CONVERGENCE ET APPROXIMATION

P mentionne des résultats particulièrement utiles et souvent oubliés dans la pratique de la convergence.

★ mentionne des erreurs à ne pas faire où des hypothèses importantes ou des mises en garde.

SD mentionne des résultats qu'il serait bon de savoir démontrer.

! Notions ou résultats qui ne semblent pas toujours très importants mais qui figurent explicitement dans le programme donc qui sont exigibles.

Dans la suite les variables aléatoires considérées le sont sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$ (sauf mention du contraire).

I CONVERGENCE EN PROBABILITÉ

▶ 1. Définition.

Déf. 1 $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On dit que $(X_n)_{n\geqslant n_0}$ **converge en probabilité** vers X si pour tout réel ε strictement positif :

$$\lim_{n\to+\infty} P(\{|X_n - X| \geqslant \varepsilon\}) = 0.$$

Nous écrirons alors $X_n \underset{P}{\longrightarrow} X$.

Th. 1 $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

$(X_n)_{n\geqslant n_0}$ converge en probabilité vers X si et seulement si, pour tout réel ε strictement positif :

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

★★ Ce résultat est en fait la définition du programme pour la convergence en probabilité. Qu'on se le dise et que l'on se le démontre.

★★ On peut encore remplacer $\lim\limits_{n\to+\infty} P(\{|X_n - X| \geqslant \varepsilon\}) = 0$ ou $\lim\limits_{n\to+\infty} P(\{|X_n - X| > \varepsilon\}) = 0$ par $\lim\limits_{n\to+\infty} P(\{|X_n - X| < \varepsilon\}) = 1$ ou $\lim\limits_{n\to+\infty} P(\{|X_n - X| \leqslant \varepsilon\}) = 1$.

★★ Notons que la convergence en probabilité est une notion relativement contraignante. Il n'est pas toujours facile de calculer la probabilité intervenant d'autant qu'elle ne résulte pas de manière immédiate des lois de X_n et de X puisqu'elle fait intervenir la loi de $X_n - X$. Les deux conditions suffisantes qui suivent sont donc les bienvenues. La convergence en loi également...

▶ 2. Une condition suffisante de convergence en probabilité.

Th. 2 *Complément* **SD** Soit $(X_n)_{n\geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

Si $\lim\limits_{n\to+\infty} E(X_n - X) = 0$ et $\lim\limits_{n\to+\infty} V(X_n - X) = 0$ alors la suite $(X_n)_{n\geqslant n_0}$ converge en probabilité vers X.

Th. 3 **PP** **SD** Soit $(X_n)_{n\geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On suppose que, pour n assez grand, $E(X - X_n) = 0$ (ou que $E(X_n) = E(X)$, ce qui n'est pas tout à fait la même chose...)

Alors si $\lim\limits_{n\to+\infty} V(X_n - X) = 0$, la suite $(X_n)_{n\geqslant n_0}$ converge en probabilité vers X.

★★ Le second théorème est de toute évidence un corollaire du premier au détail près que sa démonstration peut s'obtenir en deux lignes avec l'inégalité de Bienaymé-Tchebychev. Qu'on se le dise, qu'on se le démontre, que l'on se l'utilise et que l'on ne se l'oublie pas au niveau des estimateurs...

★★ Les conditions contenues dans ces théorèmes sont suffisantes pour avoir la convergence en probabilité mais pas nécessaires.

▶ 3. Loi faible des grands nombres.

Th. 4 $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles suivant la même loi ayant une espérance m et une variance σ^2 ($\sigma > 0$).

On suppose que les variables aléatoires réelles de cette suite sont deux à deux indépendantes.

Alors la suite de terme général $\dfrac{X_1 + X_2 + \cdots + X_n}{n}$ converge en probabilité vers la variable aléatoire réelle certaine égale à m.

★ Notons que sous les hypothèses du théorème précédent : $P\left(\left|\dfrac{X_1 + X_2 + \cdots + X_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \dfrac{\sigma^2}{n\varepsilon^2}$.

Cor. $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles suivant une loi de Bernoulli de paramètre p.

On suppose que les variables aléatoires réelles de cette suite sont deux à deux indépendantes (resp. indépendantes).

La suite de terme général $\dfrac{X_1 + X_2 + \cdots + X_n}{n}$ converge en probabilité vers la variable aléatoire réelle certaine égale à p.

★ Notons que sous les hypothèses du résultat précédent : $P\left(\left|\dfrac{X_1 + X_2 + \cdots + X_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \dfrac{p(1-p)}{n\varepsilon^2} \leqslant \dfrac{1}{4n\varepsilon^2}$.

Déf. 2 $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$.

La variable aléatoire $\dfrac{X_1 + X_2 + \cdots + X_n}{n}$ est la **moyenne empirique** d'ordre n de la suite.

▶ 4. Une conséquence de la loi faible des grands nombres.

Illustrons le dernier résultat. Considérons un événement A associé à une expérience aléatoire et notons p la probabilité de sa réalisation.

Itérons cette expérience aléatoire de manière à ce que les itérations soient indépendantes. Pour tout n dans $\mathbb{N}^*$, notons X_n la variable aléatoire qui vaut 1 si A se réalise à la $n^{\text{ème}}$ itération et 0 sinon. Pour tout n dans $\mathbb{N}^*$, X_n suit une loi de Bernoulli de paramètre p et $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires mutuellement indépendantes.

Pour tout n dans $\mathbb{N}^*$, $Y_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$ est la fréquence de réalisation de A au cours des n premières itérations.

Le résultat précédent indique que $(Y_n)_{n\geqslant 1}$ converge en probabilité vers la variable aléatoire certaine égale à p. Donc, pour tout réel strictement positif ε, la probabilité pour que Y_n prenne des valeurs à l'extérieur de l'intervalle $]-\varepsilon, \varepsilon[$ tend vers zéro lorsque n tend vers l'infini , c'est à dire lorsque l'on réalise “un grand nombre” de fois l'expérience.

Ceci légitime notre modèle probabiliste inspiré des fréquences statistiques.

A l'inverse il permet encore d'estimer la valeur de p à partir de la fréquence de réalisation de A.

II CONVERGENCE EN LOI

▶ 1. Définition.

Déf. 3 $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On dit que $(X_n)_{n\geqslant n_0}$ **converge en loi** vers X si pour tout réel x où F_X est continue :

$$\lim_{n\to+\infty} F_{X_n}(x) = F_X(x) \quad \text{ou} \quad \lim_{n\to+\infty} P(X_n \leqslant x) = P(X \leqslant x).$$

▶ 2. Pratiquement.

PP Soit $(X_n)_{n\geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$. Pour montrer que cette suite converge en loi il n'est pas utile de trouver une variable aléatoire X telle que $(X_n)_{n\geqslant n_0}$ converge en loi vers X.

En gros, il convient avant tout de trouver $\lim\limits_{n\to+\infty} F_{X_n}(x)$, pour tout x dans $\mathbb{R}$ et de montrer que l'application F : $x \to \lim\limits_{n\to+\infty} F_{X_n}(x)$ est la fonction de répartition d'une variable aléatoire. Alors deux rappels s'imposent.

Th. 5 Soit F une application de $\mathbb{R}$ dans $[0, 1]$ ou dans $\mathbb{R}$... (voir les deux premiers points).

F est la fonction de répartition d'une variable aléatoire réelle si et seulement si :

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

Th. 6 Soit F une application de $\mathbb{R}$ dans $[0,1]$ ou dans $\mathbb{R}$... (voir les deux premiers points).

F est la fonction de répartition d'une variable aléatoire réelle à densité si et seulement si :

1. F est croissante sur $\mathbb{R}$;

2. $\lim\limits_{x\to-\infty} F(x)=0$ et $\lim\limits_{x\to+\infty} F(x)=1$;

3. F est continue sur $\mathbb{R}$;

4. Il existe un ensemble fini éventuellement vide D, contenu dans $\mathbb{R}$ et tel que F soit de classe $\mathcal{C}^1$ sur $\mathbb{R}$ - D.

▶ 3. Comparaison des deux convergences.

Th. 7 $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega,\mathcal{A},P)$ et X une variable aléatoire réelle sur $(\Omega,\mathcal{A},P)$.

Si $(X_n)_{n\geqslant n_0}$ converge en probabilité vers X alors $(X_n)_{n\geqslant n_0}$ converge en loi vers X.

La réciproque est fausse.

▶ 4. Une difficulté à surmonter.

$\boxed{\textbf{PP}}$ Soit $(X_n)_{n\geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega,\mathcal{A},P)$. On souhaite montrer que cette suite converge en loi vers une variable aléatoire. Comme nous l'avons dit plus haut il convient en gros de trouver $\lim\limits_{n\to+\infty} F_{X_n}(x)$ pour x dans $\mathbb{R}$. Le problème se complique lorsque F_{X_n} est définie par intervalles et que des bornes de ses intervalles contiennent n. Donnons deux exemples.

ex. 1. $\forall x\in\mathbb{R},\ F_n(x)=\begin{cases}\left(1-\dfrac{e^{-x}}{n}\right)^n & \text{si } x\in[-\ln n,+\infty[\\ 0 & \text{sinon}\end{cases}.$

ex. 2. $\forall x\in\left]-\infty,\dfrac{2}{n}\right],\ F_n(x)=0$ et $\forall x\in[2,+\infty[,\ F_n(x)=1.$

Si x est élément de $\left[\dfrac{2}{n},1+\dfrac{1}{n}\right[: F_n(x)=\dfrac{\text{Ent}\,(nx)\ (\text{Ent}\,(nx)-1)}{2n^2}.$

Si x est élément de $\left[1+\dfrac{1}{n},2\right[: F_n(x)=\dfrac{1}{2\,n^2}\Big((4n+1)\,\text{Ent}\,(nx)-\big(\text{Ent}\,(nx)\big)^2-2\,n^2-2\,n\Big).$

Dans le premier cas il n'est pas question de passer à la limite pour x dans $[-\ln n,+\infty[$ puis pour x dans $]-\infty,-\ln n[$! ! Même type de remarque pour le deuxième exemple.

Le bon algorithme consiste à fixer x dans un intervalle indépendant de n qui permet de passer aisément à la limite.

Dans le premier cas on fixe x dans $\mathbb{R}$ et on remarque que pour n assez grand x est dans $[-\ln n,+\infty[$.

Dans le second cas on conclut aisément pour x dans $]-\infty,0]$ et dans $[2,+\infty[$.

Puis on fixe x dans $]0,1]$ et on remarque que pour n assez grand x est dans $\left[\frac{2}{n},1+\frac{1}{n}\right[$.

Puis on fixe x dans $]1,2[$ et on remarque que pour n assez grand x est dans $\left[1+\frac{1}{n},2\right[$.

III CONVERGENCE EN LOI : DU DISCRET AU DISCRET

▶ 1. Quelques caractérisations.

Th. 8 $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On suppose que toutes ses variables aléatoires prennent leurs valeurs dans $\mathbb{N}$.

$(X_n)_{n\geqslant n_0}$ converge en loi vers X si et seulement si pour tout élément k de $\mathbb{N}$:

$$\lim_{n\to+\infty} P(X_n = k) = P(X = k).$$

Ceci est le résultat du programme. C'est bien maigre et impose sans doute de savoir démontrer les quelques résultats complémentaires qui suivent.

Th. 9 *Complément* $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On suppose que toutes ses variables aléatoires prennent leurs valeurs dans un ensemble fini $\{x_1, x_2, \ldots, x_r\}$.

$(X_n)_{n\geqslant n_0}$ converge en loi vers X si et seulement si pour tout élément k de $[\![1, r]\!]$:

$$\lim_{n\to+\infty} P(X_n = x_k) = P(X = x_k).$$

Cor. *Complément* $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On suppose que toutes ses variables aléatoires prennent leurs valeurs dans $[\![a, b]\!]$ (a et b sont deux éléments de $\mathbb{Z}$ tels que $a \leqslant b$).

$(X_n)_{n\geqslant n_0}$ converge en loi vers X si et seulement si pour tout élément k de $[\![a, b]\!]$:

$$\lim_{n\to+\infty} P(X_n = k) = P(X = k).$$

Th. 10 *Complément* $(x_k)_{k\geqslant 0}$ est une suite strictement croissante de réels. $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

On suppose que toutes ses variables aléatoires prennent leurs valeurs dans l'ensemble dénombrable $\{x_k \mid k \in \mathbb{N}\}$.

$(X_n)_{n\geqslant n_0}$ converge en loi vers X si et seulement si pour tout élément k de $\mathbb{N}$:

$$\lim_{n\to+\infty} P(X_n = x_k) = P(X = x_k).$$

Cor. *Complément* $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega,\mathcal{A},P)$ et X une variable aléatoire réelle sur $(\Omega,\mathcal{A},P)$.

On suppose que toutes ses variables aléatoires prennent leurs valeurs dans $[\![r,+\infty[\![$ (r est un élément de $\mathbb{Z}$).

$(X_n)_{n\geqslant n_0}$ converge en loi vers X si et seulement si pour tout élément k de $[\![r,+\infty[\![$:

$$\lim_{n\to+\infty} P(X_n=k)=P(X=k).$$

▶ 2. Pratiquement.

PP Soit $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires discrètes réelles sur $(\Omega,\mathcal{A},P)$. Pour montrer que cette suite converge en loi il n'est pas utile de trouver une variable aléatoire X telle que $(X_n)_{n\geqslant n_0}$ converge en loi vers X.

Deux voies sont à explorer car “la limite” peut être discrète ou non discrète. Dans le second cas revenir à ce qui a été dit plus haut.

Dans le premier cas il convient, en gros, de trouver une partie finie ou dénombrable D de $\mathbb{R}$ telle que, pour tout élément d de D la suite de terme général $P(X_n=d)$ converge et de montrer que $d\to \lim\limits_{n\to+\infty} P(X_n=d)$ est une loi de probabilité discrète. Ici encore un rappel s'impose.

Déf. 4 On appelle **loi de probabilité discrète** toute application d'une partie finie ou dénombrable D de $\mathbb{R}$ dans $[0,1]$ ou dans $[0,+\infty[$ telle que :

$$\sum_{d\in D} f(d)=1$$

Prop. 1 Toute loi de probabilité discrète est la loi de probabilité d'une variable aléatoire réelle discrète .

▶ 3. Approximation 1. Approximation d'une loi hypergéométrique par une loi binômiale.

Th. 11 p est un élément de $]0,1[$, n est un élément de $\mathbb{N}^*$ et $(N_m)_{m\geqslant 0}$ est une suite d'éléments de $[\![n,+\infty[\![$ telle que :

$$\lim_{m\to+\infty} N_m=+\infty \quad \text{et} \quad \forall m\in\mathbb{N},\ N_m p\in\mathbb{N}.$$

On considère une suite $(X_m)_{m\geqslant 0}$ de variables aléatoires réelles sur $(\Omega,\mathcal{A},P)$ telle que pour tout m appartenant à $\mathbb{N}$, $X_m \hookrightarrow \mathcal{H}(N_m,n,p)$.

$(X_m)_{m\geqslant 0}$ converge en loi vers une variable aléatoire réelle suivant une loi binômiale de paramètres n et p.

Autrement dit pour tout élément k de $[\![0,n]\!]$: $\lim\limits_{m\to+\infty} P(\{X_m=k\})=\binom{n}{k}p^k(1-p)^{n-k}$.

PP Ce résultat théorique conduit dans la pratique à approximer une variable aléatoire réelle X suivant une loi hypergéométrique de paramètres N, n et p, par une variable aléatoire réelle suivant une loi binômiale de paramètres n et p lorsque $\boxed{N \text{ est sensiblement supérieur à } 10\,n}$.

Si k est un élément de $X(\Omega)$, on prend $\binom{n}{k}p^k(1-p)^{n-k}$ comme valeur approchée de $P(\{X=k\})=\dfrac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}$.

Ceci a en outre comme avantage de faire passer de trois paramètres à deux.

▶ 4. Approximation 2. Approximation d'une loi binômiale par une loi de Poisson.

Th. 12 **!** λ est un réel strictement positif. Pour tout élément n de $\mathbb{N}^*$, X_n est une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$ qui suit une loi binômiale de paramètres n et $\dfrac{\lambda}{n}\cdot$

Alors la suite $(X_n)_{n\geqslant 1}$ converge en loi vers une variable aléatoire réelle suivant une loi de Poisson de paramètre λ. Autrement dit pour tout élément k de $\mathbb{N}$:

$$\lim_{n\to+\infty} P(\{X_n = k\}) = \frac{\lambda^k}{k!}\, e^{-\lambda}.$$

Th. 13 *Complément* Pour tout élément n de $[\![n_0, +\infty [\![$, p_n est un élément de $[0,1]$ et X_n une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$ qui suit une loi binômiale de paramètres n et p_n.

On suppose de plus que la suite $(np_n)_{n\geqslant n_0}$ converge vers un réel strictement positif λ.

Alors la suite $(X_n)_{n\geqslant n_0}$ converge en loi vers une variable aléatoire réelle suivant une loi de Poisson de paramètre λ. Autrement dit pour tout élément k de $\mathbb{N}$:

$$\lim_{n\to+\infty} P(\{X_n = k\}) = \frac{\lambda^k}{k!}\, e^{-\lambda}.$$

PP Ce résultat théorique conduit dans la pratique à approximer une variable aléatoire réelle X suivant une loi binômiale de paramètres n et p par une variable aléatoire réelle suivant une loi de Poisson de paramètre np lorsque $n \geqslant 30$, $p \leqslant 0,1$ et $np \leqslant 15$ (attention il y a autant de conditions que d'auteurs, mais voir plus bas...)

Si k est un élément de $X(\Omega)$, on prend $\dfrac{(np)^k}{k!}\, e^{-np}$ comme valeur approchée de $P(\{X = k\})$.

Ceci a, en outre, comme avantage de faire passer de deux paramètres à un.

▶ 5. Une remarque du programme à propos des approximations.

★★ Le programme dit que : "toutes indications devront être fournies aux candidats quant à la justification des approximations"

IV THÉORÈME DE LA LIMITE CENTRÉE

▶ 1. Deux énoncés du théorème de la limite centrée.

Th. 14 $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles **mutuellement indépendantes** suivant la même loi ayant une espérance m et une variance σ^2 ($\sigma > 0$)

On pose pour tout élément n de $\mathbb{N}^*$, $S_n = X_1 + X_2 + \cdots + X_n$.

La suite de terme général $S_n^* = \dfrac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \dfrac{S_n - E(S_n)}{\sigma(S_n)} = \dfrac{S_n - nm}{\sqrt{n}\sigma}$ converge en loi vers une variable aléatoire réelle suivant une loi normale centrée réduite.

Ainsi, pour tout réel x : $\lim\limits_{n\to+\infty} F_{S_n^*}(x) = \lim\limits_{n\to+\infty} P(S_n^* \leqslant x) = \Phi(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}\, \mathrm{d}t$.

Th. 15 $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles **mutuellement indépendantes** suivant la même loi ayant une espérance m et une variance σ^2 ($\sigma > 0$)

On pose pour tout élément n de $\mathbb{N}^*$, $F_n = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$.

La suite de terme général $F_n^* = \dfrac{F_n - E(F_n)}{\sqrt{V(F_n)}} = \dfrac{F_n - E(F_n)}{\sigma(F_n)} = \dfrac{F_n - m}{\frac{\sigma}{\sqrt{n}}}$ converge en loi vers une variable aléatoire réelle suivant une loi normale centrée réduite.

Ainsi, pour tout réel x : $\lim\limits_{n\to+\infty} F_{F_n^*}(x) = \lim\limits_{n\to+\infty} P(F_n^* \leqslant x) = \Phi(x) = \displaystyle\int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}\, \mathrm{d}t$.

▶ 2. Approximation 3. Approximation d'une loi binômiale par une loi normale.

Th. 16 $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles de bernoulli de paramètre p ($p \in]0,1[$) **mutuellement indépendantes**. $q = 1 - p$.

On pose pour tout élément n de $\mathbb{N}^*$, $S_n = X_1 + X_2 + \cdots + X_n$.

a) Pour tout élément n de $\mathbb{N}^*$, S_n suit une loi binômiale de paramètres n et p.

b) La suite de terme général $S_n^* = \dfrac{S_n - np}{\sqrt{npq}}$ converge en loi vers une variable aléatoire réelle suivant une loi normale centrée réduite. Ainsi, pour tout réel x :

$$\lim_{n\to+\infty} P\left(\frac{S_n - np}{\sqrt{npq}} \leqslant x\right) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}\, \mathrm{d}t.$$

Cor. p est un élément de $]0,1[$. $(T_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ telle que pour tout n dans $\mathbb{N}^*$, T_n suit une loi binômiale de paramètres n et p.

La suite de terme général $T_n^* = \dfrac{T_n - np}{\sqrt{npq}}$ converge en loi vers une variable aléatoire réelle suivant une loi normale centrée réduite. Ainsi, pour tout réel x :

$$\lim_{n\to+\infty} P\left(\frac{T_n - np}{\sqrt{npq}} \leqslant x\right) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}\, \mathrm{d}t.$$

★★ Observons que le corollaire ne contient pas de condition d'indépendance. Il convient de savoir le démontrer proprement à partir du théorème.

PP Ce résultat théorique conduit dans la pratique à approximer une variable aléatoire réelle X suivant une loi binômiale de paramètres n et p par une variable aléatoire réelle Y suivant une loi normale d'espérance np et d'écart-type $\sqrt{npq}$ ($q = 1 - p$) lorsque $n \geqslant 20$ ou 30, p pas trop petit (!), $np \geqslant 10$ et $nq \geqslant 10$.

Dans ces conditions

- Pour tout réel x, on approxime $P(X \leqslant x)$ par $P(Y \leqslant x) = F_Y(x)$; donc

$$\boxed{P(X \leqslant x) \simeq \Phi\Big(\frac{x - np}{\sqrt{npq}}\Big)}$$

- Pour k dans $[\![0, n]\!]$, on approxime $P(X = k)$ par $P(k - 1/2 \leqslant Y \leqslant k + 1/2) = F_Y(k + 1/2) - F_Y(k - 1/2)$; donc

$$\boxed{P(X = k) \simeq \Phi\Big(\frac{k + 1/2 - np}{\sqrt{npq}}\Big) - \Phi\Big(\frac{k - 1/2 - np}{\sqrt{npq}}\Big)}$$

▶ 3. Approximation 4. Approximation d'une loi de Poisson par une loi normale.

Th. 17 $(X_n)_{n \geqslant 1}$ est une suite de variables aléatoires réelles **mutuellement indépendantes**. On suppose que pour tout élément n de $\mathbb{N}^*$, X_n suit une loi de Poisson de paramètre μ ($\mu > 0$)

a) $S_n = X_1 + X_2 + \cdots + X_n$ suit une loi de Poisson de paramètre $n\,\mu$.

b) La suite de terme général $S_n^* = \dfrac{S_n - n\,\mu}{\sqrt{n\,\mu}}$ converge en loi vers une variable aléatoire réelle suivant une loi normale centrée réduite. Ainsi, pour tout réel x :

$$\lim_{n \to +\infty} P\left(\frac{S_n - n\,\mu}{\sqrt{n\,\mu}} \leqslant x\right) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}\, \mathrm{d}t.$$

Cor. μ est un élément de $]0, +\infty[$. $(T_n)_{n \geqslant 1}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ telle que pour tout n dans $\mathbb{N}^*$, T_n suit une loi de Poisson de paramètre $n\,\mu$.

La suite de terme général $T_n^* = \dfrac{T_n - n\,\mu}{\sqrt{n\,\mu}}$ converge en loi vers une variable aléatoire réelle suivant une loi normale centrée réduite. Ainsi, pour tout réel x :

$$\lim_{n \to +\infty} P\left(\frac{T_n - n\,\mu}{\sqrt{n\,\mu}} \leqslant x\right) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{t^2}{2}}\, \mathrm{d}t.$$

★★ Observons que le corollaire ne contient pas de condition d'indépendance. Il convient de savoir le démontrer proprement à partir du théorème.

PP Ce résultat théorique conduit dans la pratique à approximer une variable aléatoire réelle X suivant une loi de Poisson de paramètre λ par une variable aléatoire réelle Y suivant une loi normale d'espérance λ et d'écart-type $\sqrt{\lambda}$ lorsque $\lambda > 10$.

Dans ces conditions

• Pour tout réel x, on approxime $P(X \leqslant x)$ par $P(Y \leqslant x) = F_Y(x)$; donc

$$\boxed{P(X \leqslant x) \simeq \Phi\Big(\frac{x-\lambda}{\sqrt{\lambda}}\Big)}$$

• Pour tout élément k de $\mathbb{N}$, on approxime $P(X = k)$ par $P(k-1/2 \leqslant Y \leqslant k+1/2) = F_Y(k+1/2) - F_Y(k-1/2)$; donc

$$\boxed{P(X = k) \simeq \Phi\Big(\frac{k+1/2-\lambda}{\sqrt{\lambda}}\Big) - \Phi\Big(\frac{k-1/2-\lambda}{\sqrt{\lambda}}\Big)}$$

V COMPLÉMENTS

▶ 1. De "l'unicité" de la limite dans le convergence en probabilité.

Th. 18 $(X_n)_{n\geqslant n_0}$ est une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$. X et X' deux variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$.

Si la suite $(X_n)_{n\geqslant n_0}$ converge en probabilité vers X et X', alors X et X' sont presque sûrement égales, c'est à dire que $P(X = X') = 1$.

▶ 2. Convergence presque sûre.

Déf. 5 Soit $(X_n)_{n\geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

La suite $(X_n)_{n\geqslant n_0}$ converge presque sûrement vers X s'il existe un événement négligeable A de $\mathcal{A}$ ($P(A) = 0$) tel que :

$$\forall \omega \in \Omega - A, \quad \lim_{n\to+\infty} X_n(\omega) = X(\omega).$$

On écrit alors : $X_n \underset{PS}{\longrightarrow} X$.

Déf. 6 Soit $(X_n)_{n\geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$.

Si la suite $(X_n)_{n\geqslant n_0}$ converge presque sûrement vers X alors la suite $(X_n)_{n\geqslant n_0}$ converge en probabilité vers X.

★★ La réciproque est fausse.

▶ 3. Loi "forte" des grands nombres.

Th. 19 $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires réelles suivant la même loi ayant une espérance m et une variance σ^2 ($\sigma > 0$).

On suppose que les variables aléatoires réelles de cette suite sont deux à deux indépendantes.

Alors la suite de terme général $\dfrac{X_1 + X_2 + \cdots + X_n}{n}$ converge presque sûrement vers la variable aléatoire réelle certaine égale à m.

▶ 4. Convergence en probabilité et opérations.

Prop. 2 **Théorème de Slutsky** Soit $(X_n)_{n \geqslant n_0}$ une suite de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$ et X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, P)$. Soit g une application continue de $\mathbb{R}$ dans $\mathbb{R}$.

Si la suite $(X_n)_{n \geqslant n_0}$ converge en probabilité vers X alors la suite $(g \circ X_n)_{n \geqslant n_0}$ converge en probabilité vers $g \circ X$

★ Facile à prouver pourvu que l'on sache qu'une fonction continue sur un segment y est uniformément continue.

Prop. 3 Soient $(X_n)_{n \geqslant n_0}$ et $(Y_n)_{n \geqslant n_0}$ deux suites de variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$. Soient X et Y deux variables aléatoires réelles sur $(\Omega, \mathcal{A}, P)$. a est un réel.

On suppose que $(X_n)_{n \geqslant n_0}$ et $(Y_n)_{n \geqslant n_0}$ convergent en probabilité respectivement vers X et Y.

Alors les suites de terme généraux $a\,X_n$, $X_n + Y_n$ et $X_n\,Y_n$ convergent en probabilité respectivement vers $a\,X$, $X + Y$ et $X\,Y$.