Statistique et Informatique (LI323)

Cours 5

Nicolas Baskiotis
nicolas.baskiotis@lip6.fr

Université Pierre et Marie Curie (UPMC)
Laboratoire d'Informatique de Paris 6 (LIP6)

Plan

- Mesures de probabilités sur $\mathbb{R}$ et $\mathbb{R}^n$,
- Variables aléatoires à valeurs continues,
- Fonctions de répartition et de densité,
- Approximation d'intégrales avec les méthodes de Monte-Carlo.

Mesures de probabilités sur $\mathbb{R}$

Motivation

Modéliser des résultats d'expériences aléatoires pouvant être des réels quelconques.
Par exemple : mesures de temps ou de distances.

Exemple : Probabilité continue uniforme sur $[0, 1]$

On souhaite créer une mesure telle que chaque valeur est équiprobable.
Plus exactement : si $0 \leq a < b \leq 1, P([a, b]) = b - a$.

Plus généralement : on souhaite associer une mesure de probabilité aux *intervalles* de $\mathbb{R}$.

Quelques détails techniques...

Rappel : mesure de probabilité

- Une mesure de probabilité est une fonction qui associe un *ensemble* à une valeur entre 0 et 1,
- la probabilité de l'univers Ω est de 1,
- la probabilité d'une union (dénombrable) d'ensembles disjoints est la somme des probabilités.

Fait marquant 1

Selon la théorie usuelle (théorie des ensembles + axiome du choix) :
Il *n'existe pas* de fonction P telle que:

1. P est une mesure de probabilité,
2. P peut être calculée pour n'importe quel sous-ensemble de $[0, 1]$,
3. pour tout intervalle $[a, b] \subset [0, 1]$: $P([a, b]) = b - a$.

Mesures de probabilités : définition générale (1)

Tribu

Une tribu $\mathcal{T}$ sur un ensemble Ω contient des sous-ensembles de Ω et vérifie:

1. $\Omega \in \mathcal{T}$,
2. si $E \in \mathcal{T}$, alors $\Omega \backslash E \in \mathcal{T}$,
3. si $(E_i)_{i \geq 1}$ est une suite d'ensembles appartenant à $\mathcal{T}$, alors $\bigcup_{i \geq 1} E_i \in \mathcal{T}$.

Note : l'ensemble des parties de Ω est une tribu.

Tribu de Borel

- La tribu de Borel sur $\mathbb{R}$, notée $\mathcal{B}$, est la plus petite tribu contenant tous les intervalles de $\mathbb{R}$,
- la tribu de Borel sur $\mathbb{R}^n$, notée $\mathcal{B}_n$ est la plus petite tribu contenant les produits cartésiens de n ensembles de $\mathcal{B}$:

$$\{I_1 \times I_2 \times \ldots \times I_n \,|\forall k, I_k \in \mathcal{B}\} \subset \mathcal{B}_n.$$

Les tribus de Borel contiennent les ensembles intéressants du point de vue des probabilités.

Mesures de probabilités : définition générale (2)

Espace mesurable

Un espace mesurable est un couple $(\Omega, \mathcal{T})$, où Ω est un ensemble et $\mathcal{T}$ est une tribu sur Ω. (Note : $(\mathbb{R}, \mathcal{B})$ *est donc un espace mesurable.*)

Mesure de probabilité

Soit $(\Omega, \mathcal{T})$ un espace mesurable.
Une mesure sur $(\Omega, \mathcal{T})$ est une fonction $P : \mathcal{T} \to [0, +\infty]$ telle que :

1. $P(\emptyset) = 0$,
2. $\forall E \in \mathcal{T}, P(E) \geq 0$,
3. si $(E_i)_{i \geq 1}$ sont deux à deux disjoints, et $\forall i, E_i \in \mathcal{T}$, alors

$$P\Big(\bigcup_{i \geq 1} E_i\Big) = \sum_{i \geq 1} P(E_i).$$

4. Si de plus $P(\Omega) = 1$, alors P est une *mesure de probabilité*.

Cette définition généralise la définition du cours 1 en prenant $\mathcal{P}(\Omega)$ *comme tribu sur* Ω *lorsque* Ω *est discret.*

Mesure de Borel

Fait marquant 2

Il existe une mesure λ sur $(\mathbb{R}, \mathcal{B})$, appelée mesure de Borel, telle que:

pour tout intervalle $[a,b]$ de $\mathbb{R}$ $(a \leq b)$, $\lambda([a,b]) = b - a$.

La mesure de probabilité uniforme sur un intervalle $[A,B], A < B$ est alors :

$$\forall I \in \mathcal{B}, P(I) = \frac{\lambda(I \cap [A,B])}{B - A}$$

Mesure de Borel sur $\mathbb{R}^n$

La mesure de Borel sur $(\mathbb{R}^n, \mathcal{B}_n)$, notée λ_n est définie par :

$$\forall I_1 \in \mathcal{B}, ..., I_n \in \mathcal{B}, \lambda_n\left(I_1 \times ... \times I_n\right) = \prod_{k=1}^{n} \lambda(I_k)$$

Par exemple : $\lambda_2([0,1/2] \times [0,1/2]) = 1/4$.
λ_2 correspond à l'aire d'une figure dans le plan, λ_3 au volume d'un objet dans l'espace à 3 dimensions.

Densité de probabilité

Définition

Une mesure de probabilité P sur $(\mathbb{R}^n, \mathcal{B}_n)$ admet une *fonction de densité* $p : \mathbb{R}^n \rightarrow \mathbb{R}$ si :
pour tout $I \subset \mathcal{B}_n, P(I) = \int_{x \in I} p(x) d\lambda_n(x)$

Exemples

- Si P est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B})$ qui admet comme fonction de densité p, alors pour tout $a < b$, on a :

$$P(]a, b]) = \int_a^b p(x)dx \quad \text{(avec les notations usuelles de l'intégrale sur } \mathbb{R}\text{).}$$

- La loi uniforme sur $[a, b]$ admet comme fonction de densité :

$$p(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon} \end{cases},$$

- une v.a. réelle à valeurs dans un ensemble discret *n'a pas* de fonction de densité.

Variables aléatoires à valeurs réelles

Variable aléatoire réelle

Soit P est une mesure de probabilité sur $(\Omega, \mathcal{T})$.
Une variable aléatoire réelle est une fonction $X : \Omega \mapsto \mathbb{R}$ telle que :

$$\forall I \in \mathcal{B}, X^{-1}(I) \in \mathcal{T}$$

X *induit une mesure de probabilité* P_X *sur* $(\mathbb{R}, \mathcal{B})$ *par :*

$$P_X(I) = P(X \in I) = P\big(X^{-1}(I)\big).$$

Note : cette définition généralise notre définition de variable aléatoires à valeurs réelles sur des ensembles discrets.

Fonctions de répartition et de densité d'une v.a.r.

Fonction de répartition

Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{T}, P)$.
La fonction de répartition de X, notée F_X, est définie par :

$$F_X : \left(\begin{array}{ccc} \mathbb{R} & \rightarrow & [0,1] \\ t & \mapsto & P(X \leq t) \end{array} \right)$$

On a alors $P(a < X \leq b) = F_X(b) - F_X(a)$.

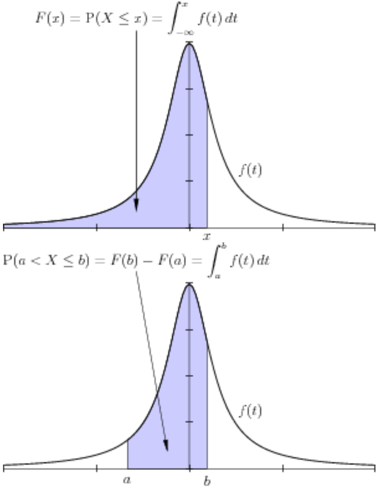
Fonction de densité

Une v.a. réelle X admet une *fonction de densité* p_X, si, pour tout $x \in \mathbb{R}$:

$$F_X(x) = \int_{-\infty}^{x} p_X(u)du.$$

- On a alors, pour $a < b$, $P(a < X \leq b) = \int_a^b p_X(u)du$.
- Si X est une $v.a.r.$ telle que sa fonction de répartition est dérivable, alors $p_X = F'_X$ est une densité de X.

Fonctions de répartition et de densité d'une v.a.r.



Espérance et variance d'une v.a.r. continue

Définition et propriétés

- Soit X une v.a.r. sur espace probabilisé $(\mathbb{R}^n, \mathcal{B}_n, P)$ où P a une fonction de densité p. L'espérance de X est alors définie par :

$$\mathbb{E}(X) = \int_{\omega \in \Omega} X(\omega) p(\omega) d\lambda_n(\omega) \,.$$

- Soit X, une v.a.r. de densité p_X. L'espérance de X est définie par :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} u p_X(u) du.$$

- La variance de X est définie par : $V(X) = \mathbb{E}\big((X - \mathbb{E}(X))^2\big)$.
- Soit $f : \mathbb{R} \mapsto \mathbb{R}$. Alors : $\mathbb{E}\big(f(X)\big) = \int_{-\infty}^{+\infty} f(u) p_X(u) du$

Les résultats montrés pour les v.a.r. à valeurs discrètes restent vraies :

- Les propriétés de l'espérance et de la variance,
- les inégalités de Markov et Tchebychev, et la loi des grands nombres.

Application : Méthodes de Monte-Carlo

Types d'algorithmes probabilistes

- Algorithmes Las Vegas : renvoient toujours la réponse correcte, mais dans un temps variable,
- algorithmes de Monte-Carlo : les ressources sont limitées a priori, mais le résultat peut ne pas être exact.

Exemple : approximation de nombres réels

Principe de la méthode (exemples : approximation de $\ln 2$ et de π) :

1. écrire le nombre comme l'espérance d'une fonction d'une v.a.r. à densité,

$$\ln 2 = \int_0^1 \frac{1}{1+u} du, \frac{\pi}{4} = \int_{u=0}^1 \int_{v=0}^1 I_{\{u^2+v^2 \leq 1\}} dudv$$

2. générer n valeurs selon la densité, calculer la fonction pour chaque valeur échantillonnée,
3. loi des grands nombres : la moyenne des n valeurs calculées tend vers le nombre qu'on souhaite calculer.