Probabilités et Statistiques

IS101

ENSEIRB

2006-2007

Programme

1. Probabilités
   Généralités sur les probabilités

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités "continues"

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités "continues" – Loi des grands nombres, théorème "central limite"

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités "continues" – Loi des grands nombres, théorème "central limite" – Chaînes de Markov

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités "continues" – Loi des grands nombres, théorème "central limite" – Chaînes de Markov
2. Statistique

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités “continues” – Loi des grands nombres, théorème “central limite” – Chaînes de Markov
2. Statistique
   Statistique descriptive

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités "continues" – Loi des grands nombres, théorème "central limite" – Chaînes de Markov
2. Statistique
   Statistique descriptive – Estimation

Programme

1. Probabilités
   Généralités sur les probabilités – Probabilités discrètes – Probabilités "continues" – Loi des grands nombres, théorème "central limite" – Chaînes de Markov
2. Statistique
   Statistique descriptive – Estimation – Test d'hypothèses

Généralités

Théorie des probabilités : théorie mathématique prévue pour modéliser le hasard, l'incertain

- **Blaise Pascal** (1623-1662), jeux de hasard.

Généralités

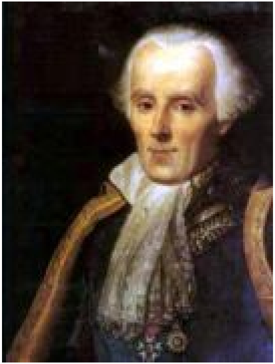
Théorie des probabilités : théorie mathématique prévue pour modéliser le hasard, l'incertain

- **Blaise Pascal** (1623-1662), jeux de hasard.
- **Jakob (Jacques) Bernoulli** (1654-1705)

Généralités

Théorie des probabilités : théorie mathématique prévue pour modéliser le hasard, l'incertain

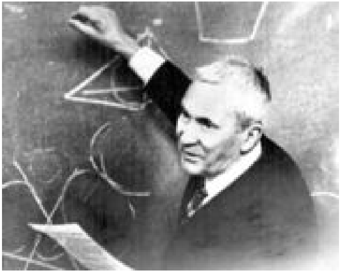
- **Blaise Pascal** (1623-1662), jeux de hasard.
- **Jakob (Jacques) Bernoulli** (1654-1705)
- **Pierre-Simon de Laplace** (1749-1827)



Généralités

Théorie des probabilités : théorie mathématique prévue pour modéliser le hasard, l'incertain

- **Blaise Pascal** (1623-1662), jeux de hasard.
- **Jakob (Jacques) Bernoulli** (1654-1705)
- **Pierre-Simon de Laplace** (1749-1827)
- **Andrei Nikolaievitch Kolmogorov** (Колмогоров) (1903-1987) – fondations mathématiques de la théorie



Espace de probabilités

On modélise une *expérience aléatoire* par un *espace de probabilités* $(\Omega, \mathcal{F}, \mathbb{P})$

- Ω : ensemble de *tous les résultats possibles* de l'expérience
- $\mathcal{F}$: ensemble de *parties de* Ω (souvent : ensemble de *toutes* les parties), les *événements* que l'on va considérer (et distinguer)
- $\mathbb{P} : \mathcal{F} \to [0, 1]$, fonction (ou *mesure*) de probabilités : attribue à chaque événement, une probabilité (un réel entre 0 (quasiment impossible) et 1 (quasiment certain))

Exemple

Lancer d'un dé classique (6 faces)

- $\Omega = \{1, 2, 3, 4, 5, 6\}$

Exemple

Lancer d'un dé classique (6 faces)

- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$

Exemple

Lancer d'un dé classique (6 faces)

- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$ ($2^6 = 64$ événements, y compris $\emptyset$ et Ω)

Exemple

Lancer d'un dé classique (6 faces)

- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$ ($2^6 = 64$ événements, y compris $\emptyset$ et Ω)
- pour tout $A \subset \Omega$, $\mathbb{P}(A) = \#A/6$

Exemple

Lancer d'un dé classique (6 faces)

- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\mathcal{F} = \mathcal{P}(\Omega)$ ($2^6 = 64$ événements, y compris $\emptyset$ et Ω)
- pour tout $A \subset \Omega$, $\mathbb{P}(A) = \#A/6$
- ainsi, $\mathbb{P}(\text{résultat pair}) = \mathbb{P}(\{2, 4, 6\}) = 1/2$

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou σ-*algèbre* ; en Anglais : σ-*algebra*), c'est-à-dire vérifier :

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou σ-*algèbre*; en Anglais : σ-*algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou σ-*algèbre* ; en Anglais : σ-*algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$
  - (stabilité par complémentaire) si $A \in \mathcal{F}$, alors $\overline{A} = \Omega - A \in \mathcal{F}$

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou σ-*algèbre*; en Anglais : σ-*algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$
  - (stabilité par complémentaire) si $A \in \mathcal{F}$, alors $\overline{A} = \Omega - A \in \mathcal{F}$
  - (stabilité par union finie) si $A, B \in \mathcal{F}$, alors $A \cup B \in \mathcal{F}$

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou σ-*algèbre*; en Anglais : σ-*algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$
  - (stabilité par complémentaire) si $A \in \mathcal{F}$, alors $\overline{A} = \Omega - A \in \mathcal{F}$
  - (stabilité par union finie) si $A, B \in \mathcal{F}$, alors $A \cup B \in \mathcal{F}$
  - (stabilité par union dénombrable) si $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, alors $\cup_n A_n \in \mathcal{F}$

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou σ*-algèbre*; en Anglais : σ*-algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$
  - (stabilité par complémentaire) si $A \in \mathcal{F}$, alors $\overline{A} = \Omega - A \in \mathcal{F}$
  - (stabilité par union finie) si $A, B \in \mathcal{F}$, alors $A \cup B \in \mathcal{F}$
  - (stabilité par union dénombrable) si $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, alors $\cup_n A_n \in \mathcal{F}$
- Notons que cela implique $\emptyset \in \mathcal{F}$, ainsi que la stabilité par intersection finie et intersection dénombrable

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou *σ-algèbre*; en Anglais : *σ-algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$
  - (stabilité par complémentaire) si $A \in \mathcal{F}$, alors $\overline{A} = \Omega - A \in \mathcal{F}$
  - (stabilité par union finie) si $A, B \in \mathcal{F}$, alors $A \cup B \in \mathcal{F}$
  - (stabilité par union dénombrable) si $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, alors $\cup_n A_n \in \mathcal{F}$
- Notons que cela implique $\emptyset \in \mathcal{F}$, ainsi que la stabilité par intersection finie et intersection dénombrable
- Notons également que $\mathcal{F} = \mathcal{P}(\Omega)$ est toujours une tribu

Axiomes (de $\mathcal{F}$)

- $\mathcal{F}$ doit être une *tribu* (ou *σ-algèbre*; en Anglais : *σ-algebra*), c'est-à-dire vérifier :
  - $\Omega \in \mathcal{F}$
  - (stabilité par complémentaire) si $A \in \mathcal{F}$, alors $\overline{A} = \Omega - A \in \mathcal{F}$
  - (stabilité par union finie) si $A, B \in \mathcal{F}$, alors $A \cup B \in \mathcal{F}$
  - (stabilité par union dénombrable) si $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}$, alors $\cup_n A_n \in \mathcal{F}$
- Notons que cela implique $\emptyset \in \mathcal{F}$, ainsi que la stabilité par intersection finie et intersection dénombrable
- Notons également que $\mathcal{F} = \mathcal{P}(\Omega)$ est toujours une tribu
- Dans les cas où Ω est fini ou dénombrable, dans la pratique on prendra toujours $\mathcal{F} = \mathcal{P}(\Omega)$; dans les autres cas, on ne s'y arrêtera pas

Axiomes (de $\mathbb{P}$)

- $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$

Axiomes (de $\mathbb{P}$)

- $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
- (additivité) si A et B sont *disjoints*, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Axiomes (de $\mathbb{P}$)

- $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
- (additivité) si A et B sont *disjoints*, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
- (σ-additivité) si $(A_n)_{n \geq 0}$ sont *deux à deux disjoints*, alors

$$\mathbb{P}\left(\bigcup_{n \geq 0} A_n\right) = \sum_{n \geq 0} \mathbb{P}(A_n)$$

Axiomes (de $\mathbb{P}$)

- $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
- (additivité) si A et B sont *disjoints*, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
- (σ-additivité) si $(A_n)_{n \geq 0}$ sont *deux à deux disjoints*, alors
$$\mathbb{P}\left(\bigcup_{n \geq 0} A_n\right) = \sum_{n \geq 0} \mathbb{P}(A_n)$$
- **On en déduit :** pour tous événements A, B,
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Axiomes (de $\mathbb{P}$)

- $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
- (additivité) si A et B sont *disjoints*, alors $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$
- (σ-additivité) si $(A_n)_{n\geq 0}$ sont *deux à deux disjoints*, alors

$$\mathbb{P}\left(\bigcup_{n\geq 0} A_n\right) = \sum_{n\geq 0} \mathbb{P}(A_n)$$

- **On en déduit également :** si $(A_n)_{n\geq 0}$ sont des événements quelconques,

$$\mathbb{P}\left(\bigcup_{n\geq 0} A_n\right) \leq \sum_{n\geq 0} \mathbb{P}(A_n)$$

(la série de droite peut diverger)

Notations

Dans la pratique, on définit très souvent un *événement* par une *condition* (un prédicat), l'événement étant l'ensemble des éléments de Ω qui vérifient la condition

- "le résultat est pair" : $\{2, 4, 6\}$

Notations

Dans la pratique, on définit très souvent un *événement* par une *condition* (un prédicat), l'événement étant l'ensemble des éléments de Ω qui vérifient la condition

- "le résultat est pair" : $\{2, 4, 6\}$
- "le résultat est inférieur ou égal à 4" : $\{1, 2, 3, 4\}$

Notations

Dans la pratique, on définit très souvent un *événement* par une *condition* (un prédicat), l'événement étant l'ensemble des éléments de Ω qui vérifient la condition

- "le résultat est pair" : $\{2, 4, 6\}$
- "le résultat est inférieur ou égal à 4" : $\{1, 2, 3, 4\}$
- "ou" se traduit par la réunion, "et" par l'intersection, la négation par le complémentaire

Notations

Dans la pratique, on définit très souvent un *événement* par une *condition* (un prédicat), l'événement étant l'ensemble des éléments de Ω qui vérifient la condition

- "le résultat est pair" : $\{2, 4, 6\}$
- "le résultat est inférieur ou égal à 4" : $\{1, 2, 3, 4\}$
- "ou" se traduit par la réunion, "et" par l'intersection, la négation par le complémentaire
- (abus de notation) on note

$$\mathbb{P}(\text{le résultat est pair}) \text{ pour } \mathbb{P}(\{2, 4, 6\})$$

Indépendance (d'événements)

Deux événements A et B sont *indépendants* si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

- "résultat pair" et "résultat inférieur ou égal à 2" : indépendants

Indépendance (d'événements)

Deux événements A et B sont *indépendants* si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

- "résultat pair" et "résultat inférieur ou égal à 2" : indépendants
- "résultat pair" et "résultat premier" : non indépendants

Indépendance (d'événements)

Deux événements A et B sont *indépendants* si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

- "résultat pair" et "résultat inférieur ou égal à 2" : indépendants
- "résultat pair" et "résultat premier" : non indépendants
- "résultat multiple de 3" et "résultat premier" : indépendants

Probabilité conditionnelle

Si B est un événement de probabilité **strictement** positive, on définit, pour tout événement A, la **probabilité conditionnelle de** A **sachant** B :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Probabilité conditionnelle

Si B est un événement de probabilité **strictement** positive, on définit, pour tout événement A, la **probabilité conditionnelle de A sachant B** :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

La probabilité conditionnelle de A sachant B représente la "nouvelle" probabilité que A se produise, si l'on sait exactement que B se produit

Probabilité conditionnelle

Si B est un événement de probabilité **strictement** positive, on définit, pour tout événement A, la **probabilité conditionnelle de A sachant B** :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

La probabilité conditionnelle de A sachant B représente la “nouvelle” probabilité que A se produise, si l'on sait exactement que B se produit

Exercice : vérifier que l'application

$$\mathbb{P}_B : A \to \mathbb{P}(A|B)$$

satisfait bien les axiomes d'une mesure de probabilités sur $\mathcal{F}$

Propriétés des probabilités conditionnelles

- **Formule des probabilités totales :** si $(B_i)_{i\in I}$ forment une *partition* de Ω, alors

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(B_i)\mathbb{P}(A|B_i)$$

Propriétés des probabilités conditionnelles

- **Formule des probabilités totales :** si $(B_i)_{i\in I}$ forment une *partition* de Ω, alors

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(B_i)\mathbb{P}(A|B_i)$$

- **Conditionnements successifs :** si $A_{n+1} \subset A_n$, alors on a

$$\mathbb{P}(A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_2)\dots\mathbb{P}(A_n|A_{n-1})$$

Propriétés (suite) : formules de Bayes

- **Formule de Bayes** (simple)

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}$$

Propriétés (suite) : formules de Bayes

- **Formule de Bayes** (simple)

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}$$

- **Formule de Bayes** (composée) si $(A_i)_{i\in I}$ forment une partition de Ω,

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(A_j)\mathbb{P}(B|A_j)}{\sum_{i\in I}\mathbb{P}(A_i)\mathbb{P}(B|A_i)}$$

Propriétés (suite) : formules de Bayes

- **Formule de Bayes** (simple)

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}$$

- **Formule de Bayes** (composée) si $(A_i)_{i\in I}$ forment une partition de Ω,

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(A_j)\mathbb{P}(B|A_j)}{\sum_{i\in I}\mathbb{P}(A_i)\mathbb{P}(B|A_i)}$$

(les formules de Bayes servent à “inverser” les conditionnements : passer de “B sachant A” à “A sachant B”)

Variable aléatoire

On appelle *variable aléatoire* (v.a.), (quasiment) toute fonction définie sur un espace de probabilités et à valeurs réelles.

Variable aléatoire

On appelle *variable aléatoire* (v.a.), (quasiment) toute fonction définie sur un espace de probabilités et à valeurs réelles.
Une variable aléatoire permet de modéliser toute situation dans laquelle l'expérience aléatoire fournit une valeur numérique fonction de "ce qui se passe"

Variable aléatoire

On appelle *variable aléatoire* (v.a.), (quasiment) toute fonction définie sur un espace de probabilités et à valeurs réelles.

Une variable aléatoire permet de modéliser toute situation dans laquelle l'expérience aléatoire fournit une valeur numérique fonction de "ce qui se passe"

Confusion possible dans les termes : une "variable aléatoire" est plus une *fonction* qu'une *variable*, et elle n'est *aléatoire* que si l'on munit l'espace de probabilités d'une mesure de probabilités fixée

Variable aléatoire

On appelle *variable aléatoire* (v.a.), (quasiment) toute fonction définie sur un espace de probabilités et à valeurs réelles.
Une variable aléatoire permet de modéliser toute situation dans laquelle l'expérience aléatoire fournit une valeur numérique fonction de "ce qui se passe"
Confusion possible dans les termes : une "variable aléatoire" est plus une *fonction* qu'une *variable*, et elle n'est *aléatoire* que si l'on munit l'espace de probabilités d'une mesure de probabilités fixée
Notation : On utilise typiquement des lettres majuscules (X, N, $G\dots$) pour désigner les variables aléatoires.

Loi d'une variable aléatoire

La *loi* d'une variable aléatoire X est la *mesure de probabilités* $\mathbb{P}_X$ sur $(\mathbb{R}, \mathcal{B})$ définie par

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

Loi d'une variable aléatoire

La *loi* d'une variable aléatoire X est la *mesure de probabilités* $\mathbb{P}_X$ sur $(\mathbb{R}, \mathcal{B})$ définie par

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

(Qu'est-ce que la tribu $\mathcal{B}$? c'est la *tribu des boréliens* de $\mathbb{R}$, c'est-à-dire la **plus petite tribu** qui contienne tous les intervalles de $\mathbb{R}$

Loi d'une variable aléatoire

La *loi* d'une variable aléatoire X est la *mesure de probabilités* $\mathbb{P}_X$ sur $(\mathbb{R}, \mathcal{B})$ définie par

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

(Qu'est-ce que la tribu $\mathcal{B}$? c'est la *tribu des boréliens* de $\mathbb{R}$, c'est-à-dire la **plus petite tribu** qui contienne tous les intervalles de $\mathbb{R}$

Dans la pratique, pour définir ou décrire la loi d'une v.a., on peut se contenter de décrire $\mathbb{P}_X(A)$ lorsque A est un intervalle

Indépendance (de variables aléatoires)

Deux variables aléatoires X et Y, définies sur le **même** espace de probabilités Ω (i.e., calculées sur une "même" expérience, éventuellement composée de plusieurs étapes) sont **indépendantes** si, pour **tous** les intervalles I et J, on a

$$\mathbb{P}(X \in I \text{ et } Y \in J) = \mathbb{P}(X \in I)\mathbb{P}(Y \in J)$$

L'égalité s'étend alors au cas où I et J sont des boréliens quelconques.