

Chapitre 1. Modèle statistique et échantillonnage

Claude Petit, université de Rennes - claud.petit@univ-rennes.fr

Sept. 2026

1. Modèle statistique
2. Quelques types de modèles statistiques

1. Modèle statistique

Définitions

- **Expérience aléatoire** $\iff$ espace probabilisé $\iff (\mathbb{X}, \mathfrak{F}, \mathbb{P})$.
 - $\mathbb{X}$ ensemble des issues possibles de l'expérience aléatoire.
 - $\mathfrak{F}$ tribu sur $\mathbb{X}$ (événements de l'expérience).
 - $\mathbb{P}$ mesure de probabilité sur $\mathfrak{F}$.
 - **Modèle statistique**: famille d'expériences aléatoires, $\hat{m} \mathbb{X}$ et $\hat{m} \mathfrak{F}$: $(\mathbb{X}, \mathfrak{F}, \mathcal{P})$.
 - Observation x : réalisation d'une variable aléatoire $X \sim \mathbb{P}$.
 - **Objectif de l'inférence statistique**: déterminer $\mathbb{P} \in \mathcal{P}$ sachant x .
- ⚠ Savoir distinguer $(\mathbb{X}, \mathfrak{F}, \mathbb{P})$ et $(\mathbb{X}, \mathfrak{F}, \mathcal{P})$!

Inférence, modélisation

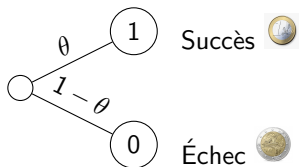
Données réelles, observations

Modèle aléatoire sous-jacent

Tirage

Exemple 1: épreuve de Bernoulli

- X v.a. de loi de **Bernoulli** $b(\theta)$, $\theta \in [0, 1]$ inconnu.
- $x \in \{0, 1\}$ résultat d'un lancer à pile ou face, x réalisation de X .
- Choix raisonnable pour $\mathcal{P}$: ensemble des lois de Bernoulli de paramètre θ , θ variant de 0 à 1.
- $\mathbb{X} = \{0, 1\} = \{\pi, \phi\} = \{\text{succès}, \text{échec}\}$.
- $\mathfrak{F}$ ensemble des parties de $\mathbb{X}$: $\mathfrak{F} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.
- $\mathcal{P} = \{\mathbb{P}_\theta; \theta \in [0, 1]\}$ avec $\mathbb{P}_\theta[X = 1] = \theta$ et $\mathbb{P}_\theta[X = 0] = 1 - \theta$.

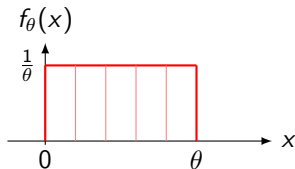


Exemple 2: loi uniforme

- X v.a. de **loi uniforme** sur $[0, \theta]$, θ inconnu.
- X résultat d'un tirage aléatoire équitable entre 0 et θ .
 - $\mathbb{X} = [0, \theta]$.
 - $\mathfrak{F}$ ensemble des boréliens de $\mathbb{X}$.
 - $\mathcal{P} = \{\mathbb{P}_\theta; \theta > 0\}$. $\mathbb{P}_\theta$ défini par sa densité:

$$f(x) = \frac{1}{\theta} \mathbb{1}_{[0, \theta]}(x) \quad (1)$$

$\mathbb{P}_\theta$ caractérisé par le paramètre réel θ . Toute valeur de θ définit une mesure.



Modèle d'échantillonnage

- Observons n réalisations $x_1, \dots, x_n$ d'une v.a. X , de manière indépendante.
- $X_1, \dots, X_n$ v.a.i.i.d. sous-jacentes à ces observations. X_i à valeurs dans $(\mathbb{X}, \mathfrak{F})$, $\mathbb{P}$ la loi commune des X_i . $(X_1, \dots, X_n)$ n -échantillon de X .

Definition

On appelle **modèle d'échantillonnage** associé au n -échantillon, le triplet

$$(\mathbb{X}^n, \mathfrak{F}^{\otimes n}, \{\mathbb{P}^{\otimes n}, \mathbb{P} \in \mathcal{P}\}) \quad (2)$$

où $\mathcal{P}$ est une famille de probabilité définie sur $(\mathbb{X}, \mathfrak{F})$.

Par $\perp$ et identique distribution des X_i , $\mathbb{P}^{\otimes n}$ loi du vecteur $(X_1, \dots, X_n)$:

$$\mathbb{P}^{\otimes n}(dx_1, \dots, dx_n) = \prod_{i=1}^n \mathbb{P}(dx_i) \quad (3)$$

Exemple 3: répétition de n épreuves de Bernoulli $\perp$

- Le modèle d'échantillonnage associé est $(\{0, 1\}^n, \mathfrak{F}^{\otimes n}, \{b(\theta)^{\otimes n}, \theta \in [0, 1]\})$, avec $\mathfrak{F} = \{\emptyset, \{0\}, \{1\}, \Omega\}$.

Ce modèle traduit simplement le fait que dans une **expérience répétée de façon indépendante**, la loi de probabilité du n -échantillon est la loi produit et l'espace sous-jacent est le produit cartésien de l'espace associé à l'un des éléments de l'échantillon.

- Supposons que $n = 10$ et mettons qu'une réalisation du n -échantillon soit $(0, 0, 1, 1, 1, 0, 1, 0, 0, 1)$. Nous souhaitons évaluer θ à partir de cet échantillon. **Comment peut-on procéder ?**

Exemple 3: répétition de n épreuves de Bernoulli $\perp$

- Le modèle d'échantillonnage associé est $(\{0, 1\}^n, \mathfrak{F}^{\otimes n}, \{b(\theta)^{\otimes n}, \theta \in [0, 1]\})$, avec $\mathfrak{F} = \{\emptyset, \{0\}, \{1\}, \Omega\}$.

Ce modèle traduit simplement le fait que dans une **expérience répétée de façon indépendante**, la loi de probabilité du n -échantillon est la loi produit et l'espace sous-jacent est le produit cartésien de l'espace associé à l'un des éléments de l'échantillon.

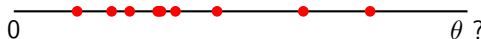
- Supposons que $n = 10$ et mettons qu'une réalisation du n -échantillon soit $(0, 0, 1, 1, 1, 0, 1, 0, 0, 1)$. Nous souhaitons évaluer θ à partir de cet échantillon. **Comment peut-on procéder ?**

$$\theta = \bar{x} = \sum x_i / 10 = 1/2.$$

répétition de n épreuves $\perp$ d'une loi uniforme sur $[0, \theta]$

$(\mathbb{X}^n, \mathfrak{F}^{\otimes n}, \{\mathbb{P}^{\otimes n}, \mathbb{P} \in \mathcal{P}\})$ avec

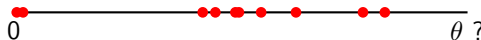
- $\mathbb{X} = [0, \theta]$.
- $\mathfrak{F} = \mathcal{B}_{[0, \theta]}$.
- $\mathbb{P} = \mathbb{P}_\theta$ loi uniforme sur $[0, \theta]$.
- Supposons que $n = 10$ et que nous disposions uniquement de l'échantillon $x = (0, 1; 0, 6; 0, 2; 1, 2; 3; 0, 5; 1, 5; 2; 1; 1, 9)$.
- Comment estimer θ ?



répétition de n épreuves $\perp$ d'une loi uniforme sur $[0, \theta]$

$(\mathbb{X}^n, \mathfrak{F}^{\otimes n}, \{\mathbb{P}^{\otimes n}, \mathbb{P} \in \mathcal{P}\})$ avec

- $\mathbb{X} = [0, \theta]$.
- $\mathfrak{F} = \mathcal{B}_{[0, \theta]}$.
- $\mathbb{P} = \mathbb{P}_\theta$ loi uniforme sur $[0, \theta]$.
- Supposons que $n = 10$ et que nous disposions uniquement de l'échantillon $x = (0, 1; 0, 6; 0, 2; 1, 2; \textcolor{red}{3}; 0, 5; 1, 5; 2; 1; 1, 9)$.
- **Comment estimer θ ?**



$$\theta = \max_i x_i = \textcolor{red}{3}.$$

Definition

Une **statistique** S sur un modèle $(\mathbb{X}, \mathfrak{F}, \mathcal{P})$, à valeurs dans $(\mathbb{Y}, \mathcal{G})$, est une application mesurable

$$S : \mathbb{X} \longrightarrow \mathbb{Y} \quad (4)$$

$$x \mapsto y = S(x) \quad (5)$$

- Pour une réalisation x de X , $y = S(x)$ réalisation de la v.a. $S(X)$.



Statistique = fonction mesurable des observations de l'expérience.
C'est simplement une variable aléatoire !

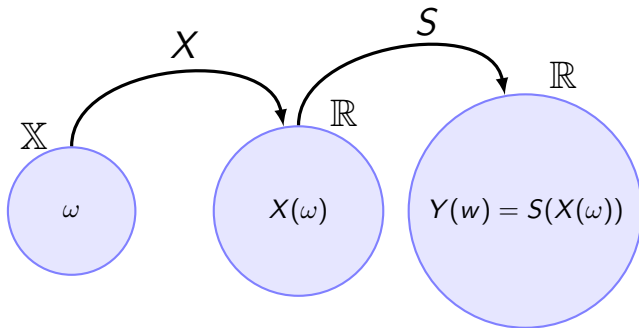


Figure 1: La statistique $S : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction mesurable qui transforme X en $Y = S(X)$.

Exemple: loi binomiale

- $X = (X_1, \dots, X_n)$ échantillon de Bernoulli de paramètre θ .

$$S(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i \quad (6)$$

Statistique = somme de toutes les observations du n -échantillon.

- **Statistique = fonction mesurable des données** $\Rightarrow$ définit une v.a. à partir des données initiales.
- **Loi image** $\mathbb{P}_S$ de $\mathbb{P}$ par $S \Rightarrow$ **modèle image**.
- Dans l'exemple, $S = S(X) = \sum_{i=1}^n X_i$.
- Loi image: loi binomiale $\mathbb{P}_S \sim \mathcal{B}(n, \theta)$. Modèle image: $(\mathbb{Y}, \mathcal{G}, \{\mathbb{P}_S : \mathbb{P} \in \mathcal{P}\})$. Considérer S plutôt que X :

$$\mathbb{P}[S \in B] = \mathbb{P}[S(X) \in B] = \mathbb{P}_S(B) = \mathbb{P}[X \in S^{-1}(B)], \quad \forall B \in \mathcal{B}_{\mathbb{R}}$$

Rappels sur la loi binomiale

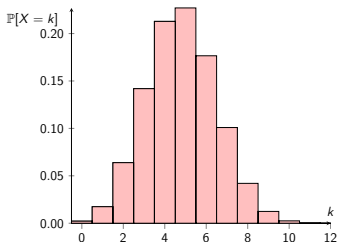
Definition

S suit une loi **binomiale** de paramètres $n \in \mathbb{N}$ et $p \in [0, 1]$ ssi pour tout $k = 0, \dots, n$,

$$\mathbb{P}[S = k] = \binom{n}{k} p^k (1-p)^{n-k} \quad (7)$$

et $\mathbb{P}[S = k] = 0$ ailleurs.

On note $S \sim \mathcal{B}(n, p)$. On a $\mathbb{E}[S] = np$ et $\mathbb{V}(S) = np(1-p)$.



Exemple: maximum d'une loi uniforme sur $[0, \theta]$.

- $X = (X_1, \dots, X_n)$ échantillon uniforme sur $[0, \theta]$.

$$S(X_1, X_2, \dots, X_n) = \max_{i=1}^n X_i \quad (8)$$

- L'information concernant la loi inconnue de X au travers de la statistique $S(X)$ est contenue dans la tribu $\mathfrak{F}(S) \subset \mathfrak{F}(X)$.
- $\mathfrak{F}(S) = \mathfrak{F}(X) \iff S$ bijective.
- Habituellement, $\mathfrak{F}(S)$ plus petite que $\mathfrak{F}(X)$ car on attend d'une statistique qu'elle réduise la tribu et condense l'information.
- Tribu = information.

2. Quelques types de modèles statistiques

Rappel: mesure dominée et théorème de Radon-Nikodym

- μ et ν deux mesures σ -finies sur $(\mathbb{X}, \mathfrak{F})$.
- On dit que ν est **absolument continue par rapport à μ** et l'on note $\nu \ll \mu$, si

$$\forall B \in \mathfrak{F}, \mu(B) = 0 \Rightarrow \nu(B) = 0 \quad (9)$$

- Le **théorème de Radon-Nikodym** assure que ν possède une **densité f** par rapport à μ définie par

$$\nu(B) = \int_B f(x) d\mu(x) \quad (10)$$

On note $f = d\nu/d\mu$. f est positive, mesurable et s'appelle densité ou dérivée au sens de Radon-Nikodym de ν par rapport à μ .

Definition

Un modèle statistique $(\mathbb{X}, \mathfrak{F}, \mathcal{P})$ est **dominé** si, et seulement si, tout $\mathbb{P} \in \mathcal{P}$ est dominé par une même mesure μ . μ est la **mesure dominante du modèle**.

Si un modèle est dominé par une mesure σ -finie μ , alors le modèle image est dominé par la mesure μ_S : $\forall B$ mesurable,

$$\mu_S(B) = 0 \iff \mu(S^{-1}(B)) = 0 \Rightarrow (\mathbb{P}(S^{-1}(B)) = 0 \iff \mathbb{P}_S(B) = 0)$$

Definition

Modèle statistique $(\mathbb{X}, \mathfrak{F}, \mathcal{P})$ **paramétré** si les éléments de $\mathcal{P}$ peuvent être décrit par un paramètre: c.-à-d. il existe Λ surjection:

$$\Lambda : \Theta \longrightarrow \mathcal{P} \quad (11)$$

$$\theta \mapsto \mathbb{P}_\theta \quad (12)$$

$\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Theta\}$. Si Λ bijection, on dit que le modèle est **identifiable**.

Modèle **paramétrique** si $\Theta \subset \mathbb{R}^p$ pour $p \in \mathbb{N}$. Sinon, **non-paramétrique**.

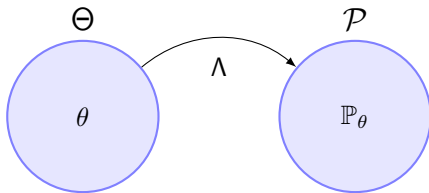


Figure 2: Modèle paramétrique.

Modèles paramétriques: exemples

- **Épreuve de Bernoulli**: $X \sim b(\theta)$ paramètre: $\theta \in [0, 1]$. Modèle paramétrique, identifiable, dominé par mesure de comptage sur $\mathbb{N}$. Si paramètre = $\cos^2 \theta$, modèle non identifiable car $\mathbb{P}_\theta = \mathbb{P}_{\theta+\pi}$.
- **Modèle d'échantillonnage gaussien**: X n -échantillon de loi $\mathcal{N}(m, \sigma^2)$.
Modèle statistique correspondant:

$$(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^{\otimes n}, \{\mathcal{N}(m, \sigma^2)^{\otimes n} : m \in \mathbb{R}, \sigma^2 > 0\}) \quad (13)$$

est paramétrique, dominé et identifiable. Paramétré par (m, σ^2) , dominé par la mesure de Lebesgue sur $\mathbb{R}^n$ et identifiable car...

- **Modèles non paramétriques**: ensemble des lois de probabilités sur $\mathbb{R}$ (paramètre $\notin$ à un espace vectoriel de dimension finie).
- **Modèles semi-paramétriques**: dépendant d'un paramètre $\theta \in \mathbb{R}^p$, mais aussi d'un second paramètre $\in$ à espace de dim. infinie. = terme de nuisance.

Definition

Soit $(\mathbb{X}, \mathfrak{F}, \{\mathbb{P}_\theta; \theta \in \Theta\})$ dominé par μ σ -finie.

La **vraisemblance** du modèle est la fonction $\theta \mapsto L(x, \theta)$ définie par

$$L(x, \theta) = d\mathbb{P}_\theta(x)/d\mu(x) \quad (14)$$

La vraisemblance est exactement la densité de la loi $\mathbb{P}_\theta$, vue comme fonction de θ : $\mathbb{P}_\theta(dx) = L(x, \theta)d\mu(x)$.

Soient $(\mathbb{X}, \mathfrak{F}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ et $\mathbb{P}_\theta \ll \lambda$ mesure de Lebesgue sur $\mathbb{R}$; supposons que $\mathbb{P}_\theta$ possède une densité continue $f_\theta(x)$ par rapport à λ et plaçons-nous dans un modèle d'échantillonnage. C'est un modèle paramétrique dominé par λ sur $\mathbb{R}^n$ et la vraisemblance de $\mathbb{P}_\theta^{\otimes n}$ relativement à $\lambda^{\otimes n}$ est

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f_\theta(x_i) \quad (15)$$

Exemple: vraisemblance d'un échantillon uniforme -1-

Soit $X = (X_1, \dots, X_n)$ n -échantillon uniforme sur $[0, \theta]$.

$$L(x, \theta) = L(x_1, \dots, x_n, \theta) \quad (16)$$

$$= \prod_{i=1}^n \left(\frac{1}{\theta} \mathbb{1}_{[0, \theta]}(x_i) \right) \quad (17)$$

$$= \frac{1}{\theta^n} \prod_{i=1}^n \mathbb{1}_{[0, \theta]}(x_i) \quad (18)$$

$$= \frac{1}{\theta^n} \mathbb{1}_{[0, \theta]^n}(x) = \frac{1}{\theta^n} \mathbb{1}_{[0 \leq \min x_i \leq \max x_i \leq \theta]} \quad (19)$$

avec $x = (x_1, \dots, x_n)$. Ici, **vraisemblance = produit des densités**.

L'indicatrice de $[0, \theta]^n \subset \mathbb{R}^n$ vaut 1 si, et seulement si, toutes les coordonnées x_i du vecteur x appartiennent à $[0, \theta]$.

Exemple: vraisemblance d'un échantillon uniforme -2-



Bien distinguer la densité d'une v.a. (fonction de x qui dépend de θ) de la vraisemblance d'un échantillon ou d'une v.a. (fonction de θ qui dépend de x).

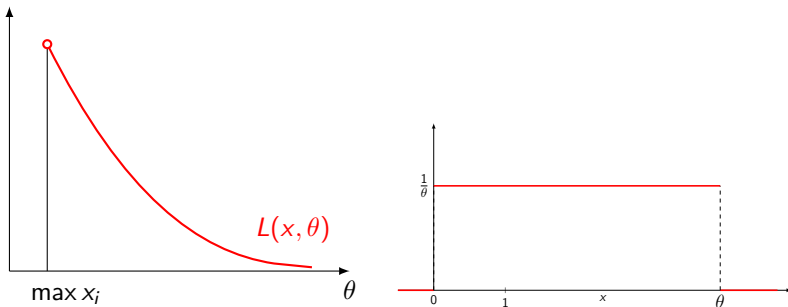


Figure 3: Vraisemblance échantillon uniforme (gauche), densité loi uniforme (droite).

Exemple: vraisemblance d'un échantillon de loi discrète

- Si $(\mathbb{X}, \mathfrak{F}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, où $\mathbb{P}_\theta \ll \delta$ mesure de comptage sur $\mathbb{N}$.

$$\mathbb{P}_\theta[X \in B] = \sum_{x \in B} \mathbb{P}_\theta[X = x] = \sum_{x \in \mathbb{N}} \mathbb{P}_\theta[X = x] \delta_x(B), \quad \forall B \in \mathbb{N} \quad (20)$$

Le modèle d'échantillonnage correspondant est dominé par la mesure discrète sur $\mathbb{N}^n$ et la **vraisemblance** de $\mathbb{P}_\theta^{\otimes n}$ par rapport à cette mesure est donnée par le **produit des probabilités**

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n \mathbb{P}_\theta[X_i = x_i] \quad (21)$$

Exemple: échantillon de Bernoulli de paramètre θ .

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n (\theta^{x_i} (1 - \theta)^{1-x_i} \mathbb{1}_{\{0,1\}}(x_i)) = \theta^s (1 - \theta)^{n-s} \mathbb{1}_{\{0,1\}^n}(x)$$

où $x = (x_1, \dots, x_n)$, $s = \sum_{i=1}^n x_i$ et $\mathbb{1}_{\{0,1\}^n}(x) = 1$ si $x_i = 0$ ou $1 \quad \forall i$.

Modèle homogène

- **Support** d'une mesure de probabilité $\mathbb{P}$ de densité f par rapport à μ :

$$\text{supp}(f) = \{x \in \mathbb{R} : f(x) > 0\} \quad (22)$$

Support défini à un ensemble de mesure nulle près.

Definition

Modèle paramétrique **homogène** si support des lois $\mathbb{P}_\theta$ est le même.

$\Rightarrow$ support ne dépend pas de $\theta \Rightarrow$ toutes les lois de $\mathcal{P}$ sont $\sim$:

$$\forall \theta, \theta', \mathbb{P}_\theta \ll \mathbb{P}_{\theta'} \quad (23)$$

Homogène $\Rightarrow$ dominé $\Rightarrow$ ensembles négligeables sont les mêmes $\forall \mathbb{P}_\theta$.

Homogène $\iff$

$$\exists \nu \text{ } \sigma\text{-finie} / \forall \theta, \mathbb{P}_\theta \ll \nu \text{ et } d\mathbb{P}_\theta/d\nu > 0 \text{ } \nu\text{-p.p.} \quad (24)$$

Modèle exponentiel: définition 1/7

Definition

Un modèle statistique $(\mathbb{X}, \mathfrak{F}, \{\mathbb{P}_\theta; \theta \in \Theta\})$ est **exponentiel** $\iff$

- Il est paramétrique, dominé et homogène.
- Sa vraisemblance s'écrit

$$L(x, \theta) = h(x) \exp(\langle \lambda(\theta), T(x) \rangle - \beta(\theta)) = h(x) e^{\lambda \bullet T - \beta} \quad (25)$$

- $h : \mathbb{R}^n \rightarrow \mathbb{R}$ mesurable positive (**mesure de base**).
- $T = (T_1, \dots, T_k) : \mathbb{X} \rightarrow \mathbb{R}^k$ mesurable (**statistique naturelle**).
- $\lambda(\theta) = (\lambda_1(\theta), \dots, \lambda_k(\theta)) : \Theta \subset \mathbb{R}^p \rightarrow \mathbb{R}^k$ (**paramètre du modèle**).
- $\beta : \mathbb{R}^p \rightarrow \mathbb{R}$ (**fonction de log-partition**).

Le produit scalaire de T par λ s'écrit également

$$\lambda \bullet T = \langle \lambda(\theta), T(x) \rangle = \sum_{i=1}^k \lambda_i(\theta) T_i(x) \quad (26)$$

- $\lambda(\theta) = \theta \Rightarrow$ paramètre canonique.
- $\Rightarrow T$ est dite statistique canonique du modèle ou statistique naturelle.
- Identifiable $\iff \theta \longrightarrow \lambda(\theta)$ injective.
- $\Rightarrow$ on peut reparamétriser le modèle avec le nouveau paramètre $\lambda \in \Lambda = \{\lambda(\theta); \theta \in \Theta\}$.
- $\lambda(\theta)$ non injective $\Rightarrow$ famille surparamétrée en θ .
- De plein rang si les $T_1(x), \dots, T_k(x)$ ne sont pas linéairement dépendants.
- Dans le polycopié: beaucoup plus de propriétés: transformée de Laplace de la mesure, convexité de l'espace canonique des paramètres, etc.

- Épreuve de Bernoulli.

Modèle dominé, paramétré par θ , homogène (support $\{0, 1\}$) et la vraisemblance s'écrit

$$L(x, \theta) = \theta^x (1 - \theta)^{1-x} = \exp(x \ln \theta + (1 - x) \ln(1 - \theta)) \quad (27)$$

$$= \exp\left(x \ln\left(\frac{\theta}{1 - \theta}\right) + \ln(1 - \theta)\right) \quad (28)$$

Le modèle est donc exponentiel avec pour paramètre

$$\lambda(\theta) = \ln\left(\frac{\theta}{1 - \theta}\right) \quad (29)$$

et pour statistique naturelle $T(x) = x$ et $\beta(\theta) = -\ln(1 - \theta)$.

- Échantillon de loi uniforme sur $[0, \theta]$.

Densité du n -échantillon par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$:

$$f_{\theta}(x_1, \dots, x_n) = \frac{1}{\theta^n} \mathbb{1}_{[0, \theta]^n}(x_1, \dots, x_n) \quad (30)$$

où l'indicatrice vaut 1 ssi $0 \leq x_i \leq \theta \ \forall \ i = 1, \dots, n$.



Modèle non homogène donc non exponentiel.

Modèle exponentiel: exemples 5/7

- Loi multinomiale.

$X = (X_1, \dots, X_p)$ vecteur multinomial, de paramètres (n, θ) n fixe, θ inconnu,

$$\Theta = \left\{ \theta = (\theta_1, \dots, \theta_p) \in [0, 1]^p / \sum_{i=1}^p \theta_i = 1 \right\} \quad (31)$$

Modélise le tirage de n objets appartenant à p catégories différentes: θ_i proportion d'objets de catégorie i , X_i nombre d'objets de catégorie i parmi les n objets tirés.

Modèle exponentiel, dominé par la mesure de comptage sur $\mathbb{N}^p$, homogène car le support de la loi est

$$\text{supp}(X) = B_n = \left\{ x \in \mathbb{N}^p : \sum_{i=1}^n x_i = n \right\} = \{x \in \mathbb{N}^p : s = n\} \quad (32)$$

Loi multinomiale

On effectue m tirages indépendants avec remise dans une urne contenant des boules de d couleurs.

Urne
Probabilités
des couleurs



$$\sum_{j=1}^d \theta_{jk} = 1$$



On tire m fois
avec remise
et on compte
les couleurs

Résultat d'un tirage de m boules



x_1
boules
rouges



x_2
boules
vertes

...



x_d
boules
bleues

Nombre total de boules : m

Le vecteur des comptages $X = (X_1, X_2, \dots, X_d)$ suit une loi multinomiale de paramètres m et $(\theta_{1k}, \theta_{2k}, \dots, \theta_{dk})$.

$$X \sim \mathcal{M}(m, \theta_{1k}, \theta_{2k}, \dots, \theta_{dk})$$

avec $\sum_{j=1}^d X_j = m$ et $X_j \geq 0$.

$$\begin{aligned} \mathbb{P}(X_1 = x_1, \dots, X_d = x_d) &= \frac{m!}{x_1! x_2! \dots x_d!} \theta_{1k}^{x_1} \theta_{2k}^{x_2} \dots \theta_{dk}^{x_d} \\ &\text{si } \sum_{j=1}^d x_j = m \text{ et } x_j \geq 0. \end{aligned}$$

Modèle exponentiel: exemples 6/7

- Loi multinomiale - suite et fin. Densité (discrète) :

$$\begin{aligned}h_{\theta}(x) &= \mathbb{P}_{\theta} [X_1 = x_1, \dots, X_p = x_p] \\&= \frac{n!}{x_1! \dots x_p!} \mathbb{1}_{B_n}(x) \prod_{i=1}^n \theta_i^{x_i} = \frac{n!}{x_1! \dots x_p!} \mathbb{1}_{B_n}(x) \exp \left(\sum_{i=1}^p x_i \ln \theta_i \right)\end{aligned}$$

- On pourrait choisir comme paramètre $(\ln \theta_1, \dots, \ln \theta_p)$ et pour statistique $X_1, \dots, X_p$, $\Rightarrow$ surparamétrage car $\sum X_i = n \Rightarrow (X_1, \dots, X_p)$ liée. Elle est de rang $p - 1$.
- On préférera choisir comme statistique $T(X) = (X_1, \dots, X_{p-1})$ et paramètre $\lambda(\theta) = (\ln(\theta_1/\theta_p), \dots, \ln(\theta_{p-1}/\theta_p))$ et $\beta(\theta) = -n \ln \theta_p$. Ici, $k = p - 1$.
- Représentation minimale si les T_i forment une famille libre. Ordre = cardinal de la famille $(T_i)_i$.

Modèle exponentiel: exemples 7/7

- Échantillon gaussien.

$X_1, \dots, X_n$ v.a.i.i.d. de loi $\mathcal{N}(m, \sigma^2)$. $\theta = (m, \sigma^2)$. Modèle exponentiel (en exercice). Idem si σ^2 est connu et $\theta = m$, ou si m connu et $\theta = \sigma^2$ (en exercice).

- Modèles binomial, de Poisson, gamma.

Un modèle binomial de paramètres (n, θ) est exponentiel et la statistique X est le paramètre naturel (dans le modèle d'échantillonnage correspondant, $S_n = \sum X_i$ est le paramètre naturel). Idem si modèle de Poisson de paramètre θ .

Modèle gamma de paramètres (α, β) , de densité:

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{1}_{\mathbb{R}_+}(x) \quad (33)$$

le paramètre naturel est donné par la statistique $(X, \ln X)$ (le prouver).

Soit $\mathcal{O}$ une tribu sur l'ensemble Θ des paramètres.

Definition

Un **modèle bayésien** est un modèle statistique $(\mathbb{X}, \mathfrak{F}, \{\mathbb{P}_\theta; \theta \in \Theta\})$ muni d'une loi de probabilité Π sur l'espace $(\Theta, \mathcal{O})$

- Π s'appelle la loi **a priori**. Dans un tel modèle, une valeur θ du paramètre est la réalisation d'une variable aléatoire T de loi Π .

On dispose de 3 lois: la loi **a priori** Π , la **vraisemblance** $L(x|\theta)$ qui est également la loi conditionnelle de X sachant θ (loi conditionnelle des effets sachant les causes) et la loi **a posteriori** notée $L(\theta|x)$ (loi des causes sachant les effets).

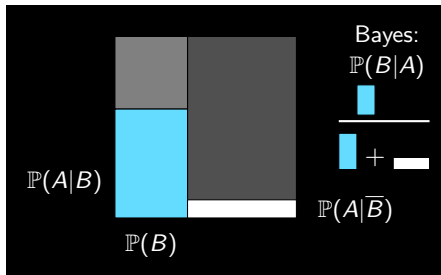
Modèle bayésien: lien *a priori*, *a posteriori*, vraisemblance 2/5

T une v.a. de loi Π . La **formule de Bayes** donne

$$f_{T/X}(\theta, x) = \frac{f_{X/T}(x, \theta) f_T(\theta)}{f_X(x)} \Rightarrow L(\theta|x) = \frac{L(x|\theta) \Pi(\theta)}{f(x)}$$

Theorem

La vraisemblance **a posteriori** $L(\theta|x)$ est proportionnelle à $L(x|\theta) \times \Pi(\theta)$



Modèle bayésien: exemple échantillon de Bernoulli 3/5

- $X = (X_1, \dots, X_n)$ ***n*-échantillon de Bernoulli** de paramètre θ inconnu.
- Modélise lancer de n tirages à pile ou face; on souhaite déterminer θ .
On pense que la pièce est truquée et des observations passées amènent à définir la loi ***a priori*** de θ comme une ***loi béta*** de paramètres (a, b) .
Vraisemblance de l'échantillon:

$$L(x, \theta) = \prod_{i=1}^n L(x_i, \theta) = \theta^s (1 - \theta)^{n-s} \quad (34)$$

avec $s = \sum_{i=1}^n x_i$. D'après la formule de Bayes, la ***vraisemblance a posteriori*** sera proportionnelle à

$$L(\theta|x) \propto \theta^s (1 - \theta)^{n-s} \frac{1}{B(a, b)} \theta^{a-1} (1 - \theta)^{b-1} \mathbb{1}_{[0,1]}(\theta) \quad (35)$$

À une constante multiplicative près, il s'agit de la densité d'une ***loi béta*** de paramètres $(s + a, n - s + b)$.

Modèle bayésien: exemple échantillon gaussien 4/5

$X = (X_1, \dots, X_n)$ **n -échantillon gaussien** $\mathcal{N}(\theta, \sigma^2)$. θ inconnu, σ^2 connue.
Loi *a priori* gaussienne $\mathcal{N}(a, b^2)$ (a, b^2 hyperparamètres):

$$\Pi(\theta) = (2\pi b^2)^{-1/2} \exp\left(-\frac{1}{2b^2}(\theta - a)^2\right) \quad (36)$$

Vraisemblance de l'échantillon:

$$L(x|\theta) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2\right) \quad (37)$$

Loi jointe de (X, T) :

$$f_{(X,T)}(x, \theta) = (2\pi b^2)^{-1/2} (2\pi\sigma^2)^{-n/2} e^{-\frac{\theta^2}{2}\left(\frac{1}{b^2} + \frac{n}{\sigma^2}\right) + \theta\left(\frac{a}{b^2} + n\frac{\bar{x}}{\sigma^2}\right) - \left(\frac{a^2}{2b^2} + \frac{n\bar{x}^2}{2\sigma^2}\right)}$$

où $\bar{x} = (\sum x_i)/n$ et $\bar{x^2} = (\sum x_i^2)/n$.

Modèle bayésien: échantillon gaussien 5/5

Posons τ et m tels que:

$$\begin{cases} \frac{1}{\tau^2} = \frac{1}{b^2} + \frac{n}{\sigma^2} \\ \frac{m}{\tau^2} = \frac{a}{b^2} + \frac{n\bar{x}}{\sigma^2} \end{cases} \Rightarrow f_{(X, \tau)}(x, \theta) = \gamma \exp(-\Lambda/2)$$

$$\text{avec } \begin{cases} \gamma = (2\pi b^2)^{-1/2} (2\pi\sigma^2)^{-n/2} \\ \Lambda = \frac{1}{\tau^2} (\theta^2 - 2m\theta) + \frac{a^2}{b^2} + \frac{n\bar{x}^2}{\sigma^2} \end{cases} \quad \text{😱}$$

$$\begin{aligned} f_{(X, \tau)}(x, \theta) &= \gamma \times \exp\left(\frac{m^2}{2\tau^2} - \frac{a}{2b^2} - \frac{n\bar{x}^2}{2\sigma^2}\right) \times \exp\left(-\frac{1}{2\tau^2}(\theta - m)^2\right) \\ &= \text{constante} \times C(x) \times \mathcal{N}(m, \tau^2) \end{aligned}$$

$C(x)$ ne dépend pas de θ . Densité de X de la forme $C(x) \times (2\pi\tau^2)^{-1/2}$
 $\Rightarrow$ **vraisemblance a posteriori** $\sim \mathcal{N}(m, \tau^2)$.

$n \uparrow \Rightarrow$ observations de + en + importantes / à l'a priori.

$$\overbrace{\mathbb{P}(A \mid B)}^{a \text{ posteriori}} = \frac{\overbrace{\mathbb{P}(B \mid A)}^{\text{vraisemblance}} \cdot \overbrace{\mathbb{P}(A)}^{a \text{ priori}}}{\underbrace{\mathbb{P}(B)}_{\text{marginale}}}$$

À retenir : ce qu'est

- un modèle aléatoire.
- une statistique.
- le théorème de Radon-Nikodym.
- la vraisemblance.
- un modèle d'échantillonnage, dominé, homogène, paramétrique, exponentiel, bayésien.