Chapitre 15

Convergence et approximation

Compétences attendues.

✓ Connaitre et savoir utiliser les inégalités de Markov et de Bienaymé-Tchebychev.

✓ Connaitre et savoir utiliser la loi faible des grands nombres.

✓ Étudier la convergence en loi d'une suite (X_n) de variables aléatoires.

✓ Connaitre et savoir utiliser le théorème limite central.

✓ Approximer une loi par une autre.

Anthony Mansuy
Professeur de Mathématiques en deuxième année de CPGE filière ECG au Lycée Clemenceau (Reims)
Page personnelle : `http://anthony-mansuy.fr`
E-mail : `mansuy.anthony@hotmail.fr`

1 Inégalités de concentration

1.1 Inégalité de Markov

Propriété 1 (Inégalité de Markov)

Si X est une variable aléatoire **positive** admettant une espérance, alors :

$$\forall a > 0, \quad P(X \geq a) \leq \frac{E(X)}{a}.$$

Preuve.

$\square$

On déduit de l'inégalité de Markov la :

Propriété 2 (Corollaire de l'inégalité de Markov)

Si X est une variable aléatoire admettant un moment d'ordre r, alors :

$$\forall a > 0, \quad P(|X| \geq a) \leq \frac{E(|X|^r)}{a^r}.$$

Preuve.

$\square$

Remarque. On retiendra en particulier que si X admet un moment d'ordre 2 (ou de manière équivalente une variance), alors :

$$\forall a > 0, \quad P(|X| \geq a) \leq \frac{E(|X|^2)}{a^2}.$$

1.2 Inégalité de Bienaymé-Tchebychev

Propriété 3 (Inégalité de Bienaymé-Tchebychev)

Si X est une variable aléatoire admettant une variance, alors :

$$\forall \varepsilon > 0, \quad P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}.$$

Preuve.

$\square$

Remarque. L'inégalité de Bienaymé-Tchebychev permet d'obtenir une majoration de $P(|X - E(X)| \geq \varepsilon)$ sans connaitre la loi de X : il suffit simplement de connaître son espérance et sa variance. Cette majoration est cependant assez grossière, souvent bien supérieure à $P(|X - E(X)| \geq \varepsilon)$.

1.3 Loi faible des grands nombres

Pour éclairer la signification de la loi faible des grands nombres, nous donnons la définition de la convergence en probabilité d'une suite de variables aléatoires.

Définition.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires, et X une variable aléatoire. On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en probabilité vers** X, et on note $X_n \xrightarrow{P} X$, lorsqu'on a :

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P(|X_n - X| \geq \varepsilon) = 0.$$

Remarque. En pratique, cela signifie que pour tout $\omega \in \Omega$ (presque sûrement), la suite $(X_n(\omega))$ converge vers $X(\omega)$ au sens de la convergence usuelle d'une suite de réels.

Théorème 4 (Loi faible des grands nombres)

On suppose que :

- $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires **mutuellement indépendantes** ;
- Les variables aléatoires X_n admettent la même espérance m et la même variance.

Pour tout $n \in \mathbb{N}^*$, on pose $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$. Alors $\overline{X_n}$ converge en probabilité vers la variable aléatoire constante égale à m, soit encore :

$$\overline{X_n} \xrightarrow{P} m.$$

Preuve.

□

En appliquant la loi faible des grands nombres au cas de variables de Bernoulli, on obtient le résultat suivant :

Propriété 5 (Théorème d'or de Bernoulli)

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables de Bernoulli indépendantes de même paramètre p. Alors $\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge en probabilité vers la variable constante égale à p, soit encore :

$$\overline{X_n} \xrightarrow{P} p.$$

Conséquence. On considère une succession d'expériences identiques indépendantes, au cours de chacune desquelles un événement A est susceptible de se produire avec probabilité $p = P(A)$. Pour tout $n \geq 1$, on définit la variable X_n par :

$$X_n = \begin{cases} 1 & \text{si l'événement } A \text{ se réalise au cours de la } n\text{-ième expérience,} \\ 0 & \text{sinon.} \end{cases}$$

Les variables aléatoires X_n sont indépendantes et suivent toutes la loi de Bernoulli de paramètre p.

Le théorème d'or de Bernoulli affirme que la fréquence $\overline{X_n}$ de réalisation de l'événement A au cours des n premières expériences converge en probabilité vers $p = P(A)$. Ainsi,

la fréquence statistique de réalisation d'un événement tend vers la probabilité de cet événement.

Ce résultat est fondamental : c'est lui qui justifie que le formalisme que nous employons correspond bien à l'intuition que l'on se fait d'une probabilité.

Remarque. On peut donner un ordre de grandeur de la vitesse de convergence de $\overline{X_n}$ vers p.

Cette inégalité permet de donner le nombre de répétitions de l'expérience à faire pour approcher p avec une marge d'erreur fixée.

Exemple. On dispose d'une pièce truquée et on souhaite évaluer la probabilité p qu'elle tombe sur pile. Si on note X_i la variable de Bernoulli qui vaut 1 si la pièce tombe sur pile au i-ème lancé, on sait d'après le théorème d'Or de Bernoulli que la fréquence $\overline{X_n} = \dfrac{X_1 + \cdots + X_n}{n}$ d'apparition du pile au cours de n lancers tend en probabilité p. On n'est cependant pas à l'abri d'une série de lancers où on aurait pas "eu de chance", c'est-à-dire telle que $\overline{X_n}$ prenne une valeur éloignée de la valeur de p. On va donc chercher le nombre n de lancers à effectuer pour approcher p à $0,1$ près en quantifiant ce risque, c'est-à-dire en étant sûr de notre résultat avec une probabilité de 99%.

Exemple. On modélise cette situation à l'aide de `Python`.

```
#Choix de la pièce
p = np.random()
#Fréquence de piles pour 2500 lancers
F = np.mean(np.binomial(1, p, 2500))
print(F)
```

En exécutant ce programme, qui correspond à 2500 lancés de la pièce truquée de probabilité p, on obtient un résultat égal à `0.7656`.

Ainsi on peut affirmer que $0.6656 \leq p \leq 0.8656$ avec une probabilité de 99%. En exécutant `print(p)`, on constate que $p = 0.7560439$ et satisfait donc bien les inégalités précédentes.

2 Convergence en loi

2.1 Définition

Définition.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire toutes définies sur $(\Omega, \mathscr{A}, P)$. Notons F_{X_n} et F_X leurs fonctions de répartition respectives.

On dit que $(X_n)_{n\in\mathbb{N}}$ **converge en loi vers** X, et on note $X_n \xrightarrow{\mathcal{L}} X$, si pour tout $x \in \mathbb{R}$ **en lequel** F_X **est continue**, on a :

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x).$$

Rappel. Si X est une variable à densité, F_X est continue sur $\mathbb{R}$, et on étudiera donc la limite de $F_{X_n}(x)$ pour tout $x \in \mathbb{R}$.

Remarques.

1. Si $X_n \xrightarrow{\mathcal{L}} X$, cela signifie concrètement que pour n assez grand, la loi de X_n "ressemble beaucoup" à la loi de X.
2. La convergence en loi ne dépend que des lois des variables aléatoires. En particulier, si $X_n \xrightarrow{\mathcal{L}} X$ et si Y suit la même loi que X, alors $X_n \xrightarrow{\mathcal{L}} Y$. Il n'y a donc pas unicité de la limite pour la convergence en loi.

Méthode.

Pour étudier la convergence en loi d'une suite (X_n) de variables aléatoires, on procède en trois étapes :

Étape 1 : Calcul des F_{X_n}.

On détermine les fonctions de répartitions F_{X_n} des X_n.

Étape 2 : Convergence ponctuelle des F_{X_n}.

*On **fixe** $x \in \mathbb{R}$, et on étudie la convergence de la suite de réels $(F_{X_n}(x))_{n\in\mathbb{N}}$. On note $F(x)$ sa limite.*

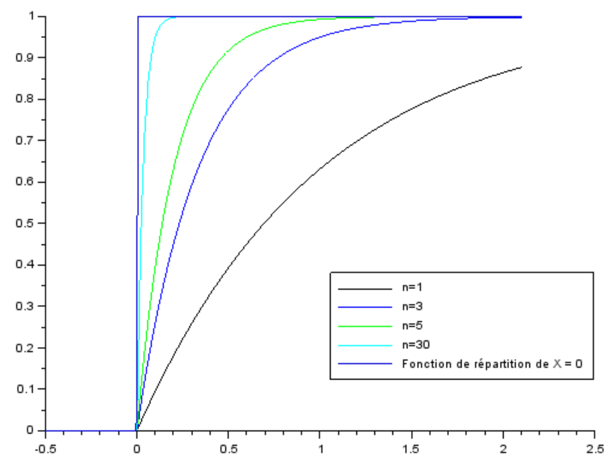
Étape 3 : F est-elle la fonction de répartition d'une variable aléatoire ?

On vérifie si F est bien la fonction de répartition d'une variable aléatoire, c'est-à-dire :

* $\lim\limits_{x\to-\infty} F(x) = 0$ *et* $\lim\limits_{x\to+\infty} F(x) = 1$ *;*
* *F est croissante ;*
* *F est continue à droite en tout point.*

Si ces conditions sont bien vérifiées, alors X_n converge en loi vers toute variable aléatoire de fonction de répartition F.

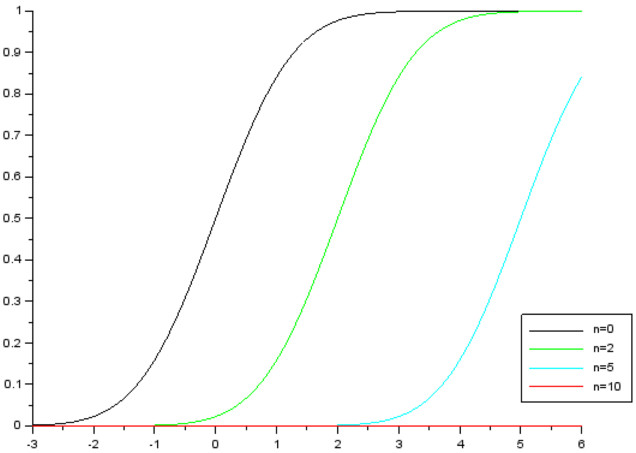
Exercice. Soit (X_n) une suite de variables aléatoires telles que $X_n \hookrightarrow \mathcal{E}(n)$.
Étudier la convergence en loi de (X_n).



Courbes représentatives de F_{X_n} et de F_X

Remarque. L'exemple précédent montre qu'une suite de variables aléatoires à densité peut converger en loi vers une variable aléatoire discrète (l'inverse étant aussi possible).

Exercice. Soit (X_n) une suite de variables aléatoires telles que $X_n \hookrightarrow \mathcal{N}(n, 1)$.
Étudier la convergence en loi de (X_n).



Courbes représentatives des F_{X_n}.

Propriété 6 (Probabilités et convergence en loi)

Si $X_n \xrightarrow{\mathcal{L}} X$, alors pour tous $a < b$ points de continuité de F_X, on a :

$$\lim_{n \to +\infty} P(a < X_n \leq b) = P(a < X \leq b).$$

Preuve.

□

2.2 Cas des variables aléatoires discrètes

Une conséquence de la propriété précédente est le théorème suivant, permettant la caractérisation de la convergence en loi pour les variables aléatoires à valeurs dans $\mathbb{N}$:

Théorème 7 (Convergence en loi - Cas des variables discrètes)

On suppose que les variables aléatoires $(X_n)_{n\in\mathbb{N}}$ et X sont **toutes à valeurs dans** $\mathbb{N}$.
Alors $X_n \xrightarrow{\mathcal{L}} X$ si et seulement si on a :

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P(X_n = k) = P(X = k).$$

Méthode.

*Pour étudier la convergence en loi d'une suite (X_n) de variables aléatoires discrètes **toutes à valeurs dans** $\mathbb{N}$, on procèdera comme suit :*

Étape 1 : Calcul des probabilités ponctuelles des X_n.
On détermine pour tout $n \in \mathbb{N}$ et $k \in \mathbb{N}$, $p_{n,k} = P(X_n = k)$.

Étape 2 : Convergence de chacune des probabilités ponctuelles.
*Pour tout $k \in \mathbb{N}$ **fixé**, on étudie la convergence **en** n de la suite de réels $(p_{n,k})$. On note p_k sa limite si elle existe.*

Étape 3 : La suite $(p_k)_{k\in\mathbb{N}}$ est-elle une loi de probabilité ?
On vérifie si $(p_k)_{k\in\mathbb{N}}$ définit une loi de probabilité, c'est-à-dire :

* $\forall k \in \mathbb{N}$, $p_k \geq 0$;
* $\sum_{k\geq 0} p_k$ *converge absolument et sa somme vaut 1.*

Si ces conditions sont bien vérifiées, alors (X_n) converge en loi vers toute variable aléatoire discrète X de loi $P(X = k) = p_k$ pour tout $k \in \mathbb{N}$.

Exemple. Soit $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires telles que, pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{P}\left(\frac{1}{n}\right)$.
Montrer que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers la variable certaine égale à 0.

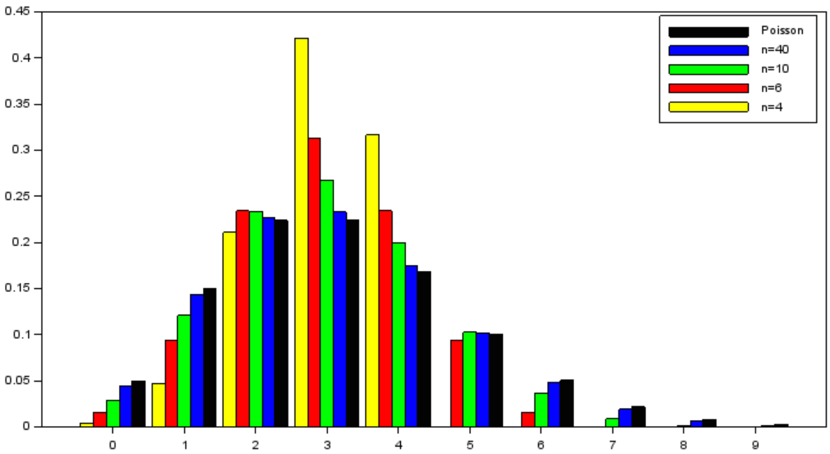
Propriété 8 (Convergence en loi de lois binomiales vers la loi de Poisson)

Soit $\lambda > 0$. On suppose que pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}\left(n, \frac{\lambda}{n}\right)$.

Alors $X_n \xrightarrow{\mathcal{L}} X$ où X est une variable aléatoire suivant une loi de Poisson $\mathcal{P}(\lambda)$.

Preuve.

□



Convergence en loi de $\mathcal{B}(n, 3/n)$ *vers* $\mathcal{P}(3)$.

2.3 Théorème limite central

Vocabulaire. Une suite $(X_n)_{n\in\mathbb{N}}$ de variables est dite **indépendante identiquement distribuées** (en abrégé i.i.d.) lorsque les variables X_n sont mutuellement indépendantes et toutes de même loi.

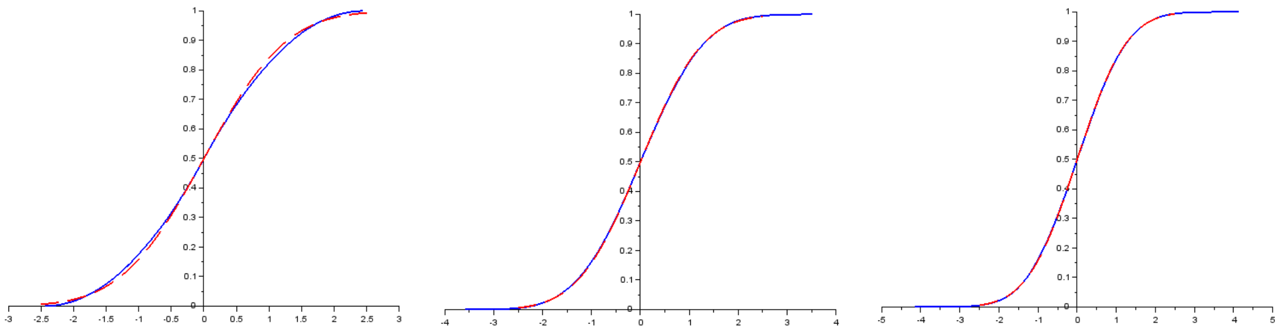
Soit (X_n) une suite de variables i.i.d. d'espérance m et de variance $\sigma^2 \neq 0$. Posons $\overline{X_n} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$. On a :

$$E(\overline{X_n}) \overset{\text{lin. de E}}{=} \frac{1}{n}\sum_{k=1}^{n} E(X_k) = m \quad \text{et} \quad V(\overline{X_n}) \overset{\text{ind. des } X_k}{=} \frac{1}{n^2}\sum_{k=1}^{n} V(X_k) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n},$$

de sorte que la variable

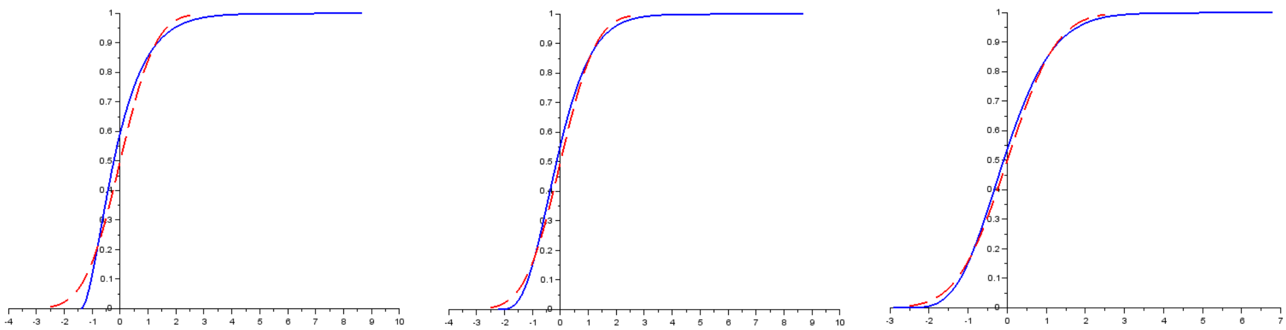
$$\overline{X_n}^* = \frac{\overline{X_n} - E(\overline{X_n})}{\sqrt{V(\overline{X_n})}} = \frac{\overline{X_n} - m}{\frac{\sigma}{\sqrt{n}}} = \sqrt{n}\left(\frac{\overline{X_n} - m}{\sigma}\right)$$

est centrée réduite. Représentons la fonction de répartition de $\overline{X_n}^*$ pour différentes lois et différentes valeurs de $n \in \mathbb{N}^*$ en bleu, ainsi que la fonction de répartition de la loi $\mathcal{N}(0,1)$ en pointillés rouges.



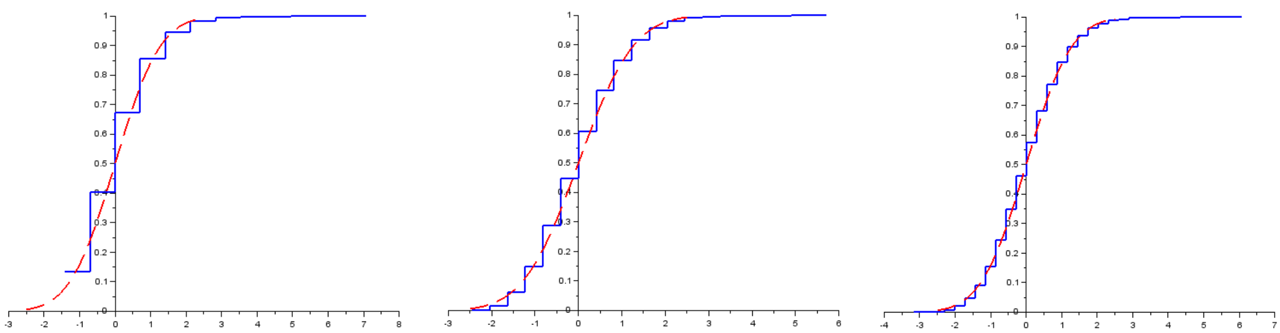
$n = 2$ $n = 6$ $n = 12$

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



$n = 2$ $n = 6$ $n = 12$

Loi exponentielle $\mathcal{E}(1)$.



$n = 2$ $n = 6$ $n = 12$

Loi de Poisson $\mathcal{P}(1)$.

On constate graphiquement la convergence de la fonction de répartition de $\overline{X_n}^*$ vers celle de la loi $\mathcal{N}(0,1)$, et donc la convergence en loi de $\overline{X_n}^*$ vers une variable suivant une loi $\mathcal{N}(0,1)$, et ceci indépendamment de la loi suivie.

On a en fait le résultat suivant :

Théorème 9 (Théorème limite central)

On suppose que :

- Les variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ sont i.i.d.
- Elles admettent une espérance m et une variance σ^2 non nulle.

Pour tout $n \in \mathbb{N}^*$, on note $\overline{X_n}^*$ la variable centrée réduite associée :

$$\overline{X_n}^* = \frac{\overline{X_n} - E(\overline{X_n})}{\sqrt{V(\overline{X_n})}} = \sqrt{n}\left(\frac{\overline{X_n} - m}{\sigma}\right).$$

Alors $\overline{X_n}^*$ converge en loi vers une variable X suivant une loi normale centrée réduite $\mathcal{N}(0,1)$.

Remarques.

- Pour tout (a,b) tel que $-\infty \le a \le b \le +\infty$, on a ainsi :
$$\lim_{n\to+\infty} P(a < \overline{X_n}^* \le b) \overset{\text{prop. } 6}{=} P(a < X \le b) = \frac{1}{\sqrt{2\pi}}\int_a^b \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t = \Phi(b) - \Phi(a)$$
où $\Phi(b) = 1$ si $b = +\infty$ et $\Phi(a) = 0$ si $a = -\infty$.
- Soit $S_n = \sum_{k=1}^{n} X_k$ et S_n^* la variable centrée réduite associée. Alors
$$E(S_n) \overset{\text{lin. de } E}{=} \sum_{k=1}^{n} E(X_k) = nm \quad \text{et} \quad V(S_n) \overset{\text{ind. des } X_k}{=} \sum_{k=1}^{n} V(X_k) = n\sigma^2$$
et on a :
$$S_n^* = \frac{S_n - nm}{\sqrt{n}\sigma} = \frac{n\overline{X_n} - nm}{\sqrt{n}\sigma} = \sqrt{n}\left(\frac{\overline{X_n} - m}{\sigma}\right) = \overline{X_n}^*.$$
Par le théorème limite central, S_n^* converge aussi en loi vers une variable suivant une loi $\mathcal{N}(0,1)$.
- Nous avons utilisé ce théorème en informatique pour simuler une loi normale centrée réduite à partir d'une loi uniforme $\mathcal{U}([0,1])$ (voir **TP 9 - Exercice 1**). Plus précisément, nous avons considéré des variables aléatoires $X_k \hookrightarrow \mathcal{U}([0,1])$ et nous avons utilisé
$$\overline{X_{12}}^* = \sqrt{12}\left(\frac{\overline{X_{12}} - m}{\sigma}\right) = \sqrt{12}\left(\frac{\frac{1}{12}\left(\sum_{k=1}^{12} X_k\right) - \frac{1}{2}}{\sqrt{\frac{1}{12}}}\right) = \left(\sum_{k=1}^{12} X_k\right) - 6$$
pour simuler une loi $\mathcal{N}(0,1)$.
On constate graphiquement que cette approximation semble pertinente, puisque les graphes des fonctions de répartition de $\overline{X_{12}}^*$ et de la loi $\mathcal{N}(0,1)$ semblent confondus.

Interprétation du théorème limite central.
D'après la loi faible des grands nombres, on a $\overline{X_n} \xrightarrow{P} m$. Le théorème limite central précise les fluctuations de $\overline{X_n}$ autour de m, puisque lorsque n est assez grand, on a :

$$\overline{X_n} - m \approx X \quad \text{avec} \quad X \hookrightarrow \mathcal{N}\left(0, \frac{\sigma^2}{n}\right).$$

En particulier, on a :

$$X_1 + \cdots + X_n = n\overline{X_n} \approx S \quad \text{avec} \quad S \hookrightarrow \mathcal{N}(nm, n\sigma^2) \quad \text{pour } n \text{ grand.}$$

Ainsi si on somme des variables aléatoires i.i.d. un grand nombre de fois, la variable aléatoire obtenue suit "approximativement" une loi normale. C'est ce résultat qui justifie l'omniprésence de cette loi dans la nature, de nombreux phénomènes étant dus à l'addition d'un grand nombre de petites perturbations aléatoires.

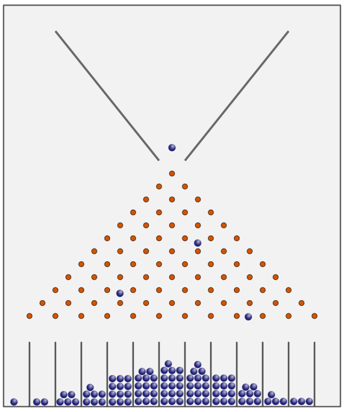
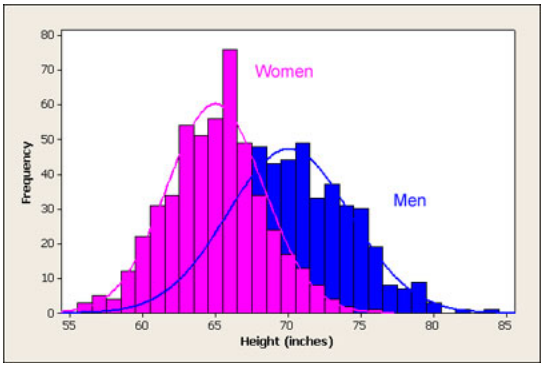


Planche de Galton

Illustration. On plante des clous sur la partie supérieure d'une planche, de telle sorte qu'une bille lâchée sur la planche passe soit à droite soit à gauche pour chaque rangée de clous. Dans la partie inférieure les billes sont rassemblées en fonction du nombre de passages à gauche et à droite qu'elles ont fait. Chaque case correspond donc à un résultat possible d'une expérience binomiale (répétition d'épreuves de Bernoulli identiques et indépendantes).

Voyons en vidéo comment les billes se répartissent dans les cases. On remarque que la répartition des billes approche la forme d'une courbe de Gauss (et ceci serait d'autant plus vrai si le nombre de rangées de clous augmentait). On observe donc ici que la loi binomiale, somme de variables de Bernoulli i.i.d., converge bien (en loi) vers une loi normale.

Exemple. La taille d'un être humain est codée génétiquement par une centaine de gènes, chacun d'eux pouvant prendre des formes variées appelées allèles (généralement un ou deux allèles pour chaque gène). Elle résulte donc d'une succession d'un grand nombre de processus aléatoires, et suit approximativement une loi normale.



3 Approximations

Objectif. Rappelons que si $X_n \xrightarrow{\mathcal{L}} X$, alors X_n suit approximativement la loi de X lorsque n est "assez grand". En pratique pour simplifier un problème ou un calcul, il peut être avantageux de considérer la loi de X plutôt que celle de X_n. On donne dans la suite des conditions empiriques sur les paramètres des lois, notamment sur n, pour que cette approximation soit satisfaisante. Ces conditions ne sont pas à apprendre par coeur, le programme officiel précisant :

Les conditions sous lesquelles l'approximation est valable doivent systématiquement être précisées.

3.1 Approximation de la loi binomiale par la loi de Poisson

Rappelons que si pour tout $n \in \mathbb{N}^*$, $X_n \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})$, alors $X_n \xrightarrow{\mathcal{L}} X$ où $X \hookrightarrow \mathcal{P}(\lambda)$.

Propriété 10 (Approximation de la loi binomiale par la loi de Poisson)

Lorsque $n \geq 30$, $p \leq 0,1$ et $np \leq 15$, on approche la loi $\mathcal{B}(n,p)$ par la loi $\mathcal{P}(np)$.

Remarque. Retenir pour cette approximation qu'il faut "n grand" et "p petit". Remarquons pour mémoire que ces deux lois ont même espérance.

Exemple. Prenons l'exemple d'une barre d'uranium. Elle contient une très grande quantité n d'atomes radioactifs. Chacun à une probabilité p très faible de se désintégrer durant un laps de temps T. On observe de plus qu'en moyenne il y a λ désintégrations durant ce laps de temps T :

$$\lambda = n \times p.$$

En supposant l'indépendance des désintégrations atomiques, il serait rigoureux de modéliser le nombre X de désintégration par une loi binomiale $\mathcal{B}(n, p)$. En pratique, les calculs numériques avec cette loi sont rendu impossible du fait que n est très grand et inconnu, p très petit. On approxime donc la loi de X par une loi de Poisson de paramètre λ, approximation satisfaisante d'après la propriété précédente.

3.2 Approximation de la loi binomiale par la loi normale

Soit (X_n) une suite de variables aléatoires i.i.d. suivant toutes une loi de Bernoulli $\mathcal{B}(p)$ avec $0 < p < 1$. Posons pour tout $n \in \mathbb{N}^*$, $S_n = \displaystyle\sum_{i=1}^{n} X_i$. Alors $S_n \hookrightarrow \mathcal{B}(n, p)$ et :

$$S_n^* = \frac{S_n - np}{\sqrt{np(1-p)}}$$

converge en loi vers (une variable aléatoire suivant une loi) $\mathcal{N}(0, 1)$ d'après le théorème limite central. Ainsi pour n suffisamment grand,

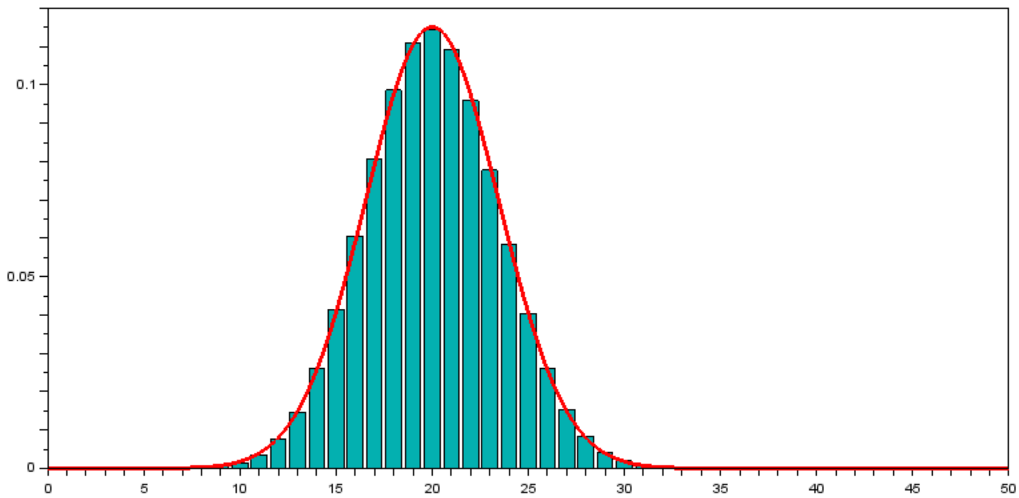
$$S_n = \sqrt{np(1-p)}S_n^* + np$$

suit approximativement une loi $\mathcal{N}(np, np(1-p))$.

Propriété 11 (Approximation de la loi binomiale par la loi normale)

Lorsque $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$, on approche la loi $\mathcal{B}(n, p)$ par la loi $\mathcal{N}(np, np(1-p))$.

Remarque. Retenir qu'il faut avoir "n grand" et "p ni trop petit, ni trop grand". Remarquons pour mémoire que ces deux lois ont même espérance et même variance.



Approximation de la loi $\mathcal{B}(50, 0.4)$ par la loi $\mathcal{N}(20, 12)$.

Correction de continuité. On approche dans le résultat précédent une variable X suivant une loi discrète $\mathcal{B}(n, p)$ par une variable Y suivant une loi continue $\mathcal{N}(np, np(1-p))$. Mais cette approximation pose une question : comment calculer les $P(X = k)$? On ne peut pas écrire :

$$P(X = k) \approx P(Y = k)$$

puisque $P(Y = k) = 0$. La règle communément acceptée est d'utiliser que :

$$P(X = k) = P(k - 0.5 < X \leq k + 0.5)$$

et donc d'approcher $P(X = k)$ par $P(k - 0.5 \leq Y \leq k + 0.5)$. De plus, afin de garantir que la somme des probabilités reste égale à 1, on remplacera $P(X = 0)$ par $P(Y \leq 0.5)$ et $P(X = n)$ par $P(n - 0.5 \leq Y)$. On parle de **correction de continuité**.

Exercice. Soit $X \hookrightarrow \mathcal{B}(100, 0.4)$. Calculer approximativement $P(X = 35)$.

Un calcul direct donne $P(X = 35) = 0.0491$.

3.3 Approximation de la loi de Poisson par la loi normale

Soit (X_n) une suite de variables aléatoires i.i.d. suivant toutes une loi de Poisson $\mathcal{P}(\lambda)$.

Posons pour tout $n \in \mathbb{N}^*$, $S_n = \sum_{i=1}^{n} X_i$. Alors $S_n \hookrightarrow \mathcal{P}(n\lambda)$ et :

$$S_n^* = \frac{S_n - n\lambda}{\sqrt{n\lambda}}$$

converge en loi vers $\mathcal{N}(0, 1)$. Pour n "suffisamment grand",

$$S_n = \sqrt{n\lambda} S_n^* + n\lambda$$

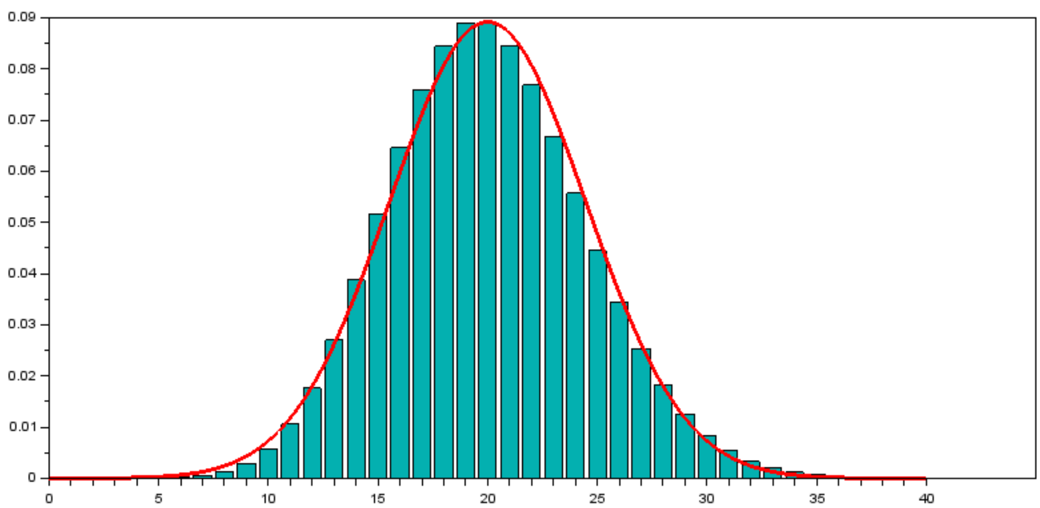
suit donc approximativement une loi $\mathcal{N}(n\lambda, n\lambda)$.

Propriété 12 (Approximation de la loi de Poisson par la loi normale)

Pour $\lambda \geq 15$, on approche la loi $\mathcal{P}(\lambda)$ par la loi $\mathcal{N}(\lambda, \lambda)$.

Remarques.

- Retenir qu'il faut avoir "λ assez grand". Remarquons pour mémoire que ces deux lois ont même espérance et même variance.
- On approche ici aussi une variable aléatoire discrète par une variable aléatoire continue. On utilisera donc une correction de continuité.



Approximation de la loi $\mathcal{P}(20)$ par la loi $\mathcal{N}(20, 20)$.