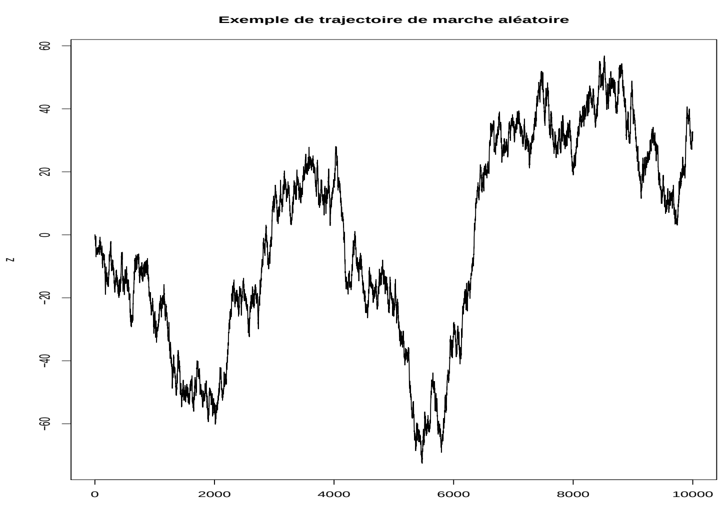
Université Paris I, Panthéon - Sorbonne

LICENCE M.I.A.S.H.S.

Cours de Probabilités

JEAN-MARC BARDET (UNIVERSITÉ PARIS 1, SAMM)



U.F.R. 27 et SAMM (Statistique, Analyse et Modélisation Multidisciplinaire)
Université Panthéon-Sorbonne, 90 rue de Tolbiac, 75013 Paris.

Plan du cours

Introduction

1. Quelques rappels de statistiques descriptives
2. Espace de probabilité, événements et indépendance
3. Variables aléatoires
4. Théorèmes limite et statistique inférentielle

References

[1] Dauxois, J et Hassenforder, C. (2004). Toutes les probabilités et Statistiques. Cours et Exercices corrigés.Ellipses.

[2] Deledicq, JC. (2000). Faîtes vos jeux! 50 grands classiques du calcul des probabilités, ACL Les éditions du kangourou

[3] Engel, A. (1990). Les certitudes du hasard, Alea éd.

[4] Leboeuf, C., Guegand, J., Roque, J.L. et Landry, P. Cours de Probabilités et de statistiques, Ellipses

[5] Leboeuf, C.,Guegand, J.,Roque, J.L. et Landry, P. Exercices corrigés de probabilités, Ellipses

[6] Ross, S.M (2007). Initiation aux probabilités, Enseignement des Mathématiques. Presses polytechniques et universitaires romandes.

[7] Saporta, G. Probabilités, analyse des données et statistique (2nd édition), éditions Technip.

Documents accessibles librement sur internet

- Des cours sur le site STAFAV: `https://www.math.u-psud.fr/ stafav/`.
- Un site de Paris V: `http://helios.mi.parisdescartes.fr/ glaunes/Proba3/CoursProbasStats.pdf`.
- Le site du LPSM de Paris VI: `https://www.lpsm.paris/formation/supports-de-cours/`. Notamment le cours de J. Bertoin.
- Un poly de maths générales (avec probas) très bien fait: `https://www.math.univ-toulouse.fr/ barthe/L1math4/`
- Le premier cours de probabilités de l'Ecole Polytechnique: `http://josselin-garnier.org/wp-content/uploads/`

Introduction

- Les maths sont partout!
- La statistique et les probabilités: qu'est-ce et pour quoi faire?
- Nécessité de savoir aussi d'autres maths (algèbre, géométrie, analyse,...)

1 Statistique descriptive

1.1 Introduction

Contenu, buts et raisons d'être de la statistique.
Statistique descriptive et inférentielle.

1.2 Statistique unidimensionnelle

Définition.
- *Variable (caractère) X.*
- *Variable quantitative.*
- *Variable qualitative $\Longrightarrow$ modalités (classes).*

Définition. *Soit X une variable et une population de n individus.*

- *Répartition en k classes.*
- *Effectifs, fréquences.*
- *Diagramme à baton, diagramme circulaire ("camembert").*

Définition. *Soit X une variable quantitative.*

- *Mode, amplitude de classe, densité d'effectif, densité de fréquence.*
- *Histogramme.*
- *Fonction de répartition empirique, polygône des fréquences cumulées.*
- *Médiane, quantiles.*
- *Moyenne, variance, écart-type.*

2 Espace de probabilité, événements et indépendance

2.1 Espace de probabilité

Définition.
- *Expérience aléatoire.*
- *Évènement élémentaire, ensemble fondamendal (univers).*
- *Événement et tribu.*

Définition.
- *Intersection, union.*
- *Évènements incompatibles.*
- *Évènement contraire.*

2.2 Mesure de probabilité d'un événement

Définition. *Une probabilité sur* $(\Omega, \mathcal{A})$ *est une application* $P\colon \mathcal{A} \to [0,1]$*, qui à un événement* $E \in \mathcal{A}$ *associe le réel* $P(E) \in [0,1]$ *et telle que:*

- $P(\Omega) = 1$.
- *Pour* $(E_i)_{i\in J}$ *événements de* $\mathcal{A}$*, incompatibles,* $P(\bigcup_{i\in J} E_i) = \sum_{i\in J} P(E_i)$.

Propriété.

- $P(\overline{E}) = 1 - P(E)$ *pour* $E \in \mathcal{A}$.
- $P(\emptyset) = 0$.
- *si* $A \subset B$*,* $0 \leq P(A) \leq P(B) \leq 1$.
- $P(E \cup F) + P(E \cap F) = P(E) + P(F)$*, pour* $(E, F) \in \mathcal{A}^2$.

Définition.

- *Cardinal d'un ensemble fini.*
- *Equiprobabilité (ou probabilité uniforme) sur un ensemble fini.*

Propriété. *Si* Ω *est fini et si* P *est une probabilité uniforme sur* $(\Omega, \mathcal{A})$*, alors pour tout* $E \in \mathcal{A}$*,* $P(E) = \dfrac{card(E)}{card(\omega)}$.

Remarque.

Pour calculer le cardinal d'un ensemble, on peut utiliser les résultats combinatoires suivants: on considère un ensemble de n éléments et on tire k éléments:

- S'il y a remise, et que l'ordre compte, un tirage est un k-uplet, et le nombre total de k-uplets est: n^k.
- S'il n'y a pas remise, et que l'ordre compte, un tirage est un arrangement, et le nombre total d'arrangements est: $A_n^k = n(n-1)..(n-k+1) = n!/k!$.
- S'il n'y a pas remise, et que l'ordre ne compte pas, un tirage est une combinaison, et le nombre total de combinaisons est: $C_n^k = A_n^k/k! = n!/(k!(n-k)!)$.

2.3 Probabilité conditionnelle et indépendance

Définition. *Soit* P *est une probabilité sur* $(\Omega, \mathcal{A})$ *et soit* $(E, F) \in \mathcal{A}^2$*. On appelle probabilité conditionnelle de* B *sachant* A*, la probabilité que* B *soit réalisé sachant que* A *est réalisé et* $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$ *si* $P(A) \neq 0$.

Remarque.

Calculer une probabilité sachant A revient à travailler avec une nouvelle probabilité sur l'ensemble fondamental A et la tribu qui lui est associée.

Définition. A *et* B*, événements de* $(\Omega, \mathcal{A})$ *sont indépendants pour* P *si* $P(B \mid A) = P(B)$.

Conséquence.

A et B, événements de $(\Omega, \mathcal{A})$ sont indépendants pour P si et seulement si $P(A \cap B) = P(A)P(B)$.

Remarque.

Ne pas confondre indépendance et incompatibilité. Plus généralement,

Définition.

- *Soit $(A_i)_{i \in I}$ une famille d'événements de $(\Omega, \mathcal{A})$. On dit que la famille $(A_i)_{i \in I}$ est une famille d'événements indépendants si et seulement si pour tout $k \in \mathbb{N}$, pour tout $j_1, \cdots, j_k \in I^k$ distincts, $P\Big(A_{j_1} \cap \cdots \cap A_{j_k}\Big) = P(A_{j_1}) \times \cdots \times P(A_{j_k})$.*
- *Si $(\Omega, \mathcal{A}_i)_{i \in I}$ est une famille de tribus sur Ω, on dit que ces tribus sont indépendantes si pour tout $A_i \in \mathcal{A}_i$ la famille $(A_i)_{i \in I}$ est indépendante.*

Définition. *Soit Ω un ensemble. On dit que $(E_i)_{i \in J}$ famille de sous-ensembles de Ω forme une partition de Ω si:*

- *Les E_i sont incompatibles deux à deux soit $E_i \cap E_j = \emptyset$ pour $i \neq j$.*
- *L'ensemble des E_i couvre Ω soit $\displaystyle\bigcup_{i \in J} E_i = \Omega$.*

Proposition. Formule des probabilités totales *Soit P est une probabilité sur $(\Omega, \mathcal{A})$ et soit $(E_i)_{i \in J}$ une partition de Ω alors:*

$$\text{pour } A \in \mathcal{A}, \quad P(A) = \sum_{i \in J} P(A \cap E_i).$$

Proposition. Formule de Bayes *Soit P est une probabilité sur $(\Omega, \mathcal{A})$ et soit $(E_i)_{i \in J}$ une partition de Ω. Soit A un événement de $\mathcal{A}$). On suppose que l'on connaît $P(E_i)$ et $P(A \mid E_i)$ pour tout $i \in J$. Alors, pour $k \in J$:*

$$P(E_k \mid A) = \frac{P(A \mid E_k)P(E_k)}{\sum_{i \in J} P(A \mid E_i)P(E_i)}.$$

3 Variables aléatoires

3.1 Définitions et propriétés générales

Définition. *Soit P une probabilité sur $(\Omega, \mathcal{A})$. On appelle variable aléatoire X dans $I \subset \mathbb{R}$, une application de $\Omega \to I$ telle que pour tout $x \in I$, l'événement $\{X \leq x\} = \{\omega \in \Omega \text{ tel que } X(\omega) \leq x\}$ soit un événement de $\mathcal{A}$. Deux cas particuliers importants sont à distinguer:*

- *Si $I = \{x_i\}_{i \in J}$ avec $J \subset \mathbb{N}$ (par exemple $I = \{P, F\}$, $I = \mathbb{Z}$,...), X est appelée variable aléatoire discrète.*
- *Si I est un intervalle de $\mathbb{R}$ (par exemple $I = [0, 1]$, $I = \mathbb{R}^+$,...), X est appelée variable aléatoire réelle.*

Définition. *Soit P une probabilité sur $(\Omega, \mathcal{A})$ et X une variable aléatoire de $(\Omega, \mathcal{A}, P)$ à valeurs dans I. On appelle fonction de répartition de X la fonction $F_X : \mathbb{R} \to [0, 1]$ telle que $F_X(x) = P(X \leq x)$, pour $x \in \mathbb{R}$.*

Propriété.

- *F_x est une fonction croissante sur $\mathbb{R}$.*
- *$\displaystyle\lim_{x \to -\infty} F_X(x) = 0$ et $\displaystyle\lim_{x \to +\infty} F_X(x) = 1$.*

- $P(a < X \leq b) = F_X(b) - F_X(a)$, *pour tout* $-\infty \leq a < b \leq +\infty$.

Définition. *Soit P une probabilité sur $(\Omega, \mathcal{A})$ et X une variable aléatoire de $(\Omega, \mathcal{A}, P)$ à valeurs dans I.*

- *Si X est une variable aléatoire discrète ($I = \{x_i\}_{i \in J}$), on appelle loi de probabilité de X l'application $P_X : \{x_i\}_{i\in J} \to [0,1]$ telle que $P(X = x_i) = P_X(x_i)$.*

- *Si X est une variable aléatoire réelle (I intervalle de $\mathbb{R}$), on appelle densité de probabilité de X* **si elle existe** *l'application $f_X : I \to [0, +\infty[$ telle que $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)dt$. X est alors appelée variable aléatoire absolument continue.*

Propriété. *Si X est une variable aléatoire discrète à valeurs dans $I = \{x_i\}_{i\in J}$ alors $\displaystyle\sum_{i\in J} P(X = x_i) = 1$.*

Propriété. *Si X est une variable aléatoire absolument continue de densité de probabilité f_X alors:*

- *f_X vérifie $\displaystyle\int_{-\infty}^{+\infty} f_X(t)dt = 1$.*

- *$F'_X(x) = f_X(x)$ pour presque tout $x \in \mathbb{R}$.*

- *$P(a < X \leq b) = \displaystyle\int_a^b f_X(t)dt$, pour tout $-\infty \leq a < b \leq +\infty$.*

- *$P(X = x) = 0$ pour tout $x \in \mathbb{R}$.*

3.2 Moments d'une variable aléatoire

Définition. *Soit P une probabilité sur $(\Omega, \mathcal{A})$, X une variable aléatoire de $(\Omega, \mathcal{A}, P)$ à valeurs dans I et h une fonction de I dans $\mathbb{R}$.*

- *Si X est une variable aléatoire discrète ($I = \{x_i\}_{i\in J}$), l'espérance de X est*

$$BBeX = \sum_{i\in J} x_i P(X = x_i) = \sum_{i \in J} x_i P_X(x_i) \quad \textbf{si cette somme existe}.$$

Plus généralement, $\mathbb{E}h(X) = \displaystyle\sum_{i\in J} h(x_i)P(X = x_i)$ **si cette somme existe**.

- *Si X est une variable aléatoire aboslument continue à valeurs dans I et de densité f_X, l'espérance de X est*

$$\mathbb{E}X = \int_I x f_X(x)dx = \int_{-\infty}^{+\infty} x f_X(x)dx \quad \textbf{si cette intégrale existe}.$$

Plus généralement, $\mathbb{E}h(X) = \displaystyle\int_I h(x)f_X(x)dx$ **si cette intégrale existe**.

Définition. *Soit X une variable aléatoire de $(\Omega, \mathcal{A}, P)$.*

- *La variance de X,* **si elle existe**, *est $var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}((X - \mathbb{E}X)^2) \geq 0$.*

- *L'écart-type de X,* **si la variance existe**, *est $\sigma(X) = \sqrt{var(X)}$.*

Propriété. *Soit X une variable aléatoire (dont l'espérance et la variance existent), et a, b deux réels.*

- *$\mathbb{E}b = b$ et $\mathbb{E}(aX + b) = a\mathbb{E}X + b$.*

- *$var(b) = 0$ et $var(aX + b) = a^2 var(X)$.*

3.3 Lois à connaître

Définition. *Les lois à connaître sont:*

- *Lois discrètes: loi uniforme, loi de Bernoulli, loi binomiale, loi de Poisson.*

- *Lois continues: loi uniforme, loi exponentielle, loi gamma, loi normale.*

3.4 Loi d'une variable aléatoire fonction d'une autre variable aléatoire

Utilisation de la fonction de répartition...

4 Théorèmes et limite et statistique inférentielle

4.1 Variables aléatoires indépendantes

Définition. *Soit $(X_i)_{i\in I}$ une famille de variables aléatoires. On dit que ces variables sont indépendantes si la famille de tribus engendrées par les $X_i^{-1}(B)$ où $B \in \mathcal{B}(\mathbb{R})$ sont indépendantes.*

Propriété. *Soit $(X_i)_{i\in I}$ une famille de variables aléatoires. Ces variables sont indépendantes si et seulement si pour tout $k \in \mathbb{N}^*$, pour tout $i_1 < \ldots < i_k \in I^k$ et tout $(B_1, \ldots, B_k) \in \mathcal{B}(\mathbb{R})^k$,*

$$P\big(X_{i_1} \in B_1 \cap \cdots \cap X_{i_k} \in B_k\big) = P\big(X_{i_1} \in B_1\big) \times \cdots \times P\big(X_{i_k} \in B_k\big).$$

Propriété. *Si $X = (X_1, \cdots, X_d)$ est un vecteur aléatoire absolument continu de densité de probabilité f_X. Alors la famille $(X_1, \cdots, X_d)$ est une famille de variables aléatoires indépendantes si et seulement s'il existe d fonctions positives f_{X_i} telles que $f_X = \prod_{i=1}^d f_{X_i}$.*

Propriété. *Si $(X_1, \cdots, X_d)$ est une famille de variables aléatoires indépendantes alors $cov(X_i, X_j) = 0$ si $i \neq j$. La réciproque est fausse sauf dans le cas où $(X_1, \cdots, X_d)$ est un vecteur gaussien.*

4.2 Convergence d'une suite variables aléatoires

Propriété (Inégalité de Markov)**.** *Soit X une variable aléatoire positive dont l'espérance existe. Alors pour tout $\varepsilon > 0$,*

$$P(X > \varepsilon) \leq \frac{\mathbb{E}X}{\varepsilon}.$$

Voici une conséquence directe de cette inégalité:

Propriété (Inégalité de Bienaymé-Tchebitchev)**.** *Soit X une variable aléatoire dont l'espérance m et la variance σ^2 existent. Alors pour tout $\varepsilon > 0$,*

$$P(|X - m| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

Définition. • *Convergence en probabilité.*

- *Convergence en loi.*

4.3 Théorèmes limite

Théorème (Loi Faible des Grands Nombres). *Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées par rapport à une loi dont l'espérance m existe. Alors :*

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n} \underset{n\to+\infty}{\overset{\mathcal{P}}{\longrightarrow}} m.$$

Théorème (Théorème de la limite central). *Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées par rapport à une loi dont l'espérance m et la variance σ^2 existent. Alors :*

$$\sqrt{n}\left(\frac{\overline{X}_n - m}{\sigma}\right) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(0,1).$$

5 Estimation et intervalles de confiance

Définition.

- *Estimateur d'un paramètre.*
- *Biais d'un estimateur.*
- *Risque quadratique d'un estimateur.*

Exemple.

Trois exemples sont à connaître:

1. Estimateur de l'espérance par la moyenne empirique.
2. Estimateur de la variance par la variance empirique.
3. Estimateurs du paramètre θ d'un n-échantillon $(X_1, \cdots, X_n)$ de loi uniforme sur $[0,\theta]$.

Définition.

- *Intervalle de confiance théorique de niveau α d'un paramètre.*
- *Réalisation d'un intervalle de confiance.*

Exemple.

Intervalle de confiance dans le cadre du théorème de la limite central.