

Probabilités – Lois usuelles

Table des matières.

MATHS – S 5

1. Simulations de lois usuelles.	1
1.1. Loi uniforme continue $\mathcal{U}(0, 1)$	1
1.2. Loi binomiale $\mathcal{B}(n, p)$	2
1.3. Loi exponentielle $\mathcal{E}(\lambda)$	2
1.4. Loi de Poisson $\mathcal{P}(\lambda)$	3
1.5. Loi normale $\mathcal{N}(m, \sigma^2)$	3
2. Convergence & loi des grands nombres.	4
2.1. Loi forte des grands nombres.	4
2.2. Convergence presque sûre et convergence en probabilité.	4
3. Théorème central limite, convergence en loi et approximations.	4
3.1. Théorème central limite.	4
3.2. Approximer une loi binomiale par loi de Poisson.	5
3.3. Approximer une loi binomiale par loi normale.	5
3.4. Approximer une loi de Poisson par loi normale.	5
4. Conclusions.	5
5. Annexe. Comment vérifier l'écart-type ?	8
5.1. Loi uniforme $\mathcal{U}(0, 1)$	8
5.2. Loi binomiale $\mathcal{B}(n, p)$	8
5.3. Loi exponentielle $\mathcal{E}(\lambda)$	8
5.4. Loi de Poisson $\mathcal{P}(\lambda)$	8
5.5. Loi normale $\mathcal{N}(m, \sigma^2)$	8

L'objectif de ce TP est de simuler des variables aléatoires suivant différentes lois, d'observer la loi forte des grands nombres, le théorème central limite, puis d'approximer des lois par d'autres lois.

En annexe (Section V), on retrouve les calculs détaillés sur la proportion de valeurs à un écart-type de l'espérance.

La densité de probabilité de la variable U est :

$$f_U(x) = \mathbb{1}_{[0,1]}(x)$$

En MATLAB, on simule une telle variable à l'aide de la fonction `rand`, et on obtient les histogrammes et les nuages de points suivants.

I. Simulations de lois usuelles.

I.1. Loi uniforme continue $\mathcal{U}(0, 1)$.

On rappelle que l'espérance et la variance d'une variable aléatoire U suivant la loi $\mathcal{U}(0, 1)$:

$$\mathbb{E}[U] = \frac{1}{2} \quad \text{et} \quad \mathbb{V}[U] = \frac{1}{12}.$$

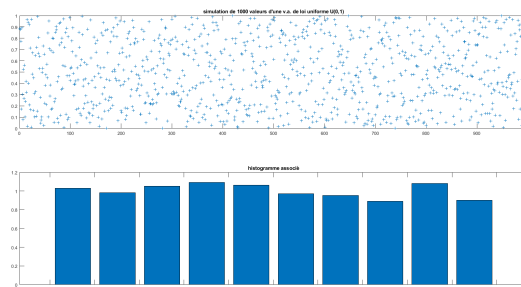


Figure 1. Simulation de U pour 1 000 valeurs

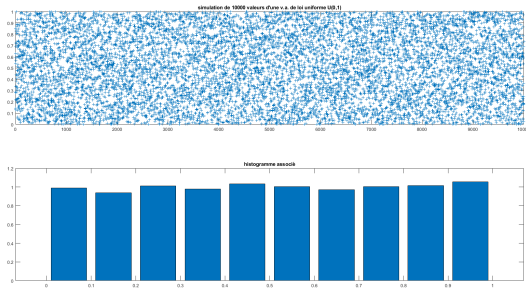


Figure 2. Simulation de U pour 10 000 valeurs

L'histogramme correspond approximativement à la densité de probabilité de U , et la différence entre les deux diminue quand N augmente, où N est le nombre de valeurs simulées.

La moyenne des résultats est d'environ $1/2$, ce qui correspond à l'espérance de U . De plus, environ 58 % des valeurs sont comprises dans l'intervalle $[\frac{1}{2} - \sigma_U, \frac{1}{2} + \sigma_U]$, avec

$$\sigma_U = \sqrt{\text{V}[U]} = \frac{\sqrt{3}}{6} \approx 0,29.$$

Ainsi, l'écart-type, et par suite la variance sont cohérents.

I.2. Loi binomiale $\mathcal{B}(n, p)$.

Pour une variable aléatoire B suivant $\mathcal{B}(n, p)$, nous avons

$$\mathbb{E}[B] = np \quad \text{et} \quad \text{V}[B] = npq,$$

où $q = 1 - p$. La densité de probabilité de B est, pour $m \in \llbracket 0, n \rrbracket$,

$$f_B(m) = \sum_{k=0}^m \binom{n}{k} p^k q^{n-k}.$$

On simule une telle variable à l'aide de la fonction `binomiale` pour obtenir 1 000 valeurs, et on obtient les histogrammes et nuages de points suivants.

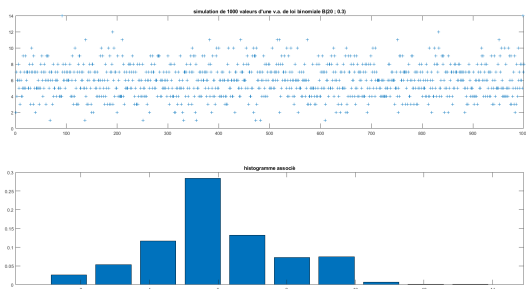


Figure 3. Simulation de B avec $p = 0,3$ et $n = 20$

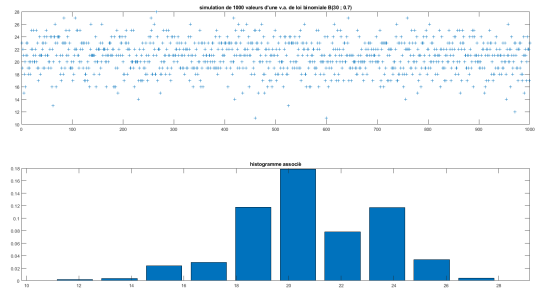


Figure 4. Simulation de B avec $p = 0,7$ et $n = 30$

La moyenne est bien de 6 pour la Figure 3, et 21 pour la Figure 4, ce qui correspond à l'espérance de B . De même, dans l'intervalle $[\mathbb{E}[B] - \sigma_B, \mathbb{E}[B] + \sigma_B]$, on retrouve bien environ 80 % des valeurs. De plus, l'histogramme correspond approximativement à la fonction de densité de probabilité de B .

I.3. Loi exponentielle $\mathcal{E}(\lambda)$.

Soit E une variable aléatoire suivant la loi exponentielle $\mathcal{E}(\lambda)$. Ainsi, sa fonction de densité de probabilité est :

$$f_E(x) = \mathbb{1}_{[0, +\infty[}(x) \cdot \lambda e^{-\lambda x}.$$

De plus, son espérance et sa variance sont :

$$\mathbb{E}[E] = \frac{1}{\lambda} \quad \text{et} \quad \text{V}[E] = \frac{1}{\lambda^2}.$$

On simule 1 000 valeurs de E avec $\lambda = 3$, puis avec $\lambda = 18$, et on obtient les figures ci-dessous.

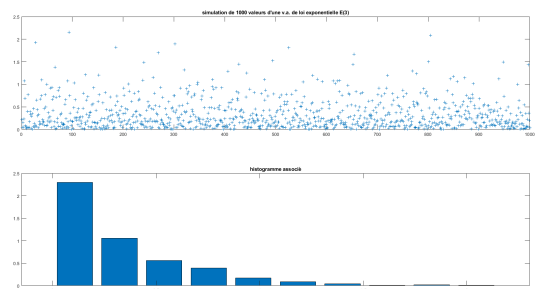


Figure 5. Simulation de E suivant $\mathcal{E}(3)$

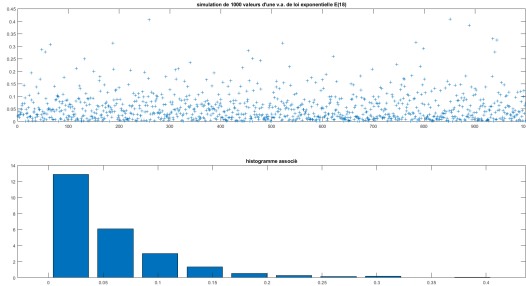


Figure 6. Simulation de E suivant $\mathcal{E}(18)$

Les deux histogrammes ci-dessous correspondent à l'expression de la densité de probabilité de E . La moyenne se situe bien vers $\frac{1}{\lambda}$, et on a bien 86 % des valeurs dans l'intervalle $[0, \frac{2}{\lambda}]$. On en déduit que l'espérance et la variance sont cohérents.

I.4. Loi de Poisson $\mathcal{P}(\lambda)$.

Pour une variable P aléatoire suivant la loi de Poisson $\mathcal{P}(\lambda)$, l'espérance et la variance sont respectivement :

$$\mathbb{E}[P] = \lambda \quad \text{et} \quad \mathbb{V}[P] = \lambda.$$

De plus, sa densité de probabilité est donnée par : pour $n \in \mathbb{N}$,

$$f_P(n) = e^{-\lambda} \frac{\lambda^n}{n!}.$$

On simule une telle variable aléatoire à l'aide de la fonction MATLAB `poissrnd`, et on obtient les figures suivantes.

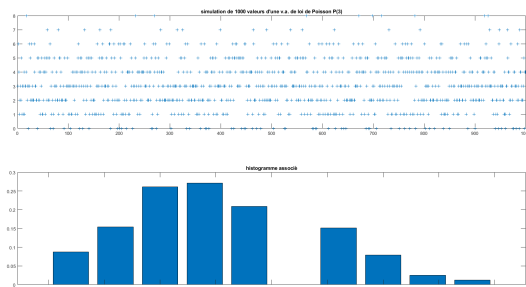


Figure 7. Simulation de P suivant $\mathcal{P}(3)$

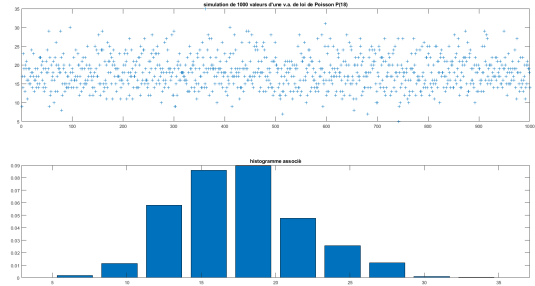


Figure 8. Simulation de P suivant $\mathcal{P}(18)$

L'histogramme correspond approximativement à la fonction de densité de probabilités de P . De plus, la moyenne est bien de λ , et on a bien 71 % et 87 % des valeurs dans l'intervalle $[\lambda - \sqrt{\lambda}, \lambda + \sqrt{\lambda}]$. On en déduit que l'espérance et la variance sont cohérents avec les résultats expérimentaux.

I.5. Loi normale $\mathcal{N}(m, \sigma^2)$.

Soit N une variable suivant la loi $\mathcal{N}(m, \sigma^2)$. Ainsi, son espérance est de $\mathbb{E}[N] = m$ et sa variance est de $\mathbb{V}[N] = \sigma^2$. Sa fonction de densité est :

$$f_N(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \cdot \left(\frac{x-m}{\sigma}\right)^2\right).$$

On utilise la fonction `randn` de MATLAB pour simuler une variable \hat{N} suivant la loi normale centrée réduite $\mathcal{N}(0, 1)$, en définissant N comme $N = \sigma\hat{N} + m$. On obtient les résultats expérimentaux suivants.

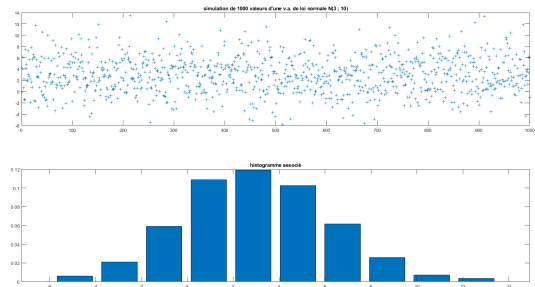


Figure 9. Simulation de N suivant $\mathcal{N}(3, 10)$

L'histogramme est cohérent avec la densité de probabilité. L'espérance et l'écart-type sont cohérents avec la moyenne et la dispersion des valeurs.

II. Convergence & loi des grands nombres.

II.1. Loi forte des grands nombres.

Dans cette section, on observe la loi forte des grands nombres. En simulant une variable X aléatoire n fois, la moyenne converge vers l'espérance $\mathbb{E}[X]$.

Par exemple, on observe cette loi avec une variable aléatoire suivant la loi $\mathcal{U}(0, 1)$, d'espérance $\frac{1}{2}$. On représente les résultats expérimentaux sur la Figure 17 (page 6). Sur cette figure, les 9 graphes représentent 9 simulations différentes. Dans chaque cas, la moyenne converge bien vers $\frac{1}{2}$.

On réalise le même test avec une variable aléatoire suivant la loi de Poisson $\mathcal{P}(\lambda)$, avec $\lambda = 3$. Son espérance est de λ . On représente les résultats expérimentaux sur la Figure 18 (page 7). On remarque que, pour toutes les simulations, la moyenne converge vers 3, *i.e.* l'espérance.

II.2. Convergence presque sûre et convergence en probabilité.

Sur la figure ci-dessous, on représente le graphe des réalisations x_n de la variable aléatoire X_n suivant la loi décrite dans l'énoncé, en fonction de l'entier $n \in \llbracket 1, n_{\max} \rrbracket$. Dans la figure ci-dessous, ainsi que la suivante, l'entier n_{\max} vaut 1 000.

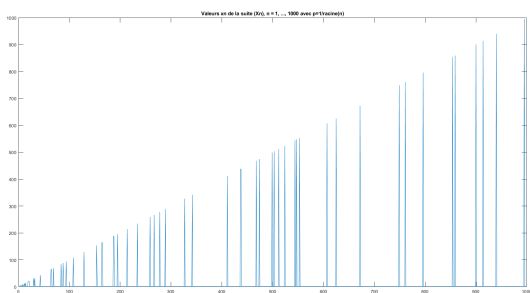


Figure 10. Réalisations x_n de X_n en fonction de n

On observe que la suite (X_n) ne converge pas presque sûrement vers 0. En effet, les « pics » ont une hauteur qui augmente strictement. Ceci est dû à la non convergence de la série $\sum 1/\sqrt{n}$, par critère de Riemann.

En remplaçant le $1/\sqrt{n}$ dans $P(X_n = n)$ par $1/n^2$, la suite (X_n) converge presque sûrement vers 0. En effet, la série $\sum 1/n^2$ converge. Sur la figure ci-dessous, on n'observe pas de « pics » passé $n = 10$.

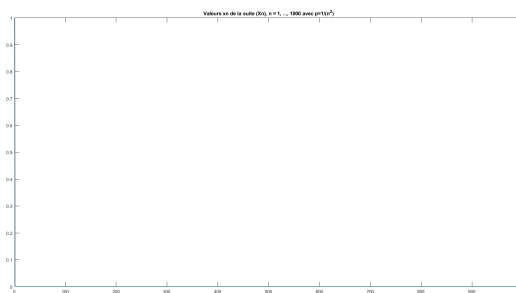


Figure 11. Réalisations x_n de X_n en fonction de n

III. Théorème central limite, convergence en loi et approximations.

III.1. Théorème central limite.

On considère (X_n) une suite de variables aléatoires indépendantes suivant la loi $\mathcal{U}(0, 1)$. Alors, d'après le théorème central limite, la suite (Z_n) converge en loi vers Z qui suit une loi normale centrée réduite $\mathcal{N}(0, 1)$, où

$$Z_n = \frac{\bar{X}_n - \mathbb{E}[X]}{\sigma_X/\sqrt{n}}.$$

Sur la figure ci-après, on représente 1 000 réalisations de la variable aléatoire $Z_{n_{\max}}$, où l'entier n_{\max} vaut 1 000.

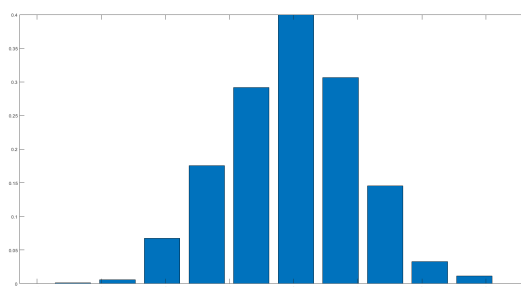


Figure 12. 1 000 réalisations de $Z_{n_{\max}}$, pour $X_n \sim \mathcal{U}(0, 1)$

Changeons la loi suivie par les variables X_n par une loi de Poisson $\mathcal{P}(\lambda)$. On définit la suite Z_n de la même manière et sur la figure ci-après, on représente 1 000 réalisations de la variable aléatoire $Z_{n_{\max}}$.

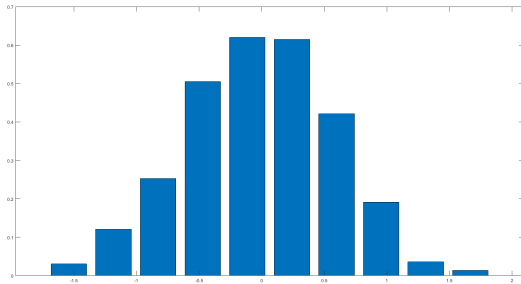


Figure 13. 1 000 réalisations de $Z_{n_{\max}}$, pour $X_n \sim \mathcal{P}(3)$

Dans les deux cas, par comparaison avec l'histogramme d'une variable suivant la loi normale centrée réduite $\mathcal{N}(0, 1)$, on en déduit que la loi de $Z_{n_{\max}}$ « ressemble » à une loi normale centrée réduite. En effet, la suite (Z_n) converge en loi vers Z suivant une loi normale centrée réduite.

III.2. Approximer une loi binomiale par loi de Poisson.

Comme indiqué dans l'énoncé, on approxime la loi binomiale $\mathcal{B}(n, p)$ par la loi de Poisson $\mathcal{P}(np)$. On représente cette approximation dans l'histogramme sur la figure ci-après. On choisit les valeurs $n = 100 > 50$, et $p = 1/10 < 18\%$.

Cette approximation est relativement correcte.

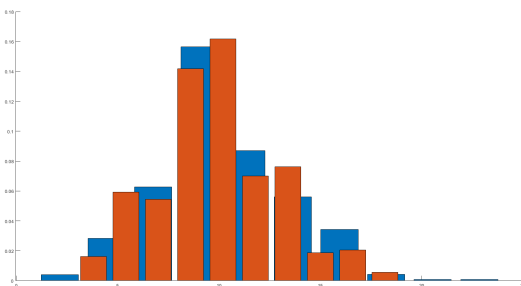


Figure 14. Approximation de $\mathcal{B}(100, 10\%)$ par $\mathcal{P}(10)$

Dans cette figure, comme dans celles d'après, on représente en orange l'histogramme généré pour la loi originelle, et en bleu celui de l'approximation.

III.3. Approximer une loi binomiale par loi normale.

On approxime la loi binomiale $\mathcal{B}(n, p)$ par la loi normale $\mathcal{N}(np, npq)$, où $q = 1 - p$. On choisit les valeurs $n = 100 > 50$ et $p = 3/10 > 18\%$. On obtient l'histogramme représenté sur la figure ci-après.

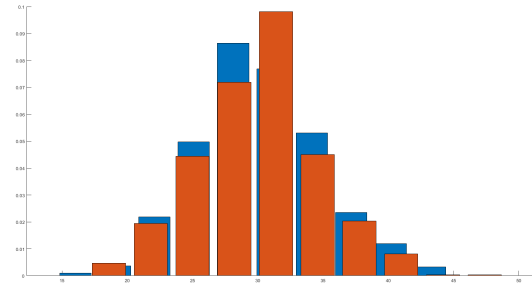


Figure 15. Approximation de $\mathcal{B}(100, 30\%)$ par $\mathcal{N}(30, 21)$

III.4. Approximer une loi de Poisson par loi normale.

On approxime la loi de Poisson $\mathcal{P}(\lambda)$ par la loi normale $\mathcal{N}(\lambda, \lambda)$. On choisit $\lambda = 30 > 18$. On obtient l'histogramme représenté sur la figure ci-après.

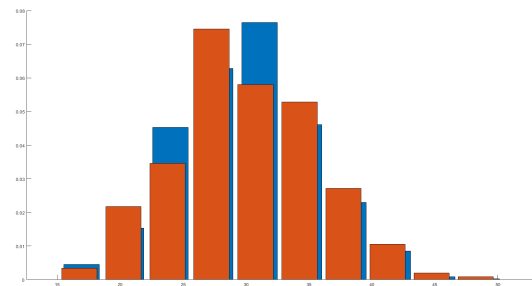


Figure 16. Approximation de $\mathcal{P}(30)$ par $\mathcal{N}(30, 30)$

IV. Conclusions.

La simulation de variables aléatoires suivant les lois usuelles en MATLAB est assez simple. De ces simulations, on peut observer la fonction de densité de probabilité, à partir d'un histogramme des résultats de la simulation. Pour un faible nombre de simulations, les résultats sont assez bruités. Dans ce TP, chaque histogramme représentait les données de 1 000 simulations (hormis la Figure 2, où 10 000 simulations étaient réalisées). Il faut donc réaliser un nombre plus important de simulations.

Observer la convergence presque sûre (ou non) et la convergence en loi est possible informatiquement, en passant par une représentation des simulations.

Finalement, il est possible d'approximer des lois par d'autres lois. Ceci peut être réalisé pour améliorer l'efficacité d'un programme basé sur ces simulations, par exemple.

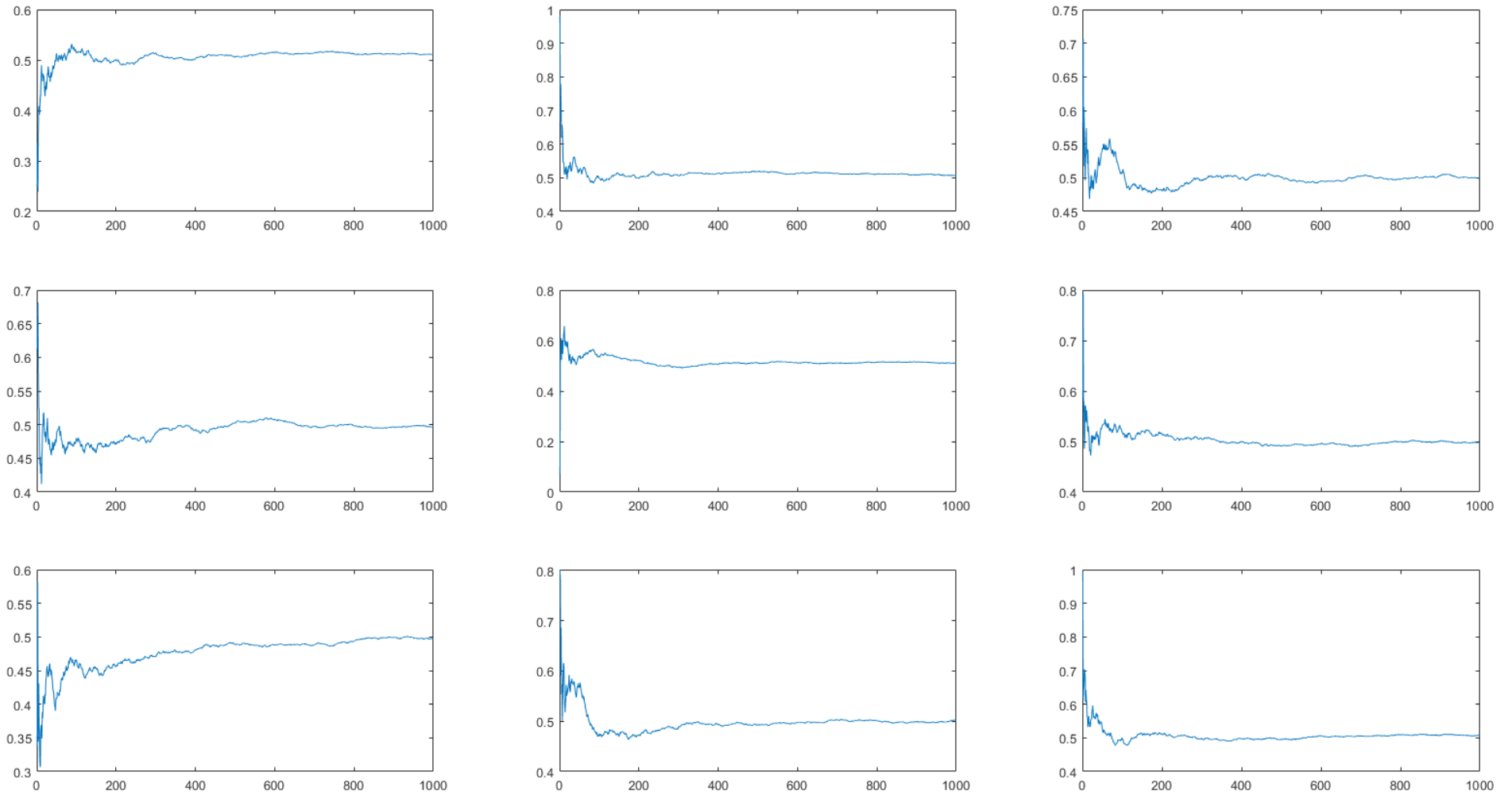


Figure 17. Observation de la loi forte des grands nombres à l'aide d'une loi uniforme $\mathcal{U}(0, 1)$

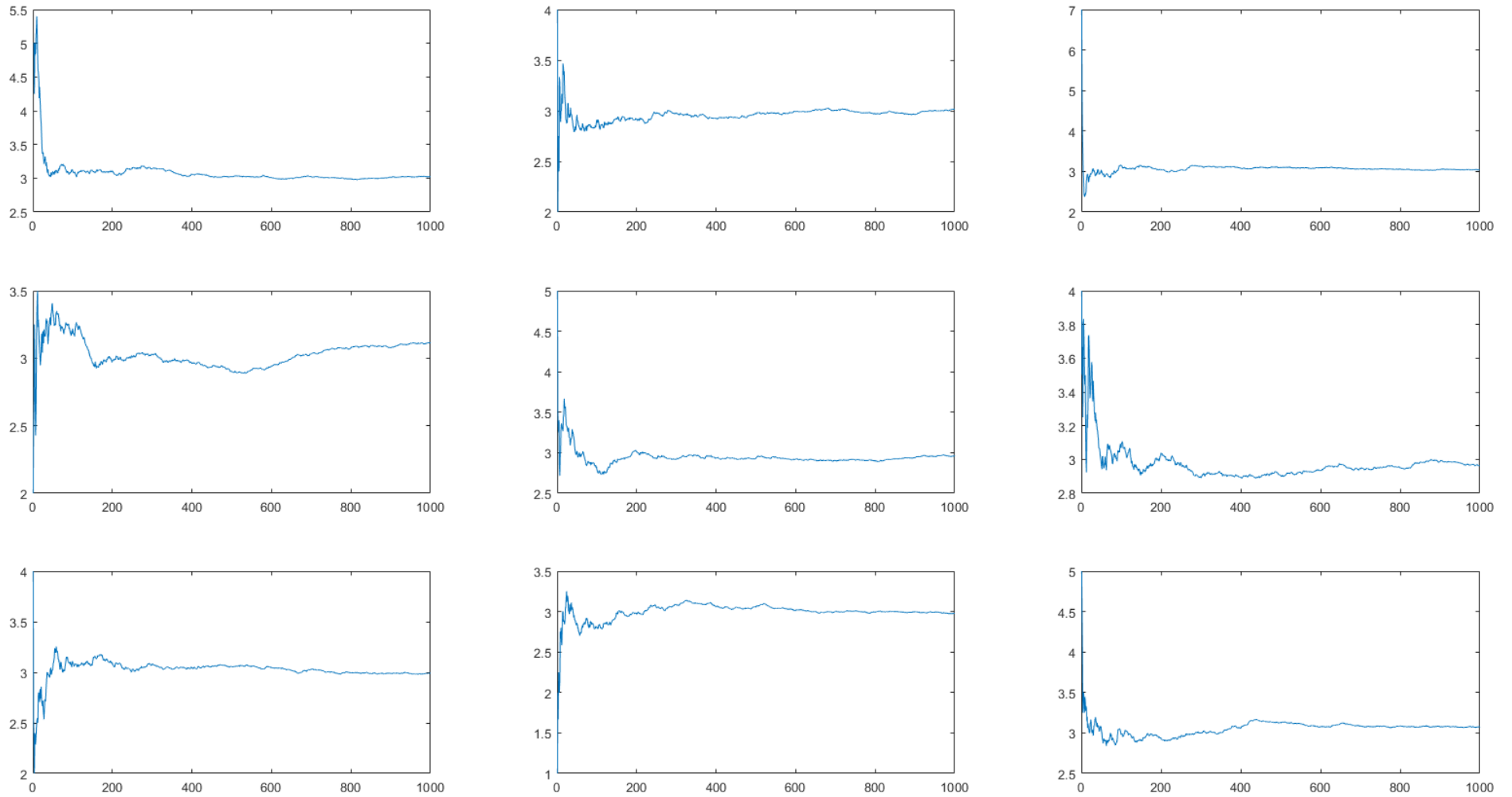


Figure 18. Observation de la loi forte des grands nombres à l'aide d'une loi de Poisson $\mathcal{P}(3)$

V. Annexe. Comment vérifier l'écart-type ?

Dans cette annexe, on détaille les calculs de la probabilité $P(|X - \mathbb{E}[X]| \leq \sigma_X)$, pour les différentes lois considérées.

On comparera la distribution des valeurs obtenues par simulation à ces résultats théoriques.

V.1. Loi uniforme $\mathcal{U}(0, 1)$.

Soit $U \sim \mathcal{U}(0, 1)$ une variable aléatoire. On calcule :

$$P(|U - \mathbb{E}[U]| \leq \sigma_U) = \int_{\mathbb{E}[U] - \sigma_U}^{\mathbb{E}[U] + \sigma_U} f_U(t) dt.$$

Numériquement, cette intégrale vaut approximativement 58 %.

V.2. Loi binomiale $\mathcal{B}(n, p)$.

Soit $B \sim \mathcal{B}(n, p)$ une variable aléatoire. On calcule

$$P(|B - \mathbb{E}[B]| \leq \sigma_B) = \sum_{k=\mathbb{E}[B] - \sigma_B}^{\mathbb{E}[B] + \sigma_B} P(B = k)$$

Numériquement, on trouve que cette somme vaut environ 78 % pour $(n, p) = (20, 0.3)$, et qu'elle vaut 84 % pour $(n, p) = (30, 0.7)$.

V.3. Loi exponentielle $\mathcal{E}(\lambda)$.

On réalise les mêmes calculs que pour la loi uniforme, mais avec $E \sim \mathcal{E}(\lambda)$. Numériquement, on trouve que $P(|E - \mathbb{E}[E]| \leq \sigma_E)$ vaut environ 86 %, quelle que soit la valeur de λ choisie.

V.4. Loi de Poisson $\mathcal{P}(\lambda)$.

On réalise les mêmes calculs que pour la loi binomiale, mais avec $P \sim \mathcal{P}(\lambda)$. Numériquement, on trouve que $P(|P - \mathbb{E}[P]| \leq \sigma_P)$ vaut environ 87 % pour $\lambda = 3$ et 71 % pour $\lambda = 18$.

V.5. Loi normale $\mathcal{N}(m, \sigma^2)$.

On réalise les mêmes calculs que pour la loi uniforme, mais avec $N \sim \mathcal{N}(m, \sigma^2)$. Numériquement, on trouve que $P(|E - \mathbb{E}[E]| \leq \sigma_E)$ vaut environ 68 %, quelle que soient les valeurs de m et de σ^2 choisies.