UNIVERSITE SIDI MOHAMED BEN ABDELLAH
FACULTE DES SCIENCES ET TECHNIQUES
Département des Mathématiques

Licence Sciences et Techniques (LST)

MATHEMATIQUES ET APPLICATIONS

MEMOIRE DE FIN D'ETUDES

Pour l'obtention du Diplôme de Licence Sciences et Techniques

Titre

Théorème de la limite centrale et ses applications

Présenté par :

- Moussa Salim

Encadré par :

- Pr F.Ezzaki

Soutenu Le 10 Juin 2016 devant le jury composé de:

- Pr F.Ezzaki
- Pr O.Ammor
- Pr R.Elkhaoulani

Année Universitaire 2015 / 2016

FACULTE DES SCIENCES ET TECHNIQUES FES – SAISS
B.P. 2202 – Route d'Imouzzer – FES
212 (0)5 35 61 16 86 – Fax : 212 (0)5 35 60 82 14
Site web : http://www.fst-usmba.ac.ma

Remerciement

Je saisie cette agréable occasion pour présenter mes chaleureux remerciements à tous ceux qui ont contribué à
la réalisation de ce travail, qui n'aurait jamais vu le jour sans leur coopération, leur encadrement et leurs recommandations. Je remercie plus particulièrement :

Madame Ezzaki Fatima : mon encadrant

Je vous remercie d'abord pour votre gentillesse, de vos compétences ensuite de tous les aides et les instructions précieux que vous m'avez fournir

Les membres de jury

Monsieur El Khaoulani Rachid :

Je vous remercie de votre rigueur, vos qualités humaines et votre Compétence professionnelle.

Madame Ouafae Ammor :

Je vous remercie de votre bienveillance,
votre gentillesse, et votre compétence professionnelle.

Je vous remercie du grand honneur que vous me faites en vous intéressant à mon travail et en acceptant de le juger.

Veuillez trouver, chers professeurs, dans ce modeste travail, l'expression de mon profond respect, ma haute
considération et ma sincère gratitude

Table des matières

Introduction

Nous abordons dans ce travail l'un des théorèmes de probabilité les plus célèbre, à savoir le théorème de la limite centrale .ce théorème a trouvé un grand succès et plusieurs applications dans la théorie de la statistique inférentielle.

Grace au théorème de la limite centrale, les statisticiens ont pu prouver et exploiter la normalité de la moyenne arithmétique d'une suite de variables aléatoire $(X_i)_{1\le i\le n}$ qui est considérée aujourd'hui l'un des résultats puissants des statistiques inférentielle.

Dans le premier chapitre nous traitons la loi faibles des grands nombres le second chapitre sera consacré au théorème de la limite centrale et dans le troisième chapitre nous donnons quelques applications de ce théorème.

Chapitre 1
Rappel sur la loi des grands nombres

Historique

C'est sur la loi des grands nombres que reposent la plupart des sondages. On interroge un nombre suffisamment important de personnes pour connaître leur opinion pour donner une idée sur le comportement de la population entière.

La loi des grands nombres sert aussi en statistique inférentielle, pour déterminer une loi de probabilité à partir d'une série d'expériences.

Les mathématiciens distinguent deux énoncés, appelés respectivement « loi faible des grands nombres » et « loi forte des grands nombres ».

On s'intéresse à la moyenne arithmétique d'une suite de variables aléatoires dans le même espace de probabilité

la loi faibles des grands parle de la convergence en probabilité c'est-à-dire que la probabilité que l'écart entre cette moyenne et son espérance en valeur absolue soit supérieur à une valeur précise vaille 0.

En revanche la loi forte des grands nombres parle de la convergence presque sûrement c'est à dire que la probabilité que la moyenne arithmétique converge vers l'espérance vaut 1.

La loi faible des grands nombres est également appelée théorème de Khinchine (rarement utilisé).

On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendants dans l'espace probabilisé$(\Omega, \mathcal{F}, p)$.

Supposons que ces variables aléatoires admettant la même espérance $\mu = E(X)$ et la même variance $\sigma^2 = V(X)$

Théorème1 **:**
soit Y_n une variable aléatoire définit sur $(\Omega, \mathcal{F}, p)$ tel que
$Y_n = \frac{\sum_{i=1}^{n} X_i}{n}$ alors pour tout ε réel strictement positif

$$\lim_{n\to+\infty} P(|Y_n - E(Y_n)\ | \geq \varepsilon) = 0$$

Autrement dit, $(Y_n)_{n\in\mathbb{N}}$ converge en probabilité vers *E(X)*.

Ce résultat est très important en statistique, puisqu'il assure que la moyenne arithmétique des $(X_i)_{1\leq i\leq n}$ est un estimateur convergent de l'espérance *E(X)*.

Proposition **:**

Soit X' une variable aléatoire sur $(\Omega, \mathcal{F}, p)$ à valeur dans $\mathbb{R}$ supposons que $\mu = E(X')$ et $\sigma^2 = V(X')$ alors $\forall\varepsilon > 0$,

$$P(|X' - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \ (*)$$

L'inégalité (*) s'appelle inégalité de Bienaymé-Tchebychev

Démonstration du théorème1 **:**

Posons $S_n = \sum_{i=1}^{n} X_i$, et $Y_n = \frac{S_n}{n}$

Alors $E(S_n) = E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = n\mu$

$V(S_n) = \sum_{i=1}^{n} V(X_i) = n\sigma^2$
L'espérance de Y_n est $E(Y_n) = \mu$ et la variance de Y_n est $V(Y_n) = \frac{\sigma^2}{n}$
Soit l'événement $A_n = \{|Y_n - \mu| \geq \varepsilon\}$ appartient à la tribu $\mathcal{F}$

D'après la proposition précédente $P(|Y_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$
nous constatons que la probabilité de A_n est majoré par une suite qui tend vers 0 quand n tend vers plus l'infini, par conséquent :

$$\lim_{n \to +\infty} P(A_n) = 0$$

La probabilité pour qu'une moyenne arithmétique s'éloigne de l'espérance est nulle donc la convergence en probabilité de la suite $(Y_n)_{n\geq 1}$ vers μ est confirmé.

La loi que nous venons de démontrer ne correspond pas exactement au fait que la moyenne arithmétique converge vers l'espérance mais équivalent à dire qu'à l'intuition qu'en lançant une pièce un grand nombre de fois la proportion de face atteigne véritablement $\frac{1}{2}$

Théorème2 (Borel-Cantelli) **:**

Soit $(\Omega, \mathcal{F}, p)$ un espace probabilisé, et soit $(A_n)_{n\geq 1}$ une suite d'événements aléatoires appartiennent à la tribu $\mathcal{F}$ on définit l'ensemble B de la manière suivante

$$B = \lim_{n\to+\infty} sup_n A_n = \cap_p \cup_{n\geq p} A_n$$

Si $\sum_{n\geq 1} p(A_n) < +\infty$ alors nous avons $p(B) = 0$

Si de plus les A_n sont indépendants alors si $\sum_{n\geq 1} p(A_n) = +\infty$ nous aurons $p(B) = 1$.

Propriété1 (la continuité monotone) :

Soit un espace probabilisé $(\Omega, \mathcal{F}, p)$ et soient $(A_n)_{n\geq 1}$ appartiennent au tribu $\mathcal{F}$

Toute suite croissante d'événements $(A_n)_{n\geq 1}$,$A_1 \subset A_2 \subset A_3 \dots \dots$ satisfait :

$$p\left(\bigcup_{n\geq 1} A_n\right) = \lim_{n\to+\infty} p(A_n)$$

Toute suite décroissante d'événements $(A_n)_{n\geq 1}$,$A_1 \supset A_2 \supset A_3 \dots ..$ satisfait :

$$p\left(\bigcap_{n\geq 1} A_n\right) = \lim_{n\to+\infty} p(A_n)$$

Axiome de probabilité :

Soit $(\Omega, \mathcal{F}, p)$ un espace probabilisé et soit $(B_n)_{n\geq 1}$ des événements qui appartiennent à la tribu $\mathcal{F}$ nous avons :

$$p(\bigcup_{n\geq 1} B_n) \leq \sum_{n\geq 1} p(B_n)$$

Théorème3 (convergence presque sûrement) :

Considérons une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes dans$(\Omega, \mathcal{F}, p)$ qui suivent la même loi de probabilité de même espérance μ et de même variance σ^2et intégrables, En reprenant les notations ci-dessus, la loi forte des grands nombres précise que $(Y_n)_{n\in\mathbb{N}}$ converge vers *E(X)* *presque sûrement* noté « p.s. »

Avec $Y_n = \frac{X_1+X_2+\cdots X_n}{n}$, $Y_n = \frac{X_1+X_2+\cdots+X_n}{n} \to m, n \to +\infty$

$$p(\{\omega : \lim_{n\to+\infty} Y_n(\omega) = m\}) = 1$$

Démonstration du théorème2 :

Nous partageons la démonstration en deux parties

- *Démonstration dans le cas où* $\sum_{n\geq 1} p(A_n) < +\infty$

Nous allons poser $B_p = \bigcup_{n\geq p} A_n$ où p est un entier fixé.

Il est facile de voir que l'ensemble B_p est un ensemble décroissant au sens de l'inclusion pour cela nous allons utiliser la propriété 1

$$\begin{aligned} p\left(\bigcap_p B_p\right) &= \lim_p p(B_p) \\ &= \lim_{p\to+\infty} p(\bigcup_{n\geq p} A_n) \\ &\leq \lim_{p\to\infty} \sum_{n\geq p} p(A_n) \end{aligned}$$

Or $\sum_{n\geq p} p(A_n)$ est le reste de la série du terme général $p(A_n)$ or on a fait l'hypothèse que la série $\sum_{n\geq 1} p(A_n)$ converge donc le reste tend vers 0 et par suite

$$p\left(\bigcap_p B_p\right) \leq 0 \Longrightarrow p\left(\bigcap_p B_p\right) = 0$$

- *Démonstration dans le cas où les* $(A_n)_{n\geq 1}$ *sont indépendants*

Sous l'hypothèse (*)

$\sum_n p(A_n) = +\infty$

Soit m un entier

$$p\left(\bigcup_{n=p}^{m} A_n\right) = 1 - p(\left(\bigcup_{n=p}^{m} A_n\right)^{C})$$
$$= 1 - p(\bigcap_{n=p}^{m} {A_n}^{C})$$

L'indépendance des événements A_n va nous permettre d'écrire

$$p\left(\bigcup_{n=p}^{m} A_n\right) = 1 - \prod_{n=p}^{m} p(\ {A_n}^{C})$$
$$= 1 - \prod_{n=p}^{m}(1 - p(A_n))$$

Nous allons utiliser une technique d'analyse qu'il s'agit d'une petite inégalité

$\forall x \in [0,1] \quad 1 - x \leq e^{-x}$ ou bien $-(x-1) \geq -e^{-x}$

Appliquons donc cette inégalité

$$p\left(\bigcup_{n=p}^{m} A_n\right) \geq 1 - \prod_{n=p}^{m} e^{-p(A_n)}$$
$$p\left(\bigcup_{n=p}^{m} A_n\right) \geq 1 - e^{-\sum_{n=p}^{m} p(A_n)}$$

Nous faisons tendre m vers plus l'infini nous aurons

$$\lim_{m\to+\infty} p\left(\bigcup_{n=p}^{m} A_n\right) \geq 1 - \lim_{m\to+\infty} e^{-\sum_{n=p}^{m} p(A_n)}$$
$$p\left(\bigcup_{n\geq p} A_n\right) \geq 1 - e^{-\sum_{n\geq p} p(A_n)}$$

L'hypothèse (*) va donner comme résultat

$$p\left(\bigcup_{n\geq p} A_n\right) \geq 1$$

Donc $p\left(\bigcup_{n\geq p} A_n\right) = 1$ (**)

Nous allons poser $N_P = (\bigcup_{n\geq p} A_n)^{C} = \bigcap_{n\geq p} {A_n}^{C}$

Le résultat (**) va nous donner $p\left(N_p\right) = 0$

Montrons que $p((\lim_{n\to+\infty} sup_n A_n)^{C}) = 0$

$$p((\lim_{n\to+\infty} sup_n A_n)^{C}) = p(\left(\bigcap_p \bigcup_{n\geq p} A_n\right)^{C})$$

$$p((\lim_{n\to+\infty} sup_n A_n)^{\complement}) = p(\cup_p N_p)$$

Nous allons utiliser l'axiome de probabilité

$$p((\lim_{n\to+\infty} sup_n A_n)^{\complement}) \leq \sum_p p(N_p)$$
$$\leq 0$$

donc $$p((\lim_{n\to+\infty} sup_n A_n)^{\complement}) = 0$$
$$p(\lim_{n\to+\infty} sup_n A_n) = 1$$

Et par suite le théorème de Borel-Cantelli est confirmé.

Démonstration du théorème3 :

* La preuve dans le cas où E(X^2) $< +\infty$

* nous allons remplacer Y_n par $Y_n - m$ et nous allons supposer que m=0

$Y_n - m$ va être supposé centré

➢ Montrons d'abord que la suite $(Y_{n^2})_{n\geq 1}$ tend p.s. vers 0 .

Pour cela nous allons utiliser l'inégalité de Tchybetchev

Soit l'événement $A_{n,q} = \left\{|Y_{n^2}| \geq \frac{1}{q}\right\}$

$$P(A_{n,q}) \leq \frac{\sigma^2 q^2}{n^2}, \quad \forall q \in \mathbb{N}^* \quad \text{Fixé}$$

$\frac{\sigma^2 q^2}{n^2}$ est le terme général d'une série convergente par conséquent

$\sum_{n\geq 1} P(A_{n,q}) < +\infty$

Posons $N_q = \lim_{n\to+\infty} \mathrm{sup_n\, A_{n,q}} = \cap_n \cup_{m\geq n} A_{m,q}$

Et N=$\cup_q N_q$

Le théorème de Borel-Cantelli implique que $P(N_q) = 0$

nous avons alors par axiome de probabilité

$P(N) \leq \sum_q P(N_q) = 0$ et $P(N^c) = 1$

Alors si $\omega \in N^c$

$$\forall q, \exists\, n_0\,, n \geq n_0 \implies |Y_{n^2}(\omega)| \leq \frac{1}{q}$$

Donc la variable aléatoire Y_{n^2} converge presque sûrement vers 0

- Montrons maintenant que Y_n converge p.s. vers 0

Pour cela nous allons poser l'entier p(n) tel que :

$p(n)^2 \leq n < (p(n)+1)^2$ (***)

$$Y_n - \frac{p(n)^2}{n} Y_{p(n)^2} = \frac{X_1+X_2+\cdots+X_n}{n} - \frac{p(n)^2}{n}\left(\frac{X_1+X_2+\cdots+X_{p(n)^2}}{p(n)^2}\right)$$

$$= \frac{1}{n}\sum_{i=p(n)^2+1}^{n} X_i$$

$$E\left[\left(Y_n - \frac{P(n)^2}{n} Y_{p(n)^2}\right)^2\right] = var\left(Y_n - \frac{p(n)^2}{n} Y_{p(n)^2}\right)$$

Car nous avons supposé que $Y_n - m$ est centré et nous avons supposé aussi que m=0 pour simplifier le travail

$$E\left[\left(Y_n - \frac{P(n)^2}{n} Y_{p(n)^2}\right)^2\right] = \frac{1}{n^2} var\left(\sum_{i=p(n)^2+1}^{n} X_i\right)$$

Les X_i sont indépendantes donc

$$E\left[\left(Y_n - \frac{P(n)^2}{n} Y_{p(n)^2}\right)^2\right] = \frac{1}{n^2}\left(n - \left(p(n)\right)^2\right) var(X_1)$$

Sous l'hypothèse (***) nous aurons

$$E\left[\left(Y_n - \frac{P(n)^2}{n} Y_{p(n)^2}\right)^2\right] \leq \frac{1}{n^2}(2\sqrt{n}+1)\sigma^2$$

Remarque :

$n - p(n)^2 \leq 2p(n) + 1$

Et $p(n) \leq \sqrt{n}$ donc $n - p(n)^2 \leq 2\sqrt{n} + 1$

Nous déduisons par suite en utilisant l'inégalité de Tchebychev que

$$\forall \varepsilon > 0, \quad p\left(\omega: \left| Y_n(\omega) - \frac{p(n)^2}{n} Y_{p(n)^2}(\omega) \right| > \varepsilon\right) \leq \frac{2\sqrt{n} + 1}{n^2 \varepsilon^2} \sigma^2$$

La conclusion se fait de la même manière que dans la convergence presque surement de Y_{n^2}

Or $p\left(\left| Y_n - \frac{p(n)^2}{n} Y_{p(n)^2} \right| > \varepsilon\right)$ est majoré par un terme général d'une Série convergente alors

$$Y_n - \frac{p(n)^2}{n} Y_{p(n)^2} \to 0 \; p.s.$$

Car $\frac{2\sqrt{n}+1}{n^2 \varepsilon^2} \sigma^2 \to 0, n \to +\infty$

Et comme $Y_{p(n)^2} \to 0 \; p.s.$et $\frac{p(n)^2}{n} \to 1$ on déduit que Y_n converge presque sûrement vers 0 c'est-à-dire :

$$Y_n \to 0 \; p.s.$$

Chapitre
2
Théorème de limite central

Le théorème central limite établit la convergence en loi de la somme d'une suite de variable aléatoires vers la loi normale. Intuitivement, ce résultat affirme que toute somme de variables aléatoires indépendantes et identiquement distribuées tend vers une variable aléatoire gaussienne.

Supposons que les $(X_i)_{1\leq i\leq n}$ une suite de variables aléatoires réelles définies sur le même espace de probabilité. $(\Omega, \mathcal{F}, p)$, indépendantes et identiquement distribuées suivant la même loi D. Supposons que l'espérance μ et l'écart-type σ de D existent et soient finis avec $\sigma \neq 0$ considérons la somme

$S_n = X_1 + X_2 + \ldots + X_n$.

Alors

- l'espérance de S_n est $n\mu$
- son écart-type vaut $\sigma\sqrt{n}$.

De plus, quand *n* est *assez grand*, la loi normale $N(n\mu, n\sigma^2)$ est une bonne approximation de la loi de S_n.

Afin de formuler mathématiquement cette approximation, nous allons poser

$$\overline{X_n} = \frac{s_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Et

$$z_n = \frac{S_n - n\mu}{\sigma/\sqrt{n}}$$

de sorte que l'espérance et l'écart-type de Z_n valent respectivement 0 et 1 la variable est ainsi dite *centrée et réduite*.

Le théorème central limite énonce alors que la suite de variables aléatoires Z_1, Z_2,..., Z_n,... converge en loi vers une variable aléatoire Z, définie sur le même espace probabilisé, et de loi normale centrée réduite $N(0,1)$ lorsque *n* tend vers l'infini.

Théorème1 :
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles indépendants, de même loi, de carrés intégrables, $E[X_1] = \mu \; et \; V[X_1] = \sigma^2$ alors les variables aléatoires $\frac{\sqrt{n}\,(\overline{X_n}-\mu)}{\sigma}$ convergent en loi vers une loi normale centrée et réduite.

Définition :
Soit X une variables aléatoire et soit $(X_n)_{n\geq 1}$ une suite de variable aléatoire définit sur l'espace $(\Omega, \mathcal{F}, p)$ toutes à valeurs dans le même espace métrique (E,d)
On dit que la suite $(X_n)_{n\geq 1}$ ***converge en loi*** vers *X* si, pour toute fonction φ *continue bornée* sur *E*, à valeurs dans $\mathbb{R}$

$$\lim_{n\to+\infty} E[\,\varphi_{X_n}(t)] = E[\varphi_X(t)]$$

Théorème2 (théorème de convergence de Lévy) :

soit une variable aléatoire X et une suite de variable aléatoire indépendantes $(X_n)_{n\geq 1}$ pas nécessairement définit sur le même l'espace $(\Omega, \mathcal{F}, p)$ et soit $\varphi_{X_n}(t)$ et $\varphi_X(t)$ les fonctions caractéristique de X_n et X alors nous avons

$$\lim_{n\to+\infty} \varphi_{X_n}(t) = \varphi_X(t) \Leftrightarrow X_n \overset{\mathcal{L}}{\rightarrow} X$$

Démonstration du théorème de la limite central

Pour un théorème d'une telle importance en statistiques et en probabilité appliquée, il existe une démonstration particulièrement simple utilisant les fonctions caractéristiques.

Soit $(X_j)_{j\geq 1}$ des variables aléatoires indépendantes et identiquement distribuées, posons $Yj = \frac{X_j-\mu}{\sigma}$

il est facile de voir que la moyenne centrée réduite des observations $X_1, X_2, \ldots, X_n$ est simplement :

$$\begin{aligned} Z_n &= \sum_{j=1}^{n} \frac{Y_j}{\sqrt{n}} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_j-\mu}{\sigma} \\ &= \frac{1}{\sqrt{n}} \frac{1}{\sigma} \left(\sum_{i=1}^{n} X_j - n\mu\right) \\ &= \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \end{aligned}$$

Avec $\bar{X} = \frac{\sum_{j=1}^{n} X_j}{n}$

D'après les propriétés élémentaires des fonctions caractéristiques, la fonction caractéristique de Z_n est

$$\begin{aligned} \varphi_{Z_n}(t) &= E[e^{itZ_n}] \\ &= E\left[e^{it\sum_{j=1}^{n}\frac{Y_j}{\sqrt{n}}}\right] \\ &= E\left[\prod_{j=1}^{n} e^{it\frac{Y_j}{\sqrt{n}}}\right] \end{aligned}$$

Les Y_j sont indépendants et identiquement distribuées

$$\varphi_{Z_n}(t) = \left[E(e^{it\frac{Y_1}{\sqrt{n}}})\right]^n$$

En utilisant le développement limité d'ordre 2 de l'exponentielle nous aurons

$$\varphi_{Z_n}(t) = \left(E[1 + \left(i\frac{t}{\sqrt{n}}\right)Y_1 - \frac{1}{2}\left(\frac{t}{\sqrt{n}}\right)^2 {Y_1}^2 + \theta((\frac{t}{\sqrt{n}})^2)]\right)^n$$

La linéarité de l'espérance va aboutir à

$$\varphi_{Z_n}(t) = (E[1] + E\left[\left(i\frac{t}{\sqrt{n}}\right)Y_1\right] - E\left[\frac{1}{2}\left(\frac{t}{\sqrt{n}}\right)^2 {Y_1}^2\right] + E(\theta((\frac{t}{\sqrt{n}})^2)))^n$$

$$= \left(1 - \frac{t^2}{2n} + \theta\left(\frac{t^2}{n}\right)\right)^n \text{ car } Y_1 \textit{est centré et réduite}$$

$$\lim_{n\to+\infty} \varphi_{Z_n}(t) = \lim_{n\to+\infty}\left(1 - \frac{t^2}{2n} + \theta\left(\frac{t^2}{n}\right)\right)^n$$

$$= \lim_{n\to+\infty} \exp(n\ln(1 - \frac{t^2}{2n} + \theta\left(\frac{t^2}{n}\right))$$

En Effectuant le D.L. d'ordre 2 de ln nous aurons comme résultat

$$\lim_{n\to+\infty} \varphi_{Z_n}(t) = \lim_{n\to+\infty} \exp(n(-\frac{t^2}{2n} - \frac{t^4}{4n^2} + \theta(\frac{t^4}{4n^2}))$$

$$= e^{-t^2/2}$$

$$= \varphi_X(t)$$

D'après le théorème de convergence de Lévy le théorème de la limite centrale est confirmé

Dans le cas de variables ne suivant pas une loi normale, le théorème peut sembler étonnant au premier abord. Nous allons donc en faire une illustration ne nécessitant pas de connaissance particulière en statistiques, mais uniquement du dénombrement.

Considérons le jeu de pile ou face et mettons des valeurs sur les faces de la pièce, par exemple 0 pour pile et 1 pour face ; on s'intéresse à la somme de *n* tirages. La pièce est équilibrée, chaque face a une chance sur deux d'être tirée. Si l'on fait un seul tirage, nous avons donc le tirage n°1 (et aucun autre), et son résultat peut être 0 ou 1 ; nous faisons la somme d'une seule valeur.

résultat d'un tirage	
résultat tirage n1	somme
0	0
1	1

pour $n = 2$ possibilités pour la valeur de la somme, apparaissant avec les fréquences suivantes :

fréquence pour tirage		
valeur de la somme	nombres d'apparitions	Fréquence
0	1	0,5
1	1	0,5

Pour n=4 possibilité pour la valeur de la somme, apparaissant avec les fréquences suivantes :

résultat de deux tirages		
résultat tirage n1	résultat tirage n2	somme
0	0	0
0	1	1
1	0	1
1	1	2

fréquence de deux tirages		
valeur de la somme	nombres d'apparitions	fréquence
0	1	0,25
1	2	0,5
2	1	0,25

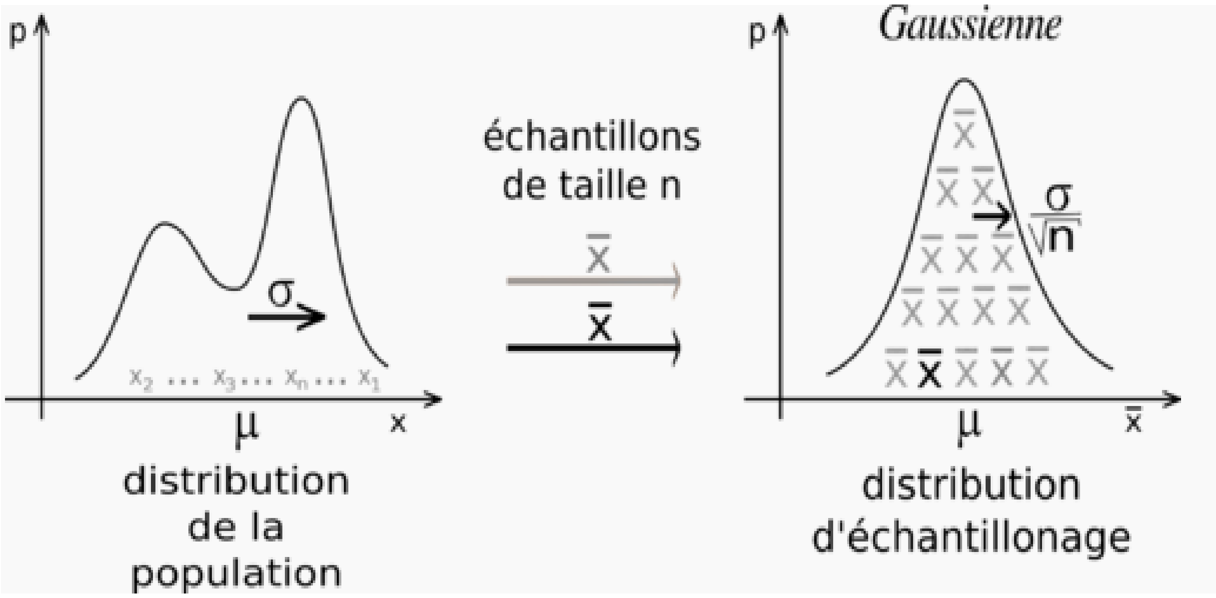
Pour n=8 possibilité pour la valeur de la somme, apparaissant avec les fréquence suivantes :

résultats de trois tirages		
résultats n1	résultats n2	résultats n3
0	0	0
0	0	1
0	1	1
0	1	2
1	0	1
1	0	2
1	1	2
1	1	3

résultats des fréquences		
valeur de la somme	nombres d'apparitions	fréquences
0	1	0,125
1	3	0,375
2	3	0,375
3	1	0,125

Dans toutes les situations ci-dessus, on a des lois uniformes ; et pourtant, la somme d'un grand nombre d'événements tend graphiquement vers une courbe en cloche symétrique. Et cela est vrai même lorsque les lois sont différentes (cas des dés polyédriques).

En effet, on ne s'intéresse pas au tirage en lui-même, mais à la *somme* du tirage. De ce point de vue, plusieurs tirages sont équivalents, donc une valeur de somme peut être obtenue par plusieurs tirages ; par exemple, pour deux dés à six faces (2d6), on peut obtenir 7 par 1+6, 2+5, 3+4, 4+3, 5+2 et 6+1, il y a six tirages équivalents. Or, il y a toujours plus de combinaisons permettant d'obtenir une valeur moyenne qu'une valeur extrême, ce qui donne la courbe en cloche.



Chapitre 3

Quelques applications du théorème De la limite centrale

- ***Intervalle de confiance d'une moyenne***

Soit une variable quantitative X dont la moyenne et la variance de la population sont $\mu\ et\ \sigma^2$.Soit un échantillon de taille n sujets tirés au sort dans la population et soit$\bar{X}$ la moyenne observée sur cette échantillon.

D'après la simulation statistique on sait que la distribution de $\bar{X}$ est approximativement normale de moyenne μ et de variance ${\sigma^2}/{n}$

D'après le théorème de la limite centrale

$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\ converge\ en\ loi\ vers\ N(0,1)$ quand n tend vers $+\infty$

Cela implique donc que $P\left(\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right| < z_{\frac{\alpha}{2}}\right) = 1-\alpha$ (*)

Où α s'appelle le seuil de signification, en pratique ce seuil vaut 1% ou 5%, et $z_{\alpha/2}$ est obtenu à partir d'une table statistique.

Si nous développons la relation (*), nous aurons :

$$P\left(\bar{X} - z_{\alpha/2}\,{\sigma}/{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\,{\sigma}/{\sqrt{n}}\right) = 1-\alpha$$

Cela signifie que l'intervalle de confiance définit par :

I=$[\bar{X} - z_{\alpha/2}\,{\sigma}/{\sqrt{n}}, \bar{X} + z_{\alpha/2}\,{\sigma}/{\sqrt{n}}]$ ne contient pas la vraie valeur de μ avec une probabilité α et contient μ avec une probabilité $1-\alpha$

- intervalle de confiance d'une proportion

Application

Un semencier a récolté 5 tonnes de graines de Tournesol. Il a besoin de connaître le taux de germination de ces graines avant de les mettre en vente. Il extrait un échantillon de 40 graines, les dépose sur un buvard humide et compte le nombre de graines ayant évolué favorablement. On remarque que ce contrôle est de type destructif :

l'échantillon ayant servi au contrôle ne peut plus être commercialisé. Il s'agit donc d'évaluer la proportion p des graines de la population à grand effectif, présentant un certain caractère X : succès de la germination. Même avec une population d'effectif restreint, un contrôle destructif impose de faire confiance à un échantillon restreint et la valeur exacte de p ne peut être calculée.

La statistique du test est $\hat{p}=\frac{X}{n}$ où X est le nombre de grains dans un échantillon de taille n ayant un succès de germination $X \rightarrow B(n,p)$

Il est naturel d'estimer p par $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i$ qui est la proportion des individus ayant le caractère X dans l'échantillon. En effet, la LGN nous assure la convergence en probabilité de la variable aléatoire

De $P = \frac{1}{n}\sum_{i=1}^{n} X_i$ vers l'espérance de X_1, c'est-à-dire p, $\hat{p}$ est l'estimateur de la proportion p et p est estimée par la réalisation de $\hat{p}$ Dans l'expérience de germination, 36 graines ont eu une issue favorable avec x_i= 1. La proportion estimée est $\hat{p} = {}^{36}/_{40}$=0,9.

on cherche à estimer la proportion p de graines défectueuses du lot de céréales. On prélève un lot de n graines et on note Xi la v.a.r. qui vaut 1 si la graine i germe, et 0 sinon. On estime p par $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i$ Les variables aléatoires réelles. Xi étant de Bernoulli, on
peut alors utiliser l'approximation donnée par le TCL.

Soit $Z \xrightarrow{loi} N(0,1)$ et $z_{1-\alpha/2}$ le quantile d'ordre $1 - \alpha/2$ de la loi normal centré et réduite.

Par le TCL $$\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{loi} Z \sim N(0,1)$$

Ceci implique que :

$$P\left(-z_{1-\alpha/2} < \frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} < z_{1-\alpha/2}\right) \xrightarrow[n\to+\infty]{} P\left(-z_{1-\alpha/2} < Z < z_{1-\alpha/2}\right) = 1 - \alpha$$

C'est-à-dire que

$$P(\hat{p} - z_{1-\alpha/2}\frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq p \leq \hat{p} + z_{1-\alpha/2}\frac{\sqrt{p(1-p)}}{\sqrt{n}}) \xrightarrow[n\to+\infty]{} 1 - \alpha$$

ceci ne fournit pas l'intervalle de confiance pour p car les bornes de l'intervalle dépendent de p mais on peut montrer que l'on a le même résultat de convergence en remplaçant p dans l'intervalle par son estimateur convergent on obtient alors

$$P(\hat{p} - z_{1-\alpha/2}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \leq p \leq \hat{p} + z_{1-\alpha/2}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}) \xrightarrow[n\to+\infty]{} 1 - \alpha$$

on dit que l'intervalle

$$I=[\hat{p} - z_{1-\alpha/2}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + z_{1-\alpha/2}\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \;]$$

est un intervalle de confiance asymptotique pour le paramètre p de coefficient de confiance $1-\alpha$.
Pour α= 5%, on lit dans les tables $z_{1-\alpha/2} = 1{,}96$ ainsi, le semencier en déduit qu'ayant observé 36 graines germées sur 40, l'intervalle de confiance asymptotique pour p est $[0{,}807; 0{,}993]$ il suffit de remplacer dans les calculs la moyenne arithmétique aléatoire $\bar{X}_n$ par l'estimation ponctuelle $\hat{p} = {}^{36}/_{40}$.

Nous avons vu au cours de ce travail le rôle crucial du théorème de la limite centrale. Grace à ce théorème nous avons pu approcher la loi de la moyenne arithmétique $\overline{X}$ d'une suite de variable aléatoire X_n par une loi normale, La force de ce théorème réside du fait de la normalité de $\overline{X}$ quel que soit la loi de X.

Nous avons également montré le lien entre la loi des grands nombres et le théorème de la limite centrale.

l'importance du théorème de la limite centrale ne se limite pas uniquement à l'approximation de la loi de $\bar{X}$, mais il joue également un rôle important dans l'estimation par intervalle de confiance de quelques paramètres de la population.

Bibliographie

[1] https://fr.wikipedia.org/wiki/Loi_des_grands_nombres

[2] https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me central limite

[3] https://fr.wikipedia.org/wiki/Loi_forte_des_grands_nombres

[4] http://testard.frederic.pagesperso-orange.fr/mathematiques/statistiques/rappels/inegalites.htm#partieB

[5] http://www.univ-orleans.fr/mapmo/membres/berglund/probamass_html/node7.html

[6] http://exo7.emath.fr/ficpdf/fic00155.pdf

[7] P. Ardilly, technique de sondage, Technip, p 67

[8] http://math.univ-lille1.fr/~suquet/Polys/TLC.pdf

[9] http://www.univ-orleans.fr/mapmo/membres/berglund/probamass_html/node7.html

[10] F .Ezzaki, licence math s6 :cours statistique inférentielle, Probabilité et statistique