Master 1 de Mathématiques MME-Mirats et MA
Probabilité 1 Année 2012-2013 cours de M. Gallardo

Chapitre 3 : Variables aléatoires discrètes.

I) Généralités :

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $d \geq 1$ un entier fixé. Considérons une application $X : \Omega \to \mathbb{R}^d$ d'image discrète i.e. l'ensemble $X(\Omega)$ est fini ou dénombrable.

Définition: On dit que X est un vecteur aléatoire (discret) de $\mathbb{R}^d$ (resp. une variable aléatoire quand $d=1$) si :
$$\forall x \in X(\Omega),\ \{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{F}$$

Notation : l'événement $\{\omega \in \Omega \mid X(\omega) = x\}$ des éventualités $\omega \in \Omega$ telles que $X(\omega) = x$, sera noté $[X = x]$.

On dira simplement que $[X=x]$ est l'événement "X prend la valeur x".

Posons $X(\Omega) = \{x_i \,;\, i \in D\}$ ($D = [\![1, N]\!]$ ou $D = \mathbb{N}$) l'ensemble des valeurs prises par X. Alors :

Exercice : Montrer que les événements $[X = x_i]$, $i \in D$, forment un système complet.

Notation : on utilisera l'abréviation v.a. pour vecteur aléatoire ou variable aléatoire si $d=1$.

On continue à noter $X(\Omega) = \{x_i \,;\, i \in D\}$ l'ensemble des valeurs prises par le v.a. X.

Définition: La suite des nombres
$$p_i = \mathbb{P}(X = x_i)\ ,\ i \in D$$
s'appelle la distribution (ou loi) de probabilité du v.a. X.

(on a noté pour simplifier $\mathbb{P}(X = x_i) := \mathbb{P}([X = x_i])$)

Probabilité d'un événement relatif au v.a. X : Les événements à prendre en considération relativement à un v.a. (discret) X sont de la forme

$[X \in A]$ où A est un sous ensemble de $X(\Omega)$.

On calcule leur probabilité de la manière suivante : Soit $J = \{i \in D \,;\, x_i \in A\}$. Alors
$$\mathbb{P}(X \in A) = \sum_{j \in J} p_j .$$

II) Exemples (lois de probabilité classiques, cas $d=1$)

1) Loi de Bernoulli : une v.a. X telle que $X(\Omega) = \{0, 1\}$ s'appelle v.a. de Bernoulli. Sa loi de probabilité est déterminée par la valeur
$$p = \mathbb{P}(X = 1)$$
qui s'appelle le paramètre de la v.a. de Bernoulli X

(Noter, ce qui est évident, que $\mathbb{P}(X=0) = 1 - p$).

<u>Exemple typique</u>: Soit $A \in \mathcal{F}$ un événement. <u>L'indicatrice de A</u> est la fonction :

$$\mathbb{1}_A : \omega \longmapsto \mathbb{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

On voit immédiatement que $X = \mathbb{1}_A$ est une v.a. de Bernoulli de paramètre $p = \mathbb{P}(A)$.

2) <u>Loi binomiale</u> : Considérons une expérience aléatoire $\mathcal{E}$ à 2 issues :

s (= succès) $\quad e$ (= échec),

avec $P(s) = p$ et $P(e) = 1-p$ $(0 < p < 1)$.

Soit $n \geq 1$ un entier fixé. On fait n répétitions indépendantes de cette expérience $\mathcal{E}$, qu'on modélise par l'espace produit

$$\Omega = \{s, e\}^n$$

muni de la probabilité produit $\mathbb{P} = P^{\otimes n}$

> <u>Définition</u>: La v.a. $X :=$ <u>le nombre total de succès</u> au cours des n répétitions de $\mathcal{E}$, est appelée <u>variable aléatoire binomiale de paramètres</u> (n, p).

(Dans ce cas $X(\Omega) = \{0, 1, \dots, n\}$)

> <u>Proposition</u>: La loi de probabilité d'une v.a. binomiale de paramètres (n,p) est donnée par :
>
> $$p_k = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$
> $$k = 0, 1, \dots, n$$

<u>démonstration</u> : dans le cours oral.

3) <u>Loi géométrique ou loi de Pascal</u> : On considère <u>la répétition illimitée</u> d'une expérience à deux issues s (= succès), e (= échec)

avec $P(s) = p$ et $P(e) = 1-p$ $(0 < p < 1)$.

qu'on modélise par l'espace produit infini $\Omega = \{s, e\}^{\mathbb{N}^*}$ muni de la probabilité (cylindrique) $\mathbb{P}$ définie à la fin du chapitre 2.

> <u>Définition</u>: La v.a. $X =$ l'instant du premier succès, s'appelle <u>v.a. géométrique</u> (<u>ou de Pascal</u>) <u>de paramètre p</u>.

(Ici on a $X(\Omega) = \{1, 2, \dots\} = \mathbb{N}^*$).

> <u>Proposition</u>: La loi de probabilité d'une v.a. de Pascal de paramètre p, est donnée par :
>
> $$\begin{cases} p_k = \mathbb{P}(X = k) = q^{k-1} p & (\text{où } q = 1-p) \\ k \in \mathbb{N}^* \end{cases}$$

<u>démonstration</u> : dans le cours oral.

4) <u>Loi de Poisson</u> : Soit $\lambda > 0$ un paramètre (fixé) et X une v.a. discrète telle que $X(\Omega) = \mathbb{N}$.

> <u>Définition</u>: On dit que la v.a. X à valeurs dans $\mathbb{N}$ suit la loi de Poisson de paramètre λ (>0)

si :

$$\forall k \in \mathbb{N},\ p_k = \mathbb{P}(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

Remarque: la loi de Poisson n'apparait pas de façon explicite comme les lois précédantes mais comme une loi approchée comme le montre le résultat fondamental suivant

Théorème (approximation de la loi binomiale par la loi de Poisson): Soient $n \in \mathbb{N}$ et $p \in]0,1[$. Supposons que $n \to +\infty$ et $p \to 0$ de telle sorte que $np \to \lambda > 0$ où λ est un paramètre fixé. Alors pour tout $k \in \mathbb{N}$ (fixé), on a

$$\binom{n}{k} p^k (1-p)^{n-k} \longrightarrow e^{-\lambda}\frac{\lambda^k}{k!}$$

quand $n \to +\infty$, $p \to 0$ et $np \to \lambda$.

démonstration:

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{1}{k!}(n-k+1)(n-k+2)\cdots n\, p^k (1-p)^{n-k}$$

$$= \frac{1}{k!}(np)^k (1-p)^n \left(1-\frac{k-1}{n}\right)\left(1-\frac{k}{n}\right)\cdots\left(1-\frac{2}{n}\right)(1-p)^{-k}.$$

Si on pose $np = \lambda + \varepsilon$ où $\varepsilon \to 0$, alors

$$(1-p)^n = \left(1-\frac{\lambda}{n}-\frac{\varepsilon}{n}\right)^n \to e^{-\lambda} \quad (\text{si } n\to+\infty \text{ et } \varepsilon \to 0)$$

d'où on déduit le résultat du théorème.

Exemple d'utilisation pratique de ce résultat:

Si n est grand, p est petit et $np = \lambda > 0$ alors la loi binomiale $\mathcal{B}(n,p)$ est approximativement égale à la loi de Poisson $\mathcal{P}(\lambda)$ de paramètre λ. Par exemple si $n = 100$ et $p = 0,1$, alors

$$\mathcal{B}(100; 0.1) \simeq \mathcal{P}(\lambda=10).$$

III Espérance d'une variable aléatoire discrète

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé modélisant une certaine expérience aléatoire et considérons $X : \Omega \to \mathbb{R}$, une v.a. discrète. Comme dans les paragraphes précédents notons $X(\Omega) = \{x_i ; i \in D\}$ l'ensemble fini ou dénombrable des valeurs prises par X et $p_i = \mathbb{P}(X = x_i)$, $i \in D$, sa loi de probabilité.

Définition: L'espérance mathématique de la v.a. X est définie par

$$\mathbb{E}(X) := \sum_{i \in D} p_i x_i \quad (*)$$

à condition, lorsque $D = \mathbb{N}$, que le second membre soit une série absolument convergente. Dans ce cas on dit que X a une espérance finie. (Lorsque $\sum_{i\in\mathbb{N}} p_i |x_i| = +\infty$, X n'a pas d'espérance).

- On dit que X est centrée lorsque $\mathbb{E}(X) = 0$.

<u>Remarque</u> : Par analogie avec la mécanique, si l'on considère chaque valeur x_i comme un "point" affecté de la "masse" p_i, <u>$E(X)$ est le barycentre des points massifs (x_i, p_i)</u>. L'espérance peut (doit) être considérée comme <u>un paramètre de position de la v.a. X</u>.
L'espérance $\mathbb{E}(X)$, quand elle existe, s'appelle aussi : <u>valeur moyenne de la v.a. X</u>.

<u>Cas particulier d'un univers fini ou dénombrable</u>

Supposons que l'espace $\Omega = \{\omega_1, \omega_2, \dots\}$ soit fini ou dénombrable. On a dans ce cas particulier, une autre formule pour calculer $\mathbb{E}(X)$ (lorsqu'elle existe)

<u>Proposition</u> (formule du transfert). On a l'égalité

$$\sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) X(\omega) = \sum_{i \in D} p_i x_i \quad (= \mathbb{E}(X))$$

valable dès que l'une des séries ci-dessus est absolument convergente (l'autre l'est alors aussi)

<u>démonstration</u> : $A_i = [X = x_i]$, $i \in D$, est un système complet d'événements (i.e. une partition de Ω)
Par sommation par paquets, on a :

$$\sum_{\omega \in \Omega} P(\{\omega\}) X(\omega) = \sum_{i \in D} \sum_{\omega \in A_i} \mathbb{P}(\{\omega\}) X(\omega) \quad (*)$$

Mais pour $\omega \in A_i$, $X(\omega) = x_i$ donc :

$$(*) = \sum_{i \in D} x_i \sum_{\omega \in A_i} \mathbb{P}(\{\omega\})$$

$$= \sum_{i \in D} x_i \, \mathbb{P}(A_i) = \sum_{i \in D} x_i \, \mathbb{P}(X = x_i) \quad \text{cqfd.}$$

<u>Exemples</u>

1) <u>Une v.a. sans espérance</u> : On lance une pièce de monnaie (régulière). Si pile apparaît pour la première fois au $n$ième lancer, on gagne $X = 2^n$ euros. Cette v.a. X prend les valeurs 2^n avec probabilité $\mathbb{P}(X = 2^n) = \frac{1}{2^n}$. Elle n'a pas d'espérance finie car $\sum_{n \geq 1} p_n x_n = \sum_{n \geq 1} \frac{1}{2^n} 2^n = +\infty$.

2) <u>Si X est une v.a. binomiale de paramètres (n, p)</u> son espérance est égale à $\mathbb{E}(X) = np$ (exercice)

3) <u>Si X est une v.a. de Poisson de paramètre</u> $\lambda > 0$, $\mathbb{E}(X)$ existe et vaut λ (exercice).

4) <u>Si X est une v.a. de Pascal de paramètre p</u>, $\mathbb{E}(X)$ existe et $\mathbb{E}(X) = \frac{1}{p}$. En effet $X(\Omega) = \{1, 2, \dots\} = \mathbb{N}^*$ $\mathbb{P}(X = n) = q^{n-1} p$ (où $q = 1 - p$). D'où $\sum_{n=1}^{+\infty} n q^{n-1} p = p \sum_{n=1}^{+\infty} n q^{n-1}$

$$= p \frac{1}{(1-q)^2} = p \frac{1}{p^2} = \frac{1}{p}$$ (on a utilisé le fait que $\sum_{n=1}^{\infty} n q^{n-1}$ est la série entière dérivée de

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q} \quad (|q|<1)$$

Propriétés fondamentales de l'espérance

Considérons deux v.a. discrètes X et Y définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. La v.a. $X+Y$ est définie par :

$$\forall \omega \in \Omega, (X+Y)(\omega) = X(\omega) + Y(\omega)$$

et pour $a \in \mathbb{R}$, on considère la v.a. aX définie par

$$\forall \omega \in \Omega, (aX)(\omega) = aX(\omega).$$

Théorème (linéarité de l'espérance mathématique)
Soient X et Y 2 v.a. ayant une espérance finie. Alors :

1) $X+Y$ a une espérance finie et
$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

2) $\forall a \in \mathbb{R}$, aX a une espérance finie et
$$\mathbb{E}(aX) = a\,\mathbb{E}(X).$$

démonstration (cas particulier) : supposons Ω fini ou dénombrable. Alors

1)
$$\sum_{\omega\in\Omega} (X+Y)(\omega)\,\mathbb{P}(\{\omega\}) = \sum_{\omega\in\Omega} (X(\omega)+Y(\omega))\,\mathbb{P}(\{\omega\})$$
$$= \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\{\omega\}) + \sum_{\omega\in\Omega} Y(\omega)\,\mathbb{P}(\{\omega\})$$
$$= \mathbb{E}(X) + \mathbb{E}(Y) \quad \text{cqfd.}$$

2) évident.

Corollaire : Soient $X_1, X_2, \dots, X_n$ des v.a. ayant une espérance. Alors la v.a. $X_1+X_2+\dots+X_n$ a une espérance : $\mathbb{E}(X_1+\dots+X_n) = \mathbb{E}(X_1)+\dots+\mathbb{E}(X_n)$.

dém^{n} : récurrence évidente sur n.

Exemple : Soit X une v.a. binomiale de paramètres (n,p). Considérons les v.a. :

$$\mathcal{X}_i = \begin{cases} 1 & \text{si succès au } i^e \text{ coup} \\ 0 & \text{sinon.} \end{cases}$$

Les $\mathcal{X}_i$ sont des v.a. de Bernoulli de paramètre p et $X = \mathcal{X}_1+\mathcal{X}_2+\dots+\mathcal{X}_n$. On $\mathbb{E}(\mathcal{X}_i) = p$ $(\forall i = 1,\dots,n)$ donc $\mathbb{E}(X) = np$.

IV Moments d'une v.a. discrète. Notion de variance

Soit X une v.a. discrète définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et soit $r \geq 1$ un entier.

Définition : 1) On dit que X a un moment d'ordre r si la v.a. X^r a une espérance finie. Dans ce cas $\mathbb{E}(X^r)$ s'appelle le moment d'ordre r de la v.a. X.
2) On appelle variance de X (lorsqu'elle existe) le moment d'ordre 2 de la v.a. centrée $X-\mathbb{E}(X)$. On note alors $\operatorname{Var} X = \mathbb{E}((X-\mathbb{E}(X))^2)$.

Calcul du moment d'ordre 2 et de la variance :

Soit X une v.a. discrète et $f : \mathbb{R} \to \mathbb{R}$ une fonction déterministe. Considérons la v.a. Y définie par :

$Y = f(X) (= f \circ X)$ définie pour tout $\omega \in \Omega$ par :

$$Y(\omega) = f(X(\omega)).$$

on dit que Y est une fonction déterministe de la v.a. X

On peut, lorsqu'elle existe, calculer $\mathbb{E}(Y)$ à partir de la loi (x_i, p_i), $i \in D$, de X :

Proposition : $\mathbb{E}(Y)$ existe si et seulement si la série $\sum_i p_i f(x_i)$ est absolument convergente et on a :

$$\mathbb{E}(Y) = \mathbb{E}(f(X)) = \sum_{i \in D} p_i f(x_i).$$

démonstration : dans le cours oral.

Corollaire 1 Le moment d'ordre 2 de la v.a. X (lorsqu'il existe) est donné par la formule

$$\mathbb{E}(X^2) = \sum_{i \in D} p_i x_i^2$$

démonstration : on applique la proposition avec la fonction $f(x) = x^2$.

Corollaire 2 : la variance de la v.a. X (si elle existe) est donnée par la formule :

$$\operatorname{Var} X = \sum_{i \in D} p_i (x_i - \mathbb{E}(X))^2.$$

Corollaire 3 (formule pratique) : La variance de X existe si et seulement si X a un moment d'ordre 2. On a alors la formule :

$$\operatorname{Var} X = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

démonstration : voir le cours oral.

Ecart type d'une v.a.

Définition : L'écart type d'une v.a. X ayant un moment deux est le nombre

$$\sigma_X = \sqrt{\operatorname{Var} X}$$

Remarque et exercice : l'écart type d'une v.a. est le nombre σ_X tel que la v.a. $\frac{X - \mathbb{E}(X)}{\sigma_X}$ soit de variance (resp. d'écart type) égal à 1.

La v.a. $\frac{X - \mathbb{E}(X)}{\sigma_X}$ s'appelle v.a. centrée et réduite associée à la v.a. X.

Remarque : La variance (resp. l'écart type) est un paramètre de dispersion car elle mesure l'écart quadratique moyen des valeurs centrées $x_i - \mathbb{E}(X)$.

Proposition : 1) La variance d'une loi binomiale de paramètres (n, p) est le nombre npq $(= np(1-p))$

2) La variance d'une loi de Poisson de paramètre $\lambda > 0$ est le nombre λ.

(démⁿ : exercice)

V Variables aléatoires indépendantes

Soient $X_1, X_2, \dots, X_n$ des v.a. discrètes définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Définition : On dit que $X_1, X_2, \dots, X_n$ sont indépendantes si $\forall\ (a_1, a_2, \dots, a_n) \in X_1(\Omega) \times X_2(\Omega) \times \dots \times X_n(\Omega)$

$$\mathbb{P}(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = a_i)$$

On a noté

$$[X_1 = a_1, \dots, X_n = a_n] = [X_1 = a_1] \cap [X_2 = a_2] \cap \dots \cap [X_n = a_n]$$

pour alléger les notations. Nous utiliserons cette notation simplifiée pour l'intersection d'événements dans la suite du cours.

Exemple typique de v.a. indépendantes : On répète n fois une même expérience aléatoire (par exemple tirages avec remise). Soient X_1 une v.a. relative au 1er essai, X_2 une v.a. relative au 2e essai, ..., X_n une v.a. relative au $n$ième essai. Alors les v.a. $X_1, X_2, \dots, X_n$ sont indépendantes. C'est toujours sous cette forme qu'apparaîtront des v.a. indépendantes.

Exemple : Soit A un événement. Posons

$$X_k = \begin{cases} 1 & \text{si } A \text{ se réalise au } k^e \text{ essai} \\ 0 & \text{sinon.} \end{cases}$$

Les v.a. X_k sont indépendantes (par construction)

Théorème (espérance d'un produit de v.a. indépendantes)

Si $X_1, \dots, X_n$ sont des v.a. indépendantes ayant une espérance alors la v.a. produit $X_1 X_2 \cdots X_n$ a aussi une espérance donnée par la formule

$$\mathbb{E}(X_1 X_2 \cdots X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2)\cdots\mathbb{E}(X_n)$$

démonstration (simplifiée) supposons $n = 2$ et considérons le cas particulier où on répète 2 fois de manière indépendante une même expérience aléatoire et que X_i $(i = 1,2)$ est une v.a. relative à la i^e expérience Ω_i $(i = 1,2)$. La v.a. $X_1 X_2$ est définie sur $\Omega_1 \times \Omega_2$ et on a

$$\mathbb{E}(X_1 X_2) = \sum_{\omega_1, \omega_2} X_1(\omega_1) X_2(\omega_2) \mathbb{P}((\omega_1, \omega_2)) \quad (*)$$

(on suppose aussi que Ω_i est dénombrable). Alors

$$(*) = \sum_{\omega_1, \omega_2} X_1(\omega_1) X_2(\omega_2) \mathbb{P}_1(\omega_1) \mathbb{P}_2(\omega_2)$$

$$= \Big(\sum_{\omega_1} X_1(\omega_1) \mathbb{P}_1(\omega_1) \Big) \Big(\sum_{\omega_2} X_2(\omega_2) \mathbb{P}_2(\omega_2) \Big) = \mathbb{E}(X_1)\mathbb{E}(X_2)$$

Corollaire (variance d'une somme de v.a. indépendantes)

Si $X_1, \dots, X_n$ sont des v.a. indépendantes ayant un moment d'ordre 2, alors $X_1 + X_2 + \dots + X_n$ admet un moment d'ordre 2 et on a

$$\mathrm{Var}(X_1 + \dots + X_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

démᵒⁿ (cas $n=2$) : $\mathrm{Var}(X_1+X_2)=\mathbb{E}\left((X_1-E(X_1)+X_2-E(X_2))^2\right)$

$$= \mathrm{Var}(X_1)+\mathrm{Var}(X_2)+2\underbrace{\left(\mathbb{E}(X_1X_2)-E(X_1)\mathbb{E}(X_2)\right)}_{0 \text{ d'après le théorème}}$$

$$= \mathrm{Var}(X_1)+\mathrm{Var}(X_2)$$

<u>Exemple</u> : si X est une v.a. binomiale de paramètres (n,p), $\mathrm{Var}\,X = np(1-p)$.

En effet écrivons $X = X_1+X_2+\dots+X_n$ où X_k vaut 1 ou 0 suivant qu'on obtient ou non succès au $k^{\text{ième}}$ essai.

Les X_k sont des v.a. de Bernoulli indépendantes et de même paramètre p et $\mathrm{Var}(X_k)=p(1-p)$

$$\mathrm{Var}\,X = \sum_{k=1}^{n}\mathrm{Var}\,X_k = np(1-p).$$

<u>Remarque et définition</u> : Etant donné deux v.a. X et Y, ayant un moment d'ordre 2, on appelle <u>covariance de X et Y</u>, la quantité :

$$\mathrm{cov}(X,Y)=\mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right) \quad (*)$$

On notera que si deux v.a. X_1 et X_2 ont un moment d'ordre 2 alors leur produit X_1X_2 a toujours un moment d'ordre 1 ceci résulte de <u>l'inégalité de Cauchy-Schwartz</u> :

$$\mathbb{E}(|X_1X_2|) \leq \sqrt{E(X_1^2)}\sqrt{E(X_2^2)}$$

(<u>exercice</u>). Ainsi la définition $(*)$ a un sens.

Or en développant le produit $(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))$, on trouve

$$\mathrm{cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)$$

La formule générale de la variance d'une somme de v.a. (non nécessairement indépendantes) s'écrit alors

$$\mathrm{Var}(X+Y)=\mathrm{Var}\,X+\mathrm{Var}\,Y+2\,\mathrm{cov}(X,Y).$$

L'indépendance de X et Y implique que $\mathrm{cov}(X,Y)=0$.
La réciproque n'est pas vraie.

VI <u>La loi faible des grands nombres</u>.

<u>Définition</u> : On dit que $X_1,X_2,\dots,X_n,\dots$ est <u>une suite de v.a. indépendantes</u> si pour tout entier n, les v.a. $X_1,X_2,\dots,X_n$ sont indépendantes.

<u>Théorème</u> : Soit $X_1,X_2,\dots,X_n,\dots$ une suite de v.a. indépendantes et de même loi ayant une espérance m et une variance σ^2. Alors :

$$\forall \varepsilon>0, \quad \lim_{n\to+\infty}\mathbb{P}\left(\left|\frac{X_1+\dots+X_n}{n}-m\right|>\varepsilon\right)=0$$

<u>Lemme</u> (<u>inégalité de Bienaymé-Tchebychev</u>) Soit X une v.a. d'espérance $E(X)$ et de variance σ_X^2. Alors pour tout $a>0$, on a

$$\mathbb{P}\left(\left|\frac{X-\mathbb{E}(X)}{\sigma_X}\right|>a\right)\leq\frac{1}{a^2}$$

démⁿ du lemme : la v.a. $Y = \frac{X - \mathbb{E}(X)}{\sigma_X}$ est d'espérance nulle et de variance $\mathrm{Var}\, Y = \mathbb{E}(Y^2) = 1$. Notons y_i, $i \in D$ les valeurs prises par Y. On a :

$$1 = \mathbb{E}(Y^2) = \sum_i y_i^2\, \mathbb{P}(Y = y_i) \geq \sum_{i : |y_i| > a} y_i^2\, \mathbb{P}(Y = y_i)$$

(dans la somme de droite on somme sur les indices i tels que $|y_i| > a$). Ainsi :

$$1 = \mathbb{E}(Y^2) \geq a^2 \sum_{i : |y_i| > a} \mathbb{P}(Y = y_i) = a^2\, \mathbb{P}(|Y| > a)$$

D'où le résultat.

démⁿ du théorème : Posons $Z_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$. On a :

$$\mathbb{E}(Z_n) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i) = m$$

$$\mathrm{Var}\, Z_n = \frac{1}{n^2} \mathrm{Var}\Big(\sum_{i=1}^{n} X_i\Big) = \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}\, X_i \quad \text{(les } X_i \text{ sont indép.)}$$

$$= \frac{1}{n^2}\, n \sigma^2 = \frac{\sigma^2}{n} \Rightarrow \sigma_{Z_n} = \frac{\sigma}{\sqrt{n}}.$$

L'inégalité de Bienaymé-Tchebychev donne alors :

$$\forall a > 0, \quad \mathbb{P}\left(\left| \frac{Z_n - \mathbb{E}(Z_n)}{\sigma_{Z_n}} \right| > a \right) \leq \frac{1}{a^2}$$

i.e. $\forall a > 0$, $\mathbb{P}\left(|Z_n - m| > a \frac{\sigma}{\sqrt{n}} \right) \leq \frac{1}{a^2}$

En posant $\varepsilon = a \frac{\sigma}{\sqrt{n}}$, on obtient

$$\mathbb{P}(|Z_n - m| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2 n} \to 0 \text{ si } n \to +\infty$$

cqfd.

Exemple : Soit $\mathcal{E}$ une expérience aléatoire et A un événement pouvant se réaliser avec la probabilité p. Soit $n(A)$ le nombre de fois où l'événement A se réalise au cours de n répétitions indépendantes de l'expérience $\mathcal{E}$.

$\frac{n(A)}{n}$ = la fréquence (de réalisations) de A au cours des n répétitions de $\mathcal{E}$. On déduit du Théorème que

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left| \frac{n(A)}{n} - p \right| > \varepsilon \right) \longrightarrow 0 \text{ si } n \to +\infty.$$

Ainsi si n est grand, il est très improbable que la fréquence relative $\frac{n(A)}{n}$ s'écarte de la valeur (théorique) p de plus de ε (quel que soit $\varepsilon > 0$ fixé). Ceci est à la base des méthodes de la Statistique : la v.a. $\frac{n(A)}{n}$ est un estimateur du paramètre p. (voir les exercices de T.D.).