Statistiques et Probabilités

UIT II Informatique - S3 -M3201

Maha Elbayad	maha.elbayad@inria.fr
Franck Corset	franck.corset@univ-grenoble-alpes.fr
Cours disponible sur	**https://fcorset.github.io/cours/cours.html**

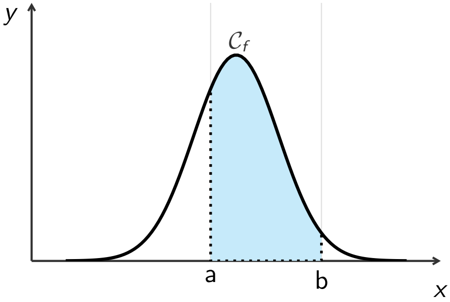
09 Sept. 2019

L'intégrale: interprétation graphique

L'intégrale $\int_a^b f(x)dx$ est l'aire sous la courbe $\mathcal{C}_f$ d'equation $y = f(x)$ dans le domaine $[a, b]$.

Interpretation valable uniquement pour $f : \mathbb{R} \to \mathbb{R}^+$ (positive)

(Le cas dans le cadre de ce cours).



Définition: dérivée

Soit f une fonction définie sur un intervalle I et $a \in I$. Si la limite $\lim_{h \to 0} \frac{f(a+h)-f(a)}{h}$ existe, on la note $f'(a)$ et on l'appelle nombre dérivé de la fonction f en a. On dit que **f est dérivable en a**.
Si f est dérivable sur I on appelle f' la fonction d'erivée de f.

Définition: primitive

Une primitive d'une fonction f est une fonction F dont f est la dérivee: $F' = f$.

Définition: intégrale

L'intégrale sur $[a, b]$ de f est égale à $F(b) - F(a)$ que l'on note $[F(x)]_a^b$, où F est la primitive de f.

$$\int_a^b f(x)dx = F(b) - F(a) = [F(x)]_a^b$$

Exemple: Pour $f : x \mapsto x$ de primitive $F(x) = \frac{x^2}{2}$

$$\int_0^1 xdx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}$$

Définition: Valeur moyenne

La valeur moyenne d'une fonction continue f sur $[a, b]$ est définie comme:

$$\mu = \frac{1}{b-a} \int_a^b f(x)dx$$

Quelques propriétés des inégrales

$$f \text{ positive sur } [a,b] \implies \int_a^b f(x)dx \geq 0 \qquad \text{(Positivité)}$$

$$\int_a^a f(x)dx = 0$$

$$\int_a^b f(x)dx = -\int_b^a f(x)dx$$

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \qquad \text{(Relation de Chasles)}$$

$$\int_a^b (f(x) + \lambda g(x))\, dx = \int_a^b f(x)dx + \lambda \int_a^b g(x)dx \qquad \text{(Linéarité)}$$

$$\forall x \in [a,b]\, f(x) \leq g(x) \implies \int_a^b f(x)dx \leq \int_a^b g(x)dx \qquad \text{(Ordre)}$$

Intǵrer une fonction

- Des primitives usuelles.

f	f'	F
$x \mapsto \lambda$	$x \mapsto 0$	$x \mapsto \lambda x + C$
$(\forall a \in \mathbb{R}\backslash\{-1\})\ \ x \mapsto x^a$	$x \mapsto a x^{a-1}$	$x \mapsto \frac{x^{a+1}}{a+1} + C$
$x \mapsto e^x$	$x \mapsto e^x$	$x \mapsto e^x + C$
$x \mapsto \frac{1}{x}$	$x \mapsto -\frac{1}{x^2}$	$x \mapsto \ln(x) + C$
$x \mapsto \cos(x)$	$x \mapsto -\sin(x)$	$x \mapsto \sin(x) + C$
$x \mapsto \sin(x)$	$x \mapsto \cos(x)$	$x \mapsto -\cos(x) + C$

Rappels

Primitives et intégrales

Intgrer une fonction

- Intégration par partie.

  Rappelons la dérivée du produit $(uv)' = u'v + uv'$, on a:

$$\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx$$

Exemple: Calculer l'integrale de $f : x \mapsto xe^x$ sur $[a, b]$.

Intgrer une fonction

- Intégration par changement de variable.

  Pour ϕ une fonction dérivable:

$$\int_a^b f(\phi(x))\,\phi'(x)dx = \int_{\phi(a)}^{\phi(b)} f(y)dy$$

Exemple: Calculer l'integrale de $f : x \mapsto x\sqrt{1+x^2}$ sur $[a, b]$.

Les probabilités traitent de modèles mathématiques et de leurs propriétés (**théorique**) alors que la statistique utilise les probabilités pour les appliquer à des données réelles (**empirique**).

Définition: Une variable aléatoire

On appelle variable aléatoire notée X tout nombre réel qui dépend du résultat d'une expérience aléatoire. Proprement dit, X est une fonction $X : \Omega \mapsto \mathbb{R}^+$ où Ω est **l'espace des observables**.

Définition: support d'une variable al'eatoire

Le support d'une variable aléatoire X est l'ensemble $X(\Omega)$ des valeurs atteintes par X.

- X est **discrète** si $X(\Omega)$ est dénombrable e.g. $X(\Omega) \subset \mathbb{N}$.
- X est **continue** si $X(\Omega)$ est indénombrable e.g. $X(\Omega) = [0, 1]$.

Exemples:

- Deux tirages à pile (0) ou face (1) : $\Omega = \{11, 10, 01, 00\}$.
  Et X la somme des deux valeurs $X(\Omega) = \{0, 1, 2\}$ (discrète).
- La taille d'une personne en cm: $\Omega = \mathbb{R}^+$.
  Et X la taille elle-même $X(\Omega) = \mathbb{R}^+$ (continue).

Définition: Masse de probabilité

On appelle fonction de masse de probabilités, la fonction suivante :

$$\forall k \in X(\Omega),\ P(X = k) = p_k$$

Où p_k est la probabilité que $X = k$ se réalise. On a les propriétés suivantes:

$$\forall k \in X(\Omega),\ p_k \geq 0 \qquad \text{et} \quad \sum_{k \in X(\Omega)} p_k = 1$$

Exemple (1) Un tiarage à pile (0) ou face (1) biaisée
X prend la valeur 1 avec $p_1 = p \in [0, 1]$ et la valeur 0 avec $p_0 = 1 - p$
On appelle cette loi la loi de Bernoulli de paramétre p que l'on note $\mathcal{B}(p)$.

Exemple (2) Loi binomiale:
On lance n tirages indépendants à pile (0) ou face (1) avec probabilités $(1-p, p)$ et on évalue X la somme des n tirages.
On dit que X suit une loi binomiale de paramètre $n \in \mathbb{N}^*$ et $p \in [0,1]$, notée $\mathcal{B}(n,p)$.
$X(\Omega) = \{0, 1, \ldots, n\}$, et les masses de probabilité $p_k = P(X = k)$, sont :

$$\forall k \in \{0 \ldots, n\},\ p_k = P(x = k) = C_n^k p^k (1-p)^{n-k}$$

Où les coefficients du binôme (cf. triangle de Pascal) sont définis par :

$$C_n^k = \frac{n!}{k!(n-k)!} \qquad \text{(combinaison de } k \text{ parmi } n\text{)}$$

avec $n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$ et par convention $0! = 1$.

Exemple (2) Loi binomiale:
La somme des probabilités est égale à 1:

$$\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = [p + (1-p)]^n = 1^n = 1$$

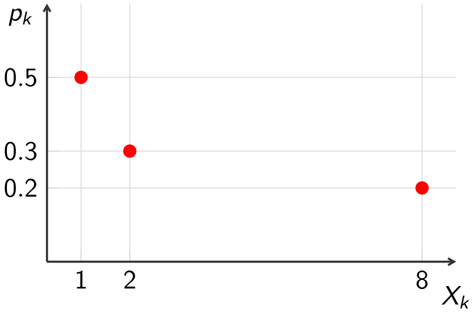
Résultat de la formule du binôme de Newton: $(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}$

Définition: Fonction de répartition

On appelle **fonction de répartition** de X (f.d.r en français, c.d.f. en anglais pour cumulative density function) la fonction suivante F_X :

$$\forall x \in \mathbb{R},\ F_X(x) = P(X \leq x) = \sum_{k:X_k \leq x} p_k$$

Exemple:



$$F_X(0.5) = ?0$$
$$F_X(1) = ?0.5$$
$$F_X(1.5) = ?0.5$$
$$F_X(2.1) = ?0.8$$
$$F_X(8) = ?1$$

Définition: Espérance

Soit X une variable aléatoire discrète de fonction de masse de probabilité $\forall k \in X(\Omega)$, $f(k) = p_k$. On appelle espérance de X, que l'on note $\mathbb{E}[X]$, la quantité suivante :

$$\mathbb{E}[X] = \sum_{k \in X(\Omega)} k p_k$$

Dans cette définition, on suppose que cette espérance est finie.

Exemple (1) Loi de Bernoulli

On suppose que X suit une loi de Bernoulli de paramètre p $B(p)$.
On rappelle que $X(\Omega) = \{0, 1\}$ et $P(X = 1) = p_1 = p$ et $P(X = 0) = p_0 = 1 - p$.
L'espérance de X est :

$$\mathbb{E}[X] = \sum_{k \in X(\Omega)} k p_k = \sum_{k=0}^{1} k p_k = 0 * (1 - p) + 1 * p = p$$

Exemple (1) Loi de Poisson

On suppose que X suit une loi de Poisson de paramètre $\mathcal{P}(\lambda)$, $\lambda > 0$. $X(\Omega) = \mathbb{N}$ et

$$\forall k \in \mathbb{N},\ P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

Interprétation: Si le nombre moyen d'occurrences d'événements dans un intervalle de temps fixé est λ, alors la probabilité qu'il existe exactement k occurrences ($k \in \mathbb{N}$) suit une lois de Poisson $\mathcal{P}(\lambda)$
La somme des masses:

$$\sum_{k=0}^{+\infty} p_k = \sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} = e^{-\lambda} * e^{\lambda} = 1 \qquad \text{(séries entières usuelles)}$$

Exemple (1) Loi de Poisson

L'espérance de X vaut :

$$\mathbb{E}[X] = \sum_{k=0}^{+\infty} kp_k = e^{-\lambda} \sum_{k=1}^{+\infty} k \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} * e^{\lambda} = \lambda$$

Définition: Variance

Soit X une variable aléatoire discrète de fonction de masse de probabilité $\forall k \in X(\Omega)$, $f(k) = p_k$. On appelle variance de X, que l'on note $\mathbb{V}[X]$, la quantité suivante :

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

où

$$\mathbb{E}[X^2] = \sum_{k \in X(\Omega)} k^2 p_k$$

Exemple (1) Loi de Bernoulli

En reprenant la loi de Bernoulli, on peut calculer le moment d'ordre 2 :

$$\mathbb{E}[X^2] = \sum_{k \in X(\Omega)} k^2 p_k = \sum_{k=0}^{1} k^2 p_k = 0^2 * (1-p) + 1^2 * p = p$$

La variance de X est donc :

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1-p)$$

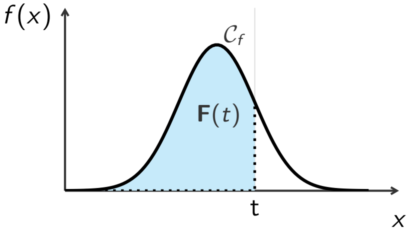
Définition: densité et fonction de répartition

La variable aléatoire X associée à une fonction f donnée et définie sur $\mathbb{R}$, continue par morceaux, représente le fait de tirer un nombre au hasard avec la probabilité suivante :

$$\mathrm{Proba}(X \leq t) = \int_{-\infty}^{t} f(x)dx = F(t)$$

où t est un réel fixé.
Cette écriture n'a de sens que si f est une fonction positive sur $\mathbb{R}$ et $\int_{-\infty}^{+\infty} f(x)dx = 1$.
f est appelée **densité** de X et F est sa **fonction de répartition**.

La valeur $F(t)$ peut être vue comme l'aire de la surface délimitée par la demi-droite $]-\infty, t]$, la droite $x = t$ et la courbe représentative de f.



Propriétés de la fonction de répartition

(1) F croissante. (2) $\forall t,\ F(t) \in [0, 1]$. (3) $\lim\limits_{t\to+\infty} F(t) = 1$. (4) $\lim\limits_{t\to-\infty} F(t) = 0$.

(5) Si F est dérivable alors $F'(x) = f(x)$.

Propriétés

- $\forall a \in \mathbb{R},\ P(X \leq a) = F(a)$.
- $\forall a \in \mathbb{R},\ P(X > a) = 1 - F(a)$.
- $\forall a, b \in \mathbb{R}$ tels que $a < b$, $P(X \in]a, b]) = F(b) - F(a)$.
- $P(X = x) = F(x) - F(x^{-})$. Si F est continue à gauche de x alors $P(X = x) = 0$

Définition: Espérance

L'espérance de X (appelée aussi moyenne de X) correspond à la valeur suivante :

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f(x)\, dx.$$

Définition: Variance - écart-type

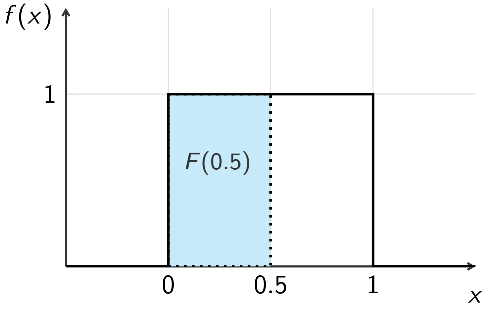
La variance de X est :

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \int_{-\infty}^{+\infty} x^2 f(x)\, dx - \mathbb{E}[X]^2.$$

L'écart-type de X notée σ_X est $\sigma_X = \sqrt{Var(X)}$.

Exemple (1) Loi uniforme sur $[0, 1]$
La fonction densité est :

$$f : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \begin{cases} 1 & \text{si } x \in [0,1] \\ 0 & \text{sinon} \end{cases} \end{array}$$



La fonction de répartition de cette loi est :

$$F : \begin{array}{ccc} \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } x \in [0,1] \\ 1 & \text{si } x > 1 \end{cases} \end{array}$$

$$\mathbb{E}[X] = \int_0^1 x dx = 1/2 \ , \ \mathbb{E}[X^2] = \int_0^1 x^2 dx = 1/3 \ , \ Var[X] = 1/12.$$

	V.a discrète	V.a continue
$X(\Omega)$:	dénombrable	indénombrable
Masse/Densité	$(p_k)_{k\in X(\Omega)}$	$f(x)$
	$p_k \geq 0,\ \sum_k p_k = 1$	$f(x) \geq 0,\ \int_{-\infty}^{\infty} f(x)dx = 1$
Fonction de répartition	$F_X(t) = \sum_{k:X_k\leq t} p_k$	$\int_{-\infty}^{t} f(x)dx$
Espérance	$\mathbb{E}[X] = \sum_k kp_k$	$\mathbb{E}[X] = \int xf(x)dx$
Moment d'ordre 2	$\mathbb{E}[X^2] = \sum_k k^2 p_k$	$\mathbb{E}[X^2] = \int x^2 f(x)dx$

Variance $\quad \mathrm{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$

Pour toute v.a X, continue ou discrète:

- $\forall a \in \mathbb{R},\ \mathbb{E}[a] = a$
- $\forall (a,b) \in \mathbb{R}^2,\ \mathbb{E}[aX + b] = a\,\mathbb{E}[X] + b$
- Pour toute v.a Y, $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- Pour toute fonction g,
  - si X est continue de densité f: $\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$
  - si X est discrète de masse $(p_k)_k$: $\mathbb{E}[g(X)] = \sum_k g(k)p_k$
- $\forall a \in \mathbb{R},\ \mathrm{Var}[a] = 0$
- $\forall a \in \mathbb{R},\ \mathrm{Var}[aX] = a^2\mathrm{Var}[X]$
- $\forall a \in \mathbb{R},\ \mathrm{Var}[-aX] = a^2\mathrm{Var}[X]$ (Une variance est toujours positive)
- $\forall (a,b) \in \mathbb{R}^2,\ \mathrm{Var}[aX + b] = a^2\mathrm{Var}[X]$