paramètre $\dfrac{1}{2}$.
Pour tout entier $n$ supérieur ou égal à 1, on pose $S_n = X_1 + \dots + X_n$.

*1.* Déterminer la loi de probabilité de la variable $S_n$.
Quelles sont l'espérance et la variance de $S_n$ ?

*2.* ***a)*** Montrer que pour tout réel $\varepsilon$ strictement positif, on peut trouver une constante $K_\varepsilon$ telle que, pour tout entier $n$ supérieur ou égal à 1, on ait l'inégalité : $\mathbb{P}\left(\left[\left|\dfrac{S_n}{n} - \dfrac{1}{2}\right| \geqslant \varepsilon\right]\right) \leqslant \dfrac{K_\varepsilon}{n}$.

***b)*** Déduire de la majoration obtenue que :

$$\forall r \in \left]0, \frac{1}{2}\right[, \lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geqslant \frac{1}{n^r}\right) = 0$$

*3.* Montrer d'autre part, à l'aide du théorème de la limite centrée, que la suite $\left(\mathbb{P}\left(\left[\left|\dfrac{S_n}{n} - \dfrac{1}{2}\right| > \dfrac{1}{\sqrt{n}}\right]\right)\right)_{n\in\mathbb{N}^*}$ admet une limite non nulle.

## Approximation

**Exercice 14.** (★★) *(extrait de* `HEC 2002 - Maths III`*)*
On appelle durée de vie d'un composant électronique la durée de fonctionnement de ce composant jusqu'à sa première panne éventuelle.
On suppose que la durée de vie d'un composant suit la loi géométrique de paramètre $p = \frac{1}{200}$.
Un premier composant est mis en service à l'instant 0 et, quand il tombe en panne, est remplacé instantanément par un composant identique qui sera remplacé à son tour à l'instant de sa première panne dans les mêmes conditions, et ainsi de suite.
Pour tout $n \in \mathbb{N}^*$, on note $U_n$ la variable aléatoire désignant le nombre de pannes (et donc de remplacements) survenues jusqu'à l'instant $n$ inclus.
On admet que $U_n$ suit la loi binomiale de paramètres $n$ et $p$.
On considère un appareillage électronique utilisant simultanément 1000 composants identiques fonctionnant indépendamment les uns des autres et dont la durée de vie suit la même loi géométrique de paramètre $p = \frac{1}{200}$.
À chaque instant, les composants en panne sont remplacés par des composants identiques comme précédemment.

*1.* Préciser la loi de la variable aléatoire $U$ désignant le nombre total de remplacements de composants effectués jusqu'à l'instant $n = 100$ inclus.

*2.* On désire qu'avec une probabilité de $0,95$, le stock de composants de rechange soit suffisant jusqu'à l'instant $n = 100$ inclus.
À combien peut-on évaluer ce stock ?
On donne : $\sqrt{\dfrac{995}{2}} \approx 22,3$ et, en désignant par $\Phi$ la fonction de répartition de la loi normale centrée réduite, $\Phi(1,65) \approx 0,95$.

**Exercice 15.** (★★) *(extrait de* `ECRICOME 2008`*)*
Pour ce jeu de hasard, la mise pour chaque partie est de 1 euro.
L'observation montre qu'une partie est gagnée avec la probabilité $\dfrac{1}{10}$, perdue avec la probabilité $\dfrac{9}{10}$.
Toute partie gagnée rapporte 3 euros. Les différentes parties sont indépendantes. Une personne décide de jouer $N$ parties ($N \geqslant 2$).
On note $X_N$ la variable aléatoire représentant le nombre de parties gagnées et $Y_N$ la variable aléatoire représentant le gain algébrique du joueur.

***1.*** Donner la loi de $X_N$ ainsi que la valeur de l'espérance et de la variance de cette variable.

***2.*** Exprimer $Y_N$ en fonction de $X_N$.
En déduire la valeur de l'espérance et de la variance de $Y_N$.

***3.*** La personne décide de jouer 60 parties.
On admet que l'on peut approcher $X_{60}$ par une loi de Poisson.

***a)*** Donner le paramètre de cette loi de Poisson.

***b)*** À l'issue des 60 parties, quelle est la probabilité que le joueur perde moins de 50 euros ?
*(cette probabilité sera impérativement calculée en utilisant l'annexe située à la fin de l'exercice)*

**Table de Poisson donnant les probabilités cumulées :** $\displaystyle\sum_{i=0}^{k} \mathrm{e}^{-\lambda} \frac{\lambda^i}{i!}$

| $k$ | $\lambda = 3$ | $\lambda = 4$ | $\lambda = 5$ | $\lambda = 6$ | $\lambda = 7$ |
|---|---|---|---|---|---|
| 0 | $0,0498$ | $0,0183$ | $0,0067$ | $0,0025$ | $0,0009$ |
| 1 | $0,1991$ | $0,0916$ | $0,0404$ | $0,0174$ | $0,0073$ |
| 2 | $0,4232$ | $0,2381$ | $0,1247$ | $0,0620$ | $0,0296$ |
| 3 | $0,6472$ | $0,4335$ | $0,2650$ | $0,1512$ | $0,0818$ |
| 4 | $0,8153$ | $0,6288$ | $0,4405$ | $0,2851$ | $0,1730$ |
| 5 | $0,9161$ | $0,7851$ | $0,6160$ | $0,4457$ | $0,3007$ |
| 6 | $0,9665$ | $0,8893$ | $0,7622$ | $0,6063$ | $0,4497$ |
| 7 | $0,9881$ | $0,9489$ | $0,8666$ | $0,7440$ | $0,5987$ |
| 8 | $0,9962$ | $0,9786$ | $0,9319$ | $0,8472$ | $0,7291$ |
| 9 | $0,9989$ | $0,9919$ | $0,9682$ | $0,9161$ | $0,8305$ |
| 10 | $0,9997$ | $0,9972$ | $0,9863$ | $0,9574$ | $0,9015$ |
| 11 | $0,9999$ | $0,9991$ | $0,9945$ | $0,9799$ | $0,9467$ |
| 12 | $1,0000$ | $0,9997$ | $0,9980$ | $0,9912$ | $0,9730$ |
| 13 | | $0,9999$ | $0,9993$ | $0,9964$ | $0,9872$ |
| 14 | | $1,0000$ | $0,9998$ | $0,9986$ | $0,9943$ |
| 15 | | | $0,9999$ | $0,9995$ | $0,9976$ |
| 16 | | | $1,0000$ | $0,9998$ | $0,9990$ |
| 17 | | | | $0,9999$ | $0,9996$ |
| 18 | | | | $1,0000$ | $0,9999$ |
| 19 | | | | | $1,0000$ |
| 20 | | | | | |

**Exercice 16.** (★★)
On veut estimer le pourcentage $p$ de réponses positives à un référendum. Pour cela, on effectue un sondage sur $n$ personnes, et on estime $p$ par la fréquence relative $F_n$ de « oui » sur les personnes sondées.
On suppose les réponses données par les personnes sondées mutuellement indépendantes.
On cherche la taille minimale $n$ de l'échantillon pour que la probabilité que la fréquence relative $F_n$ diffère de $p$ de plus de $0,01$ soit inférieure $0,05$.

*1.* Étudier la fonction $x \mapsto x(1-x)$ sur $[0,1]$.

*2.* **Première méthode** : utilisation de l'inégalité de Bienaymé-Tchebychev

*a)* Montrer, en utilisant l'inégalité de Bienaymé-Tchebychev, que :

$$\mathbb{P}\left([|F_n - p| > 0,01]\right) \leqslant \frac{2500}{n}$$

*b)* Conclure quant à la taille minimale cherchée $n$ de l'échantillon.

*3.* **Seconde méthode** : utilisation d'une approximation en loi

*a)* Expliquer pourquoi on peut approcher la loi de $\dfrac{\sqrt{n}}{\sqrt{p(1-p)}}(F_n - p)$ par la loi normale centrée réduite.

*b)* Conclure quant à la taille minimale cherchée $n$ de l'échantillon.
On donne : $\Phi(1,96) = 0,975$, et $98^2 = 9\,604$.
*(*$98^2 = (100-2)^2 = 100^2 - 4 \times 100 + 4 = 10000 - 400 + 4 = 9604$*)*

*4.* Discuter les avantages et les inconvénients des deux méthodes.
Quelle méthode l'institut de sondage va-t-il choisr ?

**Exercice 17.** (★★)
On estime à 60% la probabilité qu'une famille possède une voiture.
On admet que chaque famille possède au plus une voiture, et que chaque famille se comporte indépendamment des autres.
On construit un ensemble de logements pour $n$ famille.
On y adjoint un parking.
On note $P_n$ la variable aléatoire égale au nombre de familles possédant une voiture.

*1.* Expliquer pourquoi on peut approcher la loi de $P_n$ par une loi normale, dont on précisera les paramètres.

*2.* Dans cette question, on suppose que $n = 600$.

*a)* Calculer la probabilité de ne pas satisfaire toutes les demandes de stationnement si l'on construit 360 places de parking.

*b)* Le promoteur veut annoncer : « toute famille désirant un parking peut en bénéficier ». Quel est le nombre de places que doit offrir le parking pour que sa publicité coure un risque limité à 2% d'être mensongère ? On donne $\Phi(2,06) = 0,98$.

*3.* Le promoteur veut annoncer : « toute famille désirant un parking peut en bénéficier ». Le nombre de place de parking qu'il est possible de construire est limité à 300.
Combien le promoteur doit-il construire de logements pour que sa publicité coure un risque limité à 2% d'être mensongère ? On donne $\Phi(2,06) = 0,98$.