Cours

Charles Vin

Date

1 Rappel avec le poly

1.1 Mesure

Exemple 1.4 :
- Lebesgue : $\phi([a,b]) = b - a$
- Dirac : $\delta_x = \delta_x(A) = 1_A(x)$
- Mesure de comptage : C(A) = $\begin{cases} \infty \text{ si A infiny} \\ card(A)\text{sinon} \end{cases}$

4) Mesure discrète : Somme de masse de Dirac $\mu = \sum_{i \in I}^{p_i \delta_i}$ (99% du temps $I = N$)
5) Mesure à densité : Soit f une fonction > 0

$$\mu_f = f dx \Leftrightarrow \mu_f(A) = \int_A f(x) dx.$$

Que signifie $\int f d\mu$?
Si $\mu = dx$ alors $\int f dx = \int f(x) dx$
Si $\mu = \delta_x$ alors $\int f d\delta_x = f(x)$ car δ_x "met que du poids en x"
Si $\mu = \mu_a + \mu_2$ alors $\int f d\mu = \int f d\mu_1 + \int f d\mu_2$
Si $\mu = (1-p)\delta_0 + p\delta_1$ alors

$$\begin{aligned} \int f d\mu &= \int f(1-p)\delta_0 + \int f p \delta_1 \\ &= (1-p)\int f\delta_0 + p\int f\delta_1 \\ &= (1-p)f(0) + pf(1) \end{aligned}$$

Rappels de L2
Si μ à une densité g alors :

$$\mu = g dx \text{ avec } g(x) = (1-x)\mathbb{1}_{[0,1]}(x)$$
$$\int f d\mu = \int f(x) g(x) dx$$

exemple : $f(x) = (1-x)^3$
on a alors

$$\begin{aligned} \int f d\mu &= \int (1-x)^3 \mathbb{1}_{[0,1]}(x) * (1-x) \\ &= \int_0^1 (1-x)^4 dx = 1/5 \end{aligned}$$

Dernier exemple :
Si $\mu = \frac{1}{2}\mathbb{1}_{[0,1]}(x)dx + \frac{1}{2}\delta_3$ alors

$$\begin{aligned} \int f d\mu &= \frac{1}{2}\int f \mathbb{1}_{[0,1]}(x) dx + \frac{1}{2}\int f\delta_3 \\ &= \frac{1}{2}\int_0^1 f(x)dx + \frac{1}{2}f(3) \end{aligned}$$

1.2 Variable aléatoire

Exemple :

- Si $X \sim Bern(p), P_x = (1-p)\delta_0 + p\delta_1$ on a $P_x([2,+\infty[) = P(X \in [2,+\infty]) = 0$
- Si $X \sim N(0,1)$ alors $P_X = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{2}} dx$
- En général, la loi d'une variable aléatoire discrète à valeur dans N :

$$P_X = \sum P(X=1)\delta_1 ????????.$$

Soit $B(p)$

Définition 1.1 (Moment d'une VAR)**.** Si X discrète :

$$E(X) = \sum_{k=0}^{\infty} kP(X=k).$$

Si X à densité fdx

$$E(X) = \int_R xf(x)dx = \int xdP_x \text{ = Nouvelle notation.}$$

Preuve que la nouvelle notation unifie les deux formules de l'esperance :

Preuve :

- Pour la densité : X de densité f_X et $P_x = f_X dx$

$$\int gdP_X = \int gf_X dx = \int g-x)f(x)dx.$$

Si $g(x) = x$ alors

$$\int xdP_X = \int xf_X(x)dx = E(X).$$

- Pour une VA discrète : échaufement avec Bernouilli : $P_x = (1-p)\delta_0 + p\delta_1$

$$E(X) = \sum_{k=0}^{\infty} kP(X=k) = 0 + P + 0 = p.$$

$$\int xdP_x = (1-p)\int xd? + p\int x\delta_1 = p$$
$$= (1-p) * 0 + p = p$$

Fin de l'échauffement :

$$\int xdP_X \text{ si } P_X = \sum_{k=0}^{\infty} P(X=k)\delta_k.$$

$$\int xdP_X = \sum_{k=0}^{\infty} \int xP(X=k)\delta_k$$
$$= \sum P(X=k)\int x\delta_k \qquad = \sum P(X=k) * k = E(X)$$

□

Formule général

$$E(g(X)) = \int gdP_X$$

$$E(X^2) = \int x^2 dP_X = \begin{cases} \int x^2 f(x)dx \\ \sum k^2 P(X)k \end{cases}$$

Note. Le systeme représente soit l'un soit l'autre en fonction de densité ou discrète

Exemple 1.1. $P_X = \frac{1}{2}\delta_1 + \frac{1}{2}\frac{1}{\sqrt{\pi}}e^{-\frac{x^2}{2}}dx$

$$\begin{aligned}E(g(x)) &= \int gdP_X \\ &= \frac{1}{2}\int g\delta_1 + \frac{1}{2}\int g(x)\frac{1}{\sqrt{\pi}}e^{-\frac{x^2}{2}}dx \\ &= \frac{1}{2}g(1) + \frac{1}{2}\int g(x)\frac{1}{\sqrt{\pi}}e^{-\frac{x^2}{2}}dx\end{aligned}$$

1.3 Indépendance

$X \perp Y \Leftrightarrow E(f(X)g(Y)) = E(f(X))E(g(X))$

2 Loi des grand nombres et intervalles de confiance

BUT :

— $X, ..., X_n$ VA iid

$$\bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Que dire de la suite $(\bar{X_n})$?

— Si on cherche à estimer une quantité inconnu θ. On cherche I tel que :

$$P(\theta \in I) \geq 0.95.$$

Exemple 2.1. On a une pièce et on veut savoir si elle est équilibrée.
On modelise le résultat de cette pièce par une variable aléatoire de Bernouilli de paramètre p (p est inconnu)
0 $\rightarrow$ pile
1 $\rightarrow$ face

IL MANQUE DES TRUCS SQDGQRZG3A

2.1 Inégalité de concentration

Théorème 2.1 (Inégalité de Markov)**.** *Soit X une variable aléatoire définie sur un espace probabilisé (Ω, F, π).*

— *X est positive (X est à valeur dans R^+) (espace de dépard dans R+)*

— *X admet une espérance ($0 \leq E(X) \leq +\infty$)*

Alors on a :

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

Preuve : Soit X une va positive, d'espèrance finie. Soit $t > 0$:

$$P(X > t) = E(\mathbb{1}_{[t,+\infty[}(X))).$$

On sais que :

$$\forall x \geq 0, \forall t > 0, \mathbb{1}_{[t,+\infty[} \leq \frac{x}{t}\mathbb{1}_{[t,+\infty[} \leq \frac{x}{t}.$$

On peut dessiner les courbes pour comprendre l'inéquation.
On a donc :

$$\forall t > 0, \mathbb{1}_{[t,+\infty[} \leq \frac{X}{t}.$$

Par croissance de l'espérance, on a :

$$\forall t > 0, E(\mathbb{1}_{[t,+\infty[}) = P(X \geq t) \leq \frac{E(X)}{t}.$$

□

Théorème 2.2 (Inégalité de Tchebychev).

Note (Moment d'odre p). X admet un moment d'ordre 2

$$\Leftrightarrow E(|X|^2) < +\infty.$$

Moment d'ordre p :

$$\Leftrightarrow E(|X|^p) < +\infty.$$

Soit X une var définie sur un espace $(Omega, F, P)$ qui admet un moment d'ordre deux. On a :

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

Preuve : Soit X une var admettant un moment d'ordre 2.
Soit $Y = (X - E(X))^2$

— Y est une variable positive
— $E(Y) = VarX < +\infty$

D'après l'inégalité de Markov

$$P(Y \geq y) \leq \frac{E(Y)}{y} = \frac{VarX}{u}.$$

Soit $t > 0$

$$P(|X - E(X)| \geq t) \Leftrightarrow P((X - E(X))^2 \geq t^2).$$

D'après l'inégalité précédente avec $u = t^2 > 0$ on a :

$$P(|X - E(X)| \geq t) \leq \frac{VarX}{t^2}.$$

□

Exemple 2.2. Si $X \mathcal{N}(3, 5)$:

$$P(|X - 3| \geq 7) \leq \frac{VarX}{7^2} = \frac{5}{49}.$$

Exemple 2.3. Si $X \mathcal{E}(\lambda)$:

$$\forall t > 0, P(\left|X - \frac{1}{\lambda}\right| \geq t) \leq \frac{1}{\lambda^2 t^2}.$$

Exemple 2.4 (**MEILLEURS EXEMPLE GIGA IMPORTANT**). $X_1, ..., X_n$ va iid de loi $\mathcal{B}(p) p$ inconnu

$$\bar{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

$$\begin{aligned} E(\bar{X_n}) &= E(\frac{1}{n} \sum_{i=1}^{n} X_i) \\ &= \frac{1}{n} E(\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n} \sum_{i=1}^{n} p = p \end{aligned}$$

Que vaut la variance de $\bar{X_n}$?

$$\begin{aligned} Var\bar{X_n} &= Var(\frac{1}{n} \sum_{i=1}^{n} X_i) \\ &= \frac{1}{n^2} Var(\sum_{i=1}^{n} X_i) \\ &= \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i) \\ &= \frac{1}{n^2} n * (p * (1 - p)) \end{aligned}$$

$$Var(\bar{X_n}) = \frac{p(1-p)}{n}.$$

On applique l'inégalité de Tchebychev :

$$\forall t > 0, P(\left|\bar{X_n} - p\right| \geq t) \leq \frac{p(1-p)}{nt^2}$$
$$\leq \frac{1}{4nt^2}$$

On reviendra sur cette exemple dans deux définition

Définition 2.1 (Intervalle de confiance). Pour un modèle statistique $\Omega, F, (P_\theta)_{\theta\in\Theta}$ et une quantité à estimer $g(\theta)$, un intervalle de confiance de niveau $1-\alpha, \alpha \in [0,1]$ est un intervalle $I(X_1, ..., X_n)$ construit à partir de X_i tel que

$$\forall\theta \in \Theta, P(\theta \in I(X_1, \ldots, X_n)) \geq 1 - \alpha.$$

où $X_1, \dot{;} X_n$ sont des variables iid qui servent à construire l'estimateur

Exemple 2.5 (Retour sur l'exemple). Construison ensemble un intervale de confiance pour p de niveau 0.99 ($\alpha = 0.01$ d'erreur acceptée).
On sait que

$$\forall t > 0, P(\left|\bar{X_n} - p\right| > t) \leq \frac{1}{4nt^2}.$$

cela signifie qu'avec proba au moins $1 - \frac{1}{4nt^2}$ on a

$$\left|\bar{X_n} - p\right| < t$$
$$\Leftrightarrow p \in]\bar{X_n} - t, \bar{X_n} + t[$$
$$\Leftrightarrow P(p \in]\bar{X_n} - t, \bar{X_n} + t[) \geq 1 - \frac{1}{4nt^2}$$

Or on cherche avec un niveau $\alpha = 0.01$

$$\alpha = \frac{1}{2\sqrt{\alpha n}}.$$

$$\forall p \in [0,1], P(p \in]\bar{X_n} - \frac{1}{2\sqrt[2]{\alpha n}}; \bar{X_n} + \frac{1}{2\sqrt{\alpha n}}[).$$

Avec $\alpha = 0.01$ cela devient :

$$p \in]\bar{X_n} - \frac{1}{0.2\sqrt[2]{n}}; \bar{X_n} + \frac{1}{0.2\sqrt{n}}[.$$

Exemple 2.6. je veux une précision de 0.04 sur p. Comment choisir n? n est solution de

$$0.04 = \frac{1}{2\sqrt{\alpha n}} = \frac{1}{2 * 0.2 * \sqrt{n}}$$
$$0.4 * 0.04 = \frac{1}{\sqrt{n}}$$
$$\Leftrightarrow n = \frac{1}{(0.4 * 0.04)^2} \approx 4000$$

Définition 2.2 (Convergence en probabilité). Soit $(X_m)_{n\in N}$ une suite de va définie sur (Ω, F, P) . Soit X une va définie sur (Ω, F, P).
On dit que $(X_m)_{n\in N}$ converge en probabilité vers X si

$$\forall t > 0, P(|X_n - X| \geq t) \rightarrow_{n\to\infty} 0.$$

On le note $X_n \rightarrow^P_{n\to+\infty} X$

Théorème 2.3 (Loi faible des grands nombre). *Soit $(X_n)_{n\in N}$ une suite de va définie sur (Ω, F, P).*
Si X admet un moment d'ordre 2. Alors

$$\frac{1}{n}\sum_{x=1}^{n} X_i \rightarrow^P_{n\to+\infty} E(X_i).$$

C'est à dire :

$$\forall t > 0, P(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E(X)\right| \geq t) \rightarrow_{n\to+\infty} 0.$$

Preuve : Soit $(X_i)_{i\in N}$ une suite de v.a.r iid, admettant une espérance m et une variance.
Soit $\bar{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$

1. Calculons $E(\bar{X_n})$

$$\begin{aligned} E(\bar{X_n}) &= E(\frac{1}{n}\sum_{i=1}^n X_i) \\ &= \frac{1}{n}\sum_{i=1}^n E(X_i) \\ &= \frac{n}{n} * m = m \end{aligned}$$

2. Calculons $Var(\bar{X_n})$

$$\begin{aligned} Var(\bar{X_n}) &= Var(\frac{1}{n}\sum_{i=1}^n X_i) \\ &= \frac{1}{n^2} Var(\sum_{i=1}^n X_i) \\ &= \frac{1}{n^2}\sum_{i=1}^n Var(X_i) \\ &= \frac{n}{n^2} Var(X_i) = \frac{Var(X_1)}{n} \end{aligned}$$

D'après l'inégalité de Tchebytchev appliqué à $\bar{X_n}$ (qui possède une variance finie), on a :

$$\forall t > 0, P(\left|\bar{X_n} - E(X)\right| \geq t) \leq \frac{Var(X_1)}{nt^2} \longrightarrow_{n\to+\infty} 0.$$

□

Exemple 2.7. Si $(X_i)_{i\in N}$ est une suite iid de variable de la $\mathcal{B}(p)$. Que dire de $\frac{1}{n}\sum_{i=1}^n X_i$? $E(X_1) = p, Var(X_1) = p(1-p) < \infty$ D'après la loi faible des grands nombres :

$$\frac{1}{n}\sum_{i=1}^n X_i \to^{\mathcal{P}}_{n\to\infty} p.$$

Exemple 2.8. Si les X_i sont des va iid de la loi $\mathcal{N}(0,1)$, $E(X_1) = 0, Var(X_1) = 1$

$$\frac{1}{n}\sum_{i=1}^n X_i \to^{\mathcal{P}}_{n\to\infty} 0.$$

Si les X_i sont iid de la loi $\mathcal{E}(\lambda), \lambda > 0$ Ainsi

$$\frac{1}{n}\sum_{i=1}^n X_i \to^{\mathcal{P}}_{n\to\infty} \frac{1}{\lambda}.$$

Si les X_i de loi $\mathcal{B}(m,p)$

$$\frac{1}{n}\sum_{i=1}^n X_i \to^{\mathcal{P}}_{n\to\infty} mp.$$

Quelque exemple de convergence en probabilités

— (X_n) à X_n a une loi $\mathcal{B}(\frac{1}{n})$ Vers quoi X_n converge?
  Soit $t > 0$ fixé, $P(|X_n - 0| \geq t) = P(X_n = 1) = \frac{1}{n} \to_{n\to\infty} 0$ (Aussi faisable ave Tchebychev)

— Si X_n soit une v.a de la loi $\mathcal{E}(n)$

$$f_{X_n}(n) = ne^{-nx}\not{\mathbb{1}}_{x\geq 0}().$$

  On se doute que $X_n \to^{\mathcal{P}}_{n\to\infty} 0$

$$\forall t > 0, P(|X_n| \geq t) = \int_t^{+\infty} ne^{-xn}dx = \int_{tn}^{+\infty} e^{-u}du = e^{-tn} \to_{n\to\infty} 0.$$

Remarque. La convergence en probabilité est la notion de convergence la plus faible du cours.
Si $X_n \to^{\mathcal{P}}_{n\to\infty} X$, je ne sais pas, elle n'implique pas $E(X_n) \to_{n\to\infty} E(X)$

3 Loi forte des grands nombres

Définition 3.1 (Ensemble négligeable et ensembles presque-sûrs)**.** Soit (Ω, F, P) un espace probabilisé.
On dit que $A \in F$ est **négligeable** si $P(A) = 0$
On dit que $B \in F$ est **presque-sûr** si $P(B) = 1$

Exemple 3.1. $\Omega = \{0,1\}^{\mathbb{N}} F = P(\omega), A = \{w \in \Omega, \forall i \in N, w_i = 1\}$, $P(A) = 0$? (proba d'une infinité de pile ou face avec proba 1/2)

$$P(A_N) = \frac{1}{2^N}, P(A) = P(\bigcap_{N \in N}) = \lim_{n\to\infty} P(A_n) = 0.$$

Exemple 3.2. X v.a uniforme sur [0,1]

$$A = \{\frac{1}{2}\}, P(A) = P(X = \frac{1}{2}) = 0.$$

$$P(X \in]\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon[) = 2\epsilon = \int_{1/2-\epsilon}^{1/2+\epsilon} 1dx.$$

Remarque. Un ensemble négligeable n'est a priori pas vide.

Théorème 3.1. *Soit (A_n) une suite d'ensemble négligeable.*
- $\bigcup_{n\in N} A_n$ *est un ensemble négligeable*
- $\bigcap_{n\in N} A_n$ *est un ensemble négligeable.*

Preuve :
- $0 \leq P(\bigcup_{n\in N}) \leq \sum_{n\in N} P(A_n) = 0$
- $0 \leq P(\bigcap_{n\in N}) \leq \sum_{n\in N} P(A_n) = 0$ car $\bigcap_{n\in N} A_n \subset A_1$

□

Corollaire (immédiat)**.** *Si les $(B_n)_{n\in N}$ sont presque sur :*
- $\bigcap_{n\in N} B_n$ *est presque-sûrs*
- $\bigcup_{n\in N} B_n$ *est presque-sûrs*

Définition 3.2 (Convergence presque-sûr)**.** Soit $(X_n)_{n\in N}$ une suite de variables aléatoire définis sur (Ω, F, P)
Soit X une variable aléatoire définie sur (Ω, F, P). On dit que (X_n) converge presque-surement vers X si l'événement $\{w \in \Omega, X_n(w) \to_{n\to+\infty} X(w)\}$ est presque-sûrs.
Pour $w \in \Omega$ fixé

$$X_n(w) \to_{n\to+\infty} X(w).$$

$$P(X_n(w) \to_{n\to+\infty} X(w)) = 1.$$

$$\forall \epsilon > 0, \exists N \in N, \forall n \geq N, |X_n(w) - X_n(w)| < \epsilon.$$

On le note $X_n \to_{n\to+\infty}^{(} p.s)X$

Proposition 3.2.
- *Si $X_n \to_{n\to+\infty}^{p.s} X$ et si f est une* ***fonction continue*** *sur $\mathbb{R}$ alors :*

$$f(X_n) \to_{n\to+\infty}^{p.s} f(X).$$

- *Si $X_n \to_{n\to+\infty}^{p.s} X$ et $Y_n \to_{n\to+\infty}^{p.s} Y$ alors :*

$$X_n + Y_n \to_{n\to+\infty}^{p.s} X + Y.$$

$$X_n * Y_n \to_{n\to+\infty}^{p.s} X * Y.$$

Définition 3.3 (Loi forte des grands nombres)**.** Soit (X_n) une suite de variable aléatoire iid, définies sur (Ω, F, P), avec $\mathbb{E}(|X_i|) > +\infty$. Alors

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to_{n\to+\infty}^{p.s} E(X_i).$$

Nouveau cours du 29/09
Rappel :
Nous avons définie un nouvelle notion de convergence : la convergence presque sûr.

$$x_n \to^{p.s} X \text{ si } P(\{w \in \Omega, X_n(w) \to_{n\to+\infty}\}) = 1.$$

Ce n'est **pas** la convergence en probabilité :

$$X_n \to^{\mathbb{P}} X \text{ si } \forall \epsilon > 0, P(|X_n - X| \geq \epsilon) \to_{n\to+\infty} 0.$$

La convergence p.s est plus difficile à établir que la convergence en probabilité.
Si $X_n \to^{p.s}_{n\to\infty} \Rightarrow X_n \to^{\mathbb{P}} X$

La loi forte des grands nombres :
Voir la def.
Que veux dire avoir un moment d'ordre 1 :

$$\mathbb{E}(|X|) < +\infty.$$

C'est **LA** condition sous laquelle $\mathbb{E}(X)$ a un sens. Par exemple l'espérance d'une loi de cauchy n'existe pas. Pour vérifier ça :
— On connaît nos espérances de loi
— Ou on le montres avec la définition
— OU AVEC LA PROP SUIVANTE

Proposition 3.3. *Si X est bornée, alors X admet des moments de tous ordres*

$$\exists M, |X|^p \leq M \Rightarrow E(|X|) \leq E(M^p) = M^p < +\infty.$$

FIN DU RAPPEL

4 Application de la loi forte des grands nombres

Exemple 4.1. Si $X_1, \ldots, X_n$ sont des variable indépendante iid de loi $\mathcal{B}(p), p \in]0,1[$. Comme X_i admet une espérance, alors d'après la loi forte des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to^{p.s}_{n\to+\infty} p = E(X_1).$$

Exemple 4.2. Si $(X_i)_{i\in N}$ est une suite de v.a iid de loi $\mathcal{E}(\lambda), \lambda > 0$. X_1 admet un moment d'ordre 1 et d'après la loi forte des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to^{p.s}_{n\to+\infty} E(X_1) = \frac{1}{\lambda}.$$

Exemple 4.3. Si $(X_i)_{i\in N}$ est une suite de v.a iid de loi $\mathcal{N}(3,8)$. X_1 admet un moment d'ordre 1 donc d'après la loi forte des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to^{p.s}_{n\to+\infty} E(X_1) = 3.$$

Remarque (sur l'estimation). En stat on part d'un modèle statistique $(\Omega, \mathcal{F}, (\mathbb{P}_\theta)_{\theta\in\Theta})$ (P_θ représente une famille de loi qui décrivent notre problème). Le modèle statistique, c'est la famille de loi qu'on considère dans le modèle pratique.
On va chercher à retrouver $\theta \in \Theta$ à partir d'observation (ou d'autre quantité liées au problème)

4.1 Estimation de moyenne

$(X_i)_{i\in N}$ famille iid de la .. admettant un moment d'ordre 1. Si mon but est de retrouver la moyenne m de cette loi on introduit :

$$\bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \to^{p.s}_{n\to+\infty} E(X_i) = m.$$

4.2 Estimation du moment d'ordre p

$(X_i)_{i\in N}$ sont v.a iid de la m..., admettant un moment d'ordre p . Si mon but est de retrouver/estimer ce moment noté m_p

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i^p.$$

Si on pose $Y_i = X_i^p$ la suite $(Y_i)_{i\in N}$ est une suite de va iid admettant un moment d'ordre 1 (car $E(|Y_i|) = E(|X_i|^p) < +\infty$). D'après la loi forte des grand nombres

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i^p \rightarrow^{p.s}_{n\rightarrow+\infty} E(Y_i) = E(X_i^p) = m_p.$$

4.3 Estimation de la variance en connaissant E(X1)

Soit $(X_i)_{i\in N}$ sont v.a iid de la m..., admettant un moment d'ordre 2 . Je cherche à estimer $E((X - E(X_i)^2)) = Var(X_1)$. **Si je connais** $m = E(X_1)$ Soit $V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2$.
On pose $Y_i = (X_i - m)^2$.
La suite $(Y_i)_{i\in N}$ est une suite de va iid admettant une espérance. En effet

$$\begin{aligned} E(|Y_1|) &= E(\left|(X-m)^2\right|) \\ &= E((X_1 - m)^2) \\ &= E((X_1 - E(X_1)^2)) \\ &= Var(X_i) \end{aligned}$$

D'après la loi forte des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n}(X_1 - m)^2 \rightarrow^{p.s}_{n\rightarrow+\infty} Var(X_i).$$

4.4 Estimation de la variance sans connaître E(X1)

On remplace m par $\bar{X_n}$
On pose $V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X_n})^2$ où $\bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$
Horreur ! les $Y_i = (X_i - \bar{X_n})$ ne sont pas des variable indépendante. On réécrit :

$$\begin{aligned} V_n &= \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X_n}) \\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\bar{X_n} + \bar{X_n}^2) \\ &= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \frac{2}{n}\sum_{i=1}^{n}X_i\bar{X_n} + \frac{1}{n}\sum_{i=1}^{n}\bar{X_n}^2 \\ &= \text{''} \rightarrow^{p.s}_{n\rightarrow+\infty} E(X_1^2) - \frac{2\bar{X_n}}{n}\sum_{i=1}^{n}X_i + \frac{1}{n}\sum_{i=1}^{n}\bar{X_i}\text{''} \end{aligned}$$

Or

$$\frac{2}{n}\sum_{i=1}^{n}X_i\bar{X_n} = 2\bar{X_n}^2.$$

Et :

$$\frac{1}{n}\sum_{i=1}^{n}\bar{X_n}^2 = \bar{X}_n^2.$$

On a alors

$$V_n = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X_n}^2.$$

Ainsi

$$V_n \to^{p.s}_{n\to+\infty} E(X_1^2) - E(X_1)^2 = Var(X_1).$$

Nouveau cours du 07/10

4.5 Estimateur classiques

4.5.1 Moyenne empirique

$(X_i)_{i\in N}$ famille de v.a **i.i.d**, admettant un moment d'ordre 1. Alors :

$$\bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} \to^{p.s}_{n\to\infty} E(X_1).$$

Biais :

$$E(\bar{X_n}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{n}{n}E(X_1) = E(X_1).$$

4.5.2 Estimateur du moment d'ordre p

$(X_i)_{i\in N}$ famille de v.a **i.i.d**, admettant un moment d'ordre $p \in N$. Alors

$$M_n = \frac{1}{n}\sum_{i=1}^{n} X_i^p \to^{p.s}_{n\to\infty} E(X_1^p).$$

Biais :

$$E(M_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^p) = \frac{n}{n}E(X_1^p) = E(X_1^p).$$

Pourquoi cette convergence ? $\forall i \in N$, on pose $Y_i = X_i^p$. $(Y_i)_{i\in N}$ est une suite de v.a. i.i.d.

$$E(|Y_i|) = E(|X_i^p|) = E(|X_i|^p) < +\infty.$$

D'après la loi forte des grands nombre (appliqué à la suite de v.a. i.i.d. $(Y_i)_{i\in N}$ admettant un moment d'ordre 1), on a

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n} X_i^p \to^{p.s}_{n\to\infty} E(Y_i) = E(X_i^p).$$

4.5.3 Estimateur de la variance

1. **A moyenne connue** :
   Soit $(X_i)_{i\in N}$ un suite de v.a **i.i.d**, admettant un moment d'ordre 2 (car on veux calculer une variance) de moyenne/espérance m connue. Alors

   $$V_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2 \to^{p.s}_{n\to\infty} Var(X_1).$$

   Biais :

   $$E(V_n) = \frac{1}{n}\sum_{i=1}^{n}(E(X_i - m)^2) = \frac{n}{n}Var(X_1) = Var(X_1).$$

2. **A moyenne inconnue**
   Soit $(X_i)_{i\in N}$ un suite de v.a **i.i.d**, admettant un moment d'ordre 2 (car on veux calculer une variance). Alors

   $$V_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X_n})^2 \to^{p.s}_{n\to\infty} Var(X_1).$$

   Biais :

   $E(V_n) = Var(X_1)$ (voir le cours de Thiam).

Comment le montrer ? On pose :

$$W_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X_n})^2 = \frac{n-1}{n}V_n.$$

On ne peut pas poser $Y_i = (X_i - \bar{X_n})^2$ car ce n'est pas indépendant ! ! On ne peut pas appliquer la loi forte des grand nombre avec ça
Montrons que $W_n \rightarrow_{n\rightarrow\infty}^{p.s} Var(X_i)$

$$\begin{aligned}
W_n &= \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\bar{X_n} + \bar{X_n}^2) \\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \frac{2}{n}\sum_{i=1}^{n}X_i\bar{X_n} + \frac{1}{n}\sum_{i=1}^{n}\bar{X_n}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X_n}(\frac{1}{n}\sum_{i=1}^{n}X_i) + \frac{1}{n}\sum_{i=1}^{n}\bar{X_n}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X_n}^2 + \frac{1}{n}\sum_{i=1}^{n}\bar{X_n}^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X_n}^2 + \bar{X_n}^2 \\
W_n &= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X_n}^2
\end{aligned}$$

Or, par ce point a), $\bar{X_n} \rightarrow_{n\rightarrow\infty}^{p.s} E(X_1)$ donc $\bar{X_n}^2 \rightarrow_{n\rightarrow\infty}^{p.s} E(X_1)^2$ car la fonction x^2 continue.
Or, par le point b) pour $p = 2$, $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \rightarrow_{n\rightarrow\infty}^{p.s} E(X_1^2)$ (car les X_i sont iid et admettent un moment d'ordre 2)

Estimateur de la variance :
Ainsi, $W_n \rightarrow_{n\rightarrow\infty}^{p.s} E(X_1^2) - E(X_1)^2 = Var(X_1)$.

$$V_n = \frac{n}{n_1}W_n \rightarrow_{n\rightarrow\infty}^{p.s} Var(X_1) \text{ car } \frac{n}{n-1} \rightarrow 1.$$

4.5.4 Estimation de la fonction de répartition

$(X_i)_{i\in N}$ une suite de v.a. i.i.d.
La fonction de répartion empirique est

$$F_n : t \mapsto \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{X_i\leq t}.$$

Pour tout $t \in R$ fixé, on a :

$$F_n(t) \rightarrow_{n\rightarrow\infty}^{p.s} F_{X_1}(t) = P(X_1 \leq t).$$

Pour $t \in R$ fixé, on pose

$$Y_i = \mathbb{1}_{X_i\leq t} = \begin{cases} 1 \text{ si } X_i \leq t \\ 0 \text{ sinon} \end{cases}.$$

La suite $(Y_i)_{i\in N}$ est une suite de v.a. i.i.d.. Y_i est une variable de Bernouilli.

$$E(Y_1) = E(\mathbb{1}_{X_1\leq t}) = P(X_1 \leq t) = F_{X_1}(t).$$

$(Y_i)_{i\in N}$ est une suite de v.a iid admettant un moment d'ordre 1 (car ce sont des Bernoulli). Ainsi d'après la loi forte des grands nombres

$$\frac{1}{n}\sum_{i=1}^{n}Y_i = F_n(t) \rightarrow_{n\rightarrow\infty}^{p.s} E(Y_i) = F_{X_1}(t).$$

Théorème 4.1 (Théorème de Glivenko-Contelli, théorème fondamentale de la statistique)**.** *Soit* $(X_i)_{i\in N}$ *une suite de v.a. i.i.d. Soit*

$$\begin{aligned} F_n : & R \to [0,1] \\ & t \mapsto \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \leq t} \end{aligned}$$

Alors :

$$P(\|F_n - F_{X_1}\|_\infty \to_{n\to\infty} 0) = 1.$$

Reformulons : Presque-sûrement, la suite de fonction $(F_n)_{n\in N}$ *converge uniformément vers* F_{X_1}.

Rappel : $F_n \to^{CV}_{n\to+\infty} F$ si $\|F_n - F\|_\infty \sup |F_n(x) - F(x)| \to_{n\to+\infty} 0$ Si on aime les quantificateurs :

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall t \in R, |F_n(t) - F(t)| < \epsilon.$$

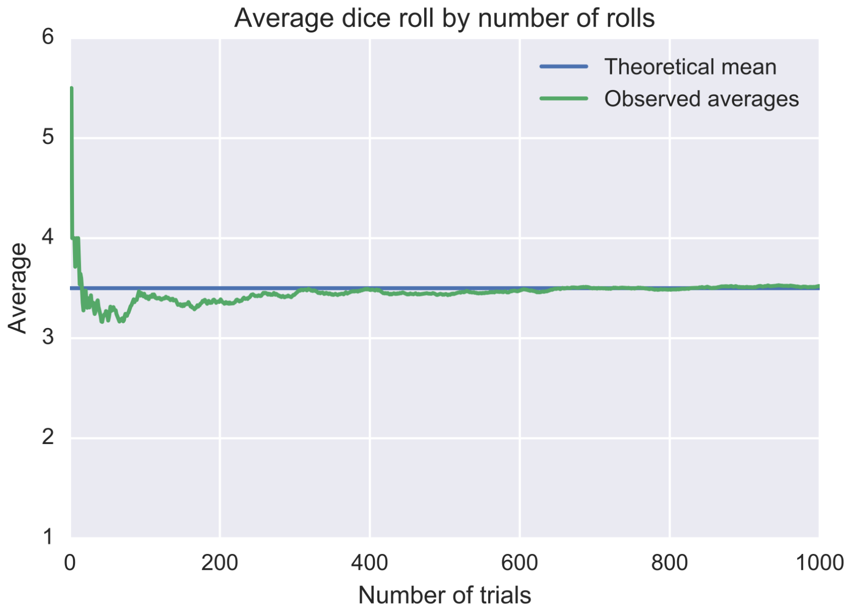

Figure 1 – Illustration de la loi des grands nombres

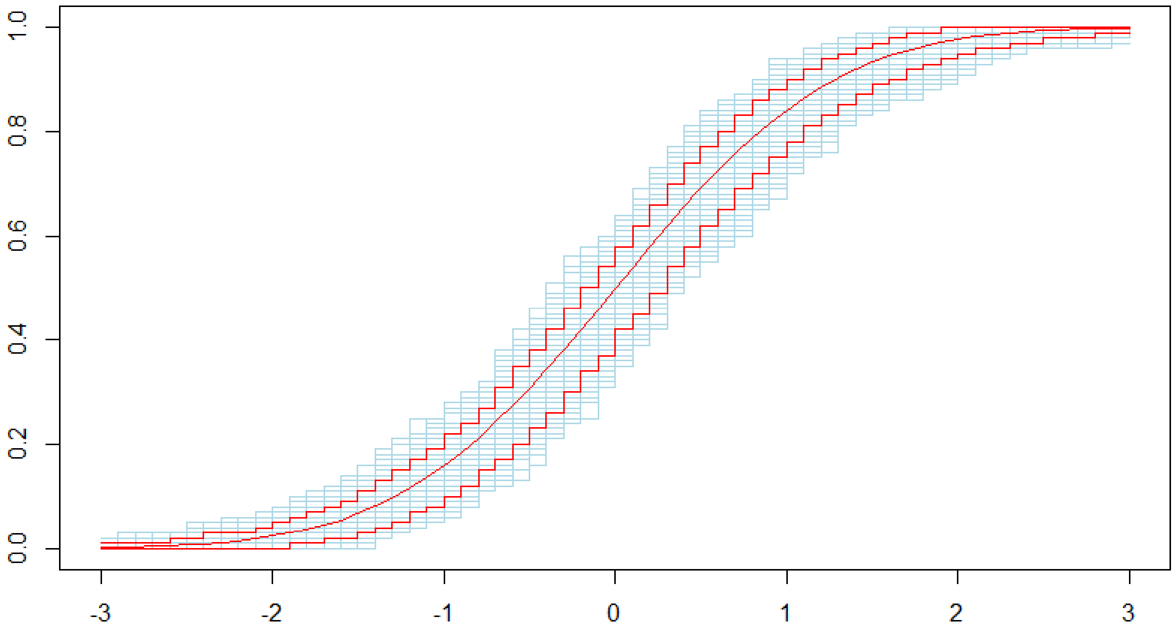

Figure 2 – Illustration du théorème de Glivenko-Contelli