Théorèmes limites

Sommaire

1 Couple de v.a.r

Définition *Soit X_1, X_2 deux v.a.r. définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.*
On appelle couple aléatoire, l'application définie par

$$\begin{aligned} X : \quad & \Omega \longrightarrow \mathbb{R}^2 \\ & \omega \longmapsto (X_1(\omega), X_2(\omega)) \end{aligned} \tag{1}$$

On note $X = (X_1, X_2)$.

1.1 Définitions et Propriétés

Définition *On appelle fonction de répartition du couple de v.a.r. $X = (X_1, X_2)$ l'application*

$$\begin{aligned} F_X : \quad & \mathbb{R}^2 \longrightarrow \mathbb{R} \\ & (x_1, x_2) \longmapsto \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \end{aligned} \tag{2}$$

avec $\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = \mathbb{P}(\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\})$

Définition *Le couple de v.a.r. $X = (X_1, X_2)$ est dit discret s'il est à valeurs dans un sous ensemble fini ou dénombrable de $\mathbb{R}^2$*
Le couple de v.a.r. $X = (X_1, X_2)$ est dit absolument continue s'il admet une représentation intégrale dans $\mathbb{R}^2$

Définition *On appelle espérance du couple aléatoire $X = (X_1, X_2)$ l'élément de $\mathbb{R}^2$ définie par*

$$\mathbb{E}(X) = \mathbb{E}(X_1, X_2) = (\mathbb{E}(X_1), \mathbb{E}(X_2))$$

Si $\mathbb{E}(X_1)$ et $\mathbb{E}(X_2))$ existent il en sera de même pour $\mathbb{E}(X)$.

Définition *Soit $X = (X_1, X_2)$ un couple de v.a.r, si $\mathbb{E}(X_1)$ et $\mathbb{E}(X_2))$ existent, on appelle covariance du couple aléatoire $X = (X_1, X_2)$ le réel*

$$Cov(X_1, X_2) = \mathbb{E}(X_1.X_2) - \mathbb{E}(X_1).\mathbb{E}(X_2))$$

notons que $Cov(X_1, X_1) = Var(X_1)$

1.2 Indépendance

Définition *Deux v.a.r. X et Y sont dites indépendantes si*

$$\begin{aligned} & \forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \\ & \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x).\mathbb{P}(Y \leq y) \end{aligned} \tag{3}$$

Théorème *Soit X et Y deux v.a.r. indépendantes, on a*

$$\mathbb{E}(X.Y) = \mathbb{E}(X).\mathbb{E}(Y)$$

Définition *On appelle coefficient de corrélation de deux v.a.r. X et Y le réel défini par*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}$$

où $\sigma_X = \sqrt{Var(X)} = \sqrt{Cov(X,X)}$.
Ce taux mesure le degré de "dépendance" de deux v.a.r.

2 Théorèmes limites

On s'intérèsse à un évènement E de probabilité inconnue $\mathbb{P}(E)$.
On veut une estimation de $\mathbb{P}(E)$.
Si l'on répète N fois l'expérience obtient t'on "en moyenne" une approximation de $\mathbb{P}(E)$?
Que se passe t'il lorsque N tend vers l'infini ?
Exemple
Nous avons un dé truqué et nous voulons savoir quelles sont les nouvelles proba.

2.1 Loi faible des grands Nombres

Théorème *Soient $X_1, X_2, \ldots, X_n$, n v.a.r. de même loi, de même espérance m, de même variance σ^2 et indépendantes (mutuellement). Alors la v.a.r. $\bar{X}_n = \frac{1}{N}\sum_{i=1}^{n} X_i$ converge en probabilité vers m. On écrira*

$$\bar{X}_n \xrightarrow{\mathbb{P}} m$$

Convergence en probabilité : pour n grand $\bar{X}_n$ est **probablement** très proche de m Cette loi confirme notre intuition : une fréquence tend à se stabiliser lors d'épreuves répétées.
On peut le vérifier expérimentalement.
Plus précisément : lorsqu'on répète un grand nombre de fois une expérience aléatoire, la fréquence d'apparition d'un évènement est "proche" de sa probabilité.
C'est cette "convergence en probabilité" qui explique par exemple comment retrouver, en sondant un échantillon, la structure d'une population.

2.2 Loi forte des grands Nombres

Théorème *Soient $X_1, X_2, \ldots, X_n$, n v.a.r. de même loi, de même espérance m et indépendantes. Alors la v.a.r. $\bar{X}_n = \frac{1}{N}\sum_{i=1}^{n} X_i$ converge* **presque sûrement** *vers m. On écrira*

$$\bar{X}_n \xrightarrow{p.s.} m$$

Convergence en p.s. : pour n grand $\bar{X}_n$ est "presque partout" très proche de m **Loi forte ou Loi faible ?**

Ces deux lois nous assurent que $\bar{X}_n$ tend vers m lorsque n tend vers l'infini.
La loi faible nous assure que pour n assez grand $\bar{X}_n$ est assez proche de m mais elle n'exclut pas que pour des valeurs p plus grandes que n, $\bar{X}_n$ s'éloigne de m.
En effet il se peut que pour certaines valeurs (dont la probabilité collective est très faible) $\bar{X}_n$ s'éloigne de m.
La loi forte exclut cette éventualité.

2.3 Théorème Central Limite : de la limite centrale

Théorème *Soient $X_1, X_2, \ldots, X_n$, n v.a.r. indépendantes de même loi, d'espérance m et de variance σ^2 alors la v.a.r. $S_n \sum_{i=0}^{n} X_i$ vérifie*

$$\frac{S_n - n.m}{\sigma\sqrt{n}} \Longrightarrow Y \quad avec \ Y \sim \mathcal{N}(0,1)$$

c'est à dire que la v.a.r. $\frac{S_n - n.m}{\sigma\sqrt{n}}$ suit pour n grand une loi normale centrée réduite. C'est le théorème le plus important en probabilité.
Les v.a.r. X_i suivent **une** lois quelconque (la même pour toutes les v.a.r.).
Ce théorème permet de donner des approximations pour le calcul des probabilités d'évènement faisant intervenir une somme de v.a.r.
Il explique aussi pourquoi beaucoup de phénomènes aléatoires naturels admettent des distributions normales.

Proposition *Soient $X_1, X_2, \ldots, X_n$ n v.a.r. indépendantes de loi de Bernoulli de paramètre p. La v.a.r. $S_n = \sum_{i=0}^{n} X_i$ suit une loi binomiale $\mathcal{B}(n,p)$. et S_n "tend" vers une loi normale de paramètres $(np, \sqrt{np(1-p)}$ lorsque n tend vers l'infini.*

En pratique en approche une loi binomiale $\mathcal{B}(n,p)$ par une loi normale de paramètres dès que $np \approx 20$

Proposition *Soient $X_1, X_2, \ldots, X_n$ n v.a.r. indépendantes de loi de Poisson de paramètre λ. La v.a.r. $S_n = \sum_{i=0}^{n} X_i$ suit une loi de Poisson de paramètre $n\lambda$ et S_n "tend" vers une loi normale de paramètres $(n\lambda, \sqrt{n\lambda}$ lorsque n tend vers l'infini.*

En pratique en approche une loi de Poisson $\mathcal{P}(\lambda)$ par une loi normale dès que $np \approx 20$

3 Statistiques

Problème : évaluer une quantité α déterministe à partir d'un échantillon.
L'échantillon $(X_1, \ldots, X_n)$ de taille à déterminer pour satisfaire une exigence de "fiabilité", on parlera de **Niveau ou seuil de confiance**
Problème dit d'estimation ponctuelle.

Définition *Soient $X_1, \ldots, X_n$ une suite de V.A. indépendantes admettant une moyenne μ et un écart type σ on note $\bar{X}$ la* **moyenne empirique** *de l'échantillon*

$$\bar{X} = \frac{1}{n}\sum_{i=1} nX_i$$

Définition *Soient $X_1, \ldots, X_n$ une suite de V.A. indépendantes admettant une moyenne μ et un écart type σ on note S^2 la* **variance empirique** *de l'échantillon*

$$S^2 = \frac{1}{n}\sum_{i=1} n(X_i - \bar{X})^2$$

Moyenne et variance empirique sont des v.a. Si n est grand, d'après le théorème central-limite

$$\sqrt{n}(\bar{X} - \mu)/\sigma \longrightarrow \mathcal{N}(0,1)$$

En pratique σ est **inconnu** on le **remplace** par S, d'où

$$\mathbb{P}(\mu \in [\bar{X} - t\frac{S}{\sqrt{n}}, \bar{X} + t\frac{S}{\sqrt{n}}])$$

et donc

$$\begin{aligned} & \mathbb{P}(\mu \in [\bar{X} - t\frac{S}{\sqrt{n}}, \bar{X} + t\frac{S}{\sqrt{n}}]) \\ = \quad & \mathbb{P}(|\sqrt{n}\frac{\bar{X} - \mu}{S}| \leq t) \\ \approx \quad & 2\Phi(t) \end{aligned}$$

où

$$\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_0^t exp(-x^2/2)$$

En pratique
On choisit un **niveau ou seuil de confiance** α (en général $\alpha = 0.95$ ou $\alpha = 0.99$.
On estime grâce à la table ou par un calcul approché t_α, par exemple $t_{0.95} \approx 1.96$.
On peut alors affirmer

$$\mathbb{P}(\mu \in [\bar{X} - t_\alpha\frac{S}{\sqrt{n}}, \bar{X} + t_\alpha\frac{S}{\sqrt{n}}]) \approx \alpha$$

L'intervalle

$$[\bar{X} - t_\alpha\frac{S}{\sqrt{n}}, \bar{X} + t_\alpha\frac{S}{\sqrt{n}}]$$

est appelé intervalle de confiance pour μ au seuil (niveau) de confiance α
la v.a. $\bar{X}$ est dite estimateur de μ Test d'hypothèse : on veut tester une hypothèse, par exemple équiprobabilité des lancés d'un dé.
Si X est une v.a. de moyenne μ on veut savoir si $\mu = \mu_0$ (appelée hypothèse nulle).
L'outil mathématique nous aide à prendre la décision : rejeter l'hypothèse ou pas.
Si on rejette l'hypothèse on a lhypothèse alternative, ici $\mu \neq \mu_0$. Deux erreurs possible :
1/ On rejette l'hypothèse alors qu'elle est vraie : erreur de première espèce dont la probabilité est appelée risque de première espèce (noté α).
2/ On accepte l'hypothèse alors qu'elle est fausse : erreur de deuxième espèce dont la probabilité est appelée risque de deuxième espèce (noté β, $1-\beta$ est appelé puissance du test).
On souhaite bien sûr minimiser les deux risques, souvent dans la pratique minimiser l'un entraine une augmentation de l'autre. D'après l'intervalle de confiance vu précédemment

$$\mathbb{P}(\mu \in [\bar{X} - t_\alpha\frac{S}{\sqrt{n}}, \bar{X} + t_\alpha\frac{S}{\sqrt{n}}]) \approx \alpha$$

d'où

$$\mathbb{P}(\mu \in [\mu_0 - t_\alpha\frac{S}{\sqrt{n}}, \mu_0 + t_\alpha\frac{S}{\sqrt{n}}])$$

représente (si on accepte l'hypothèse) le risque de première espèce $1 - \alpha$. La note obtenue par des étudiants à un examen est une v.a.r. normale $X \sim \mathcal{N}(7, 3^2)$.

1. Calculer le pourcentage d'individus ayant plus de 10, et la note en dessous de laquelle se trouvent 10 % des étudiants.

On cherche $\mathbb{P}(X \geq 10)$, il suffit de centrer et réduire pour se ramener à une v.a. normale centrée réduite (on la notera dans cet exercice$Y$)

$$\begin{aligned} \mathbb{P}(X \geq 10) &= \mathbb{P}(\frac{X-7}{3} \geq \frac{10-7}{3}) \\ &= \mathbb{P}(Y \geq 1) \\ &= 1 - \mathbb{P}(Y \leq 1) \\ &= 1 - F_Y(1) \end{aligned}$$

On consulte la table

environ 16% des étudiants ont plus de 10. on a donc $\mathbb{P}(Y \leq \frac{x_0-7}{3}) = 0.1$, en consultant la table on trouve (on cherche plutot x_1 tel que $\mathbb{P}(Y \leq \frac{x_1}{)} = 0.9$, on aura alors $\frac{x_0-7}{3} = -x_1$, vous pouvez vous en convaincre sur le graphique de la loi normale)

$$\frac{x_0 - 7}{3} = -1.28$$

on en déduit alors x_0.

Soit p la proportion de la population favorable à un candidat.
On souhaite grâce à un sondage avoir une estimation de p.
On intéroge un échantillon de n personnes et on prend comme approximation de p le nombre $\tilde{p}$ de personnes favorables à ce candidat dans l'échantillon.
Déterminer, grâce au théorème de la limite centrée, la taille de l'échantillon pour que $\tilde{p}$ approche p à 10^{-2} prés avec une probabilité de 0,95. On souhaite, à taille d'échantillon fixée, trouver la précision obtenue. calculer

1. la précision ϵ sachant que l'échantillon est de taille $n = 1000$ et pour un seuil de confiance de $0,95$ (autrement dit qu'elle est la précision obtenue dans 95% des cas si l'échantillon est de taille $n = 1000$).
2. la précision ϵ sachant que l'échantillon est de taille $n = 1000$ et pour un seuil de confiance de $0,99$

On a lancé (Weldon) un dé 315 672 fois et tiré 106 602 fois l'une des faces 5 ou 6.

1. Calculer les fréquences théorique et observée d'apparition des faces 5 ou 6.
2. Calculer la fluctuation (différence entre les fréquences observée et théorique)
3. Quelle loi représente cette expérience aléatoire.
4. En utilisant le théorème central-limite, montrer que l'hypothèse d'un dé équilibré est à rejeter

test du χ^2 (imaginé par Pearson)
But : déterminer si une v.a. X suit une loi particulière.
On souhaite donc tester l'hypothèse :
H : X suit une loi $\mathcal{L}$.
On calcule les effectifs n_i extraits de notre échantillons de taille n
d'après le théorème central limite n_i/n tend vers p_i (probabilité théorique), plus précisément :

$$T = \sum \frac{(n_i - np_i)^2}{np_i} \sim \chi^2$$

C'est cette "tendance" qui nous fournira le test.