MA201 - Séance 3
Construction d'Estimateurs Bayésiens
Modèle de mélanges et Algorithme EM

H. Piet-Lahanier - L. Meyer

Estimateur

Rappel du problème

- On dispose d'un ensemble d'*observations* $z_1, \dots, z_n$.
- Ces observations sont les réalisations d'un n-échantillon $Z_1, \dots, Z_n$ suivant une loi P_θ, avec $\theta \in \Theta$.
- L'*estimation* consiste à retrouver θ, le ou les paramètres caractéristiques de la loi P_θ ayant servi à générer ces observations
- Approche fréquentiste : le paramètre θ est déterministe.
- Approche Bayésienne : le paramètre θ est une variable aléatoire.

Estimateurs Bayésiens

Les principaux estimateurs Bayésiens

- Estimateur du Maximum (de Vraisemblance) a Posteriori
- Estimateur de la moyenne quadratique minimale
- Estimateur de Bayes : $\widehat{\theta} = T^{\pi}(\boldsymbol{z}) = arg \min_{T} \mathrm{E}^{\pi}[C(\theta, T)|z]$, où C est un coût.

Exemples de définitions d'estimateurs de Bayes

Risque associé au coût en valeur absolue

La fonction de coût en valeur absolue $\ell(\theta, T(z))$ pour l'estimateur $T(z)$ est la fonction définie par :

$$\ell(\theta, T(z)) = k_2(\theta - T(z)), \theta > T(z) \quad (1)$$

$$\ell(\theta, T(z)) = k_1(T(z) - \theta), \theta \leq T(z) \quad (2)$$

$$(3)$$

Un estimateur de Bayes associé à $\pi(\theta)$ et à ℓ est un fractile (ou quantile) d'ordre $\frac{k_2}{k_1+k_2}$ de la loi conditionnelle de $\theta|z$

Exemples de définitions d'estimateurs de Bayes

Risque associé au coût en valeur absolue

$$\begin{aligned} \mathrm{E}^{p_{\theta|z}}[\ell(\theta, T(z))] &= \int_{\Theta} \ell(\xi, T(z)) p_{\theta|z} d\xi \\ &= \int_{-\infty}^{T(z)} k_1(T(z) - \xi) p_{\theta|z} d\xi \\ &+ \int_{T(z)}^{+\infty} k_2(\xi - T(z)) p_{\theta|z} d\xi \end{aligned}$$

avec $p_{\theta|z} d\xi = df_{\theta|z} = -d(1 - f_{\theta|z})$

Exemples de définitions d'estimateurs de Bayes

Risque associé au coût en valeur absolue

$$
\begin{aligned}
\mathsf{E}^{p_{\theta|z}}[\ell(\theta, T(z))] &= [k_1(T(z)-\xi)(f_{\theta|z})]_{-\infty}^{T(z)} \\
&+ \int_{-\infty}^{T(z)} k_1 P_{\theta|z}(\theta<\xi)d\xi \\
&+ [k_2(\xi - T(z))(1-f_{\theta|z})]_{T(z)}^{+\infty} \\
&+ \int_{T(z)}^{+\infty} k_2 P_{\theta|z}(\theta>\xi)d\xi \\
&= \int_{-\infty}^{T(z)} k_1 P_{\theta|z}(\theta<\xi)d\xi + \int_{T(z)}^{+\infty} k_2 P_{\theta|z}(\theta>\xi)d\xi
\end{aligned}
$$

Exemples de définitions d'estimateurs de Bayes

Risque associé au coût en valeur absolue

$$\mathrm{E}^{p_{\theta|z}}[\ell(\theta, T(z))] = \int_{-\infty}^{T(z)} k_1 P_{\theta|z}(\theta < \xi) d\xi + \int_{T(z)}^{+\infty} k_2 P_{\theta|z}(\theta > \xi) d\xi$$

Recherche de l'estimateur

- Dérivation de l'espérance par rapport à $T(z)$ et de la valeur d'annulation de la dérivée
- $T(z)$ doit satisfaire : $P_{\theta|z}(\theta < T(z)) = \frac{k_2}{k_1+k_2}$.
- Valeur optimale de $T(z)$, estimateur de Bayes : valeur pour laquelle la probabilité que θ soit inférieur à la valeur de l'estimateur est égale à $\frac{k_2}{k_1+k_2}$

On retrouve $T(z)$ médiane pour la loi a posteriori pour $k_1 = k_2$

Exemples de définitions d'estimateurs de Bayes

Risque associé au coût 0-1

Fonction de coût 0-1 $\ell(\theta, T(z))$ pour l'estimateur $T(z)$ définie par :

$$\begin{aligned} \ell(\theta, T(z)) &= 0, H_0 \ \theta = T(z) \\ \ell(\theta, T(z)) &= 1, H_1 \ \theta \neq T(z) \end{aligned} \tag{4}$$

Le coût est nul lorsque le choix est bon,
Le coût est identique pour toutes les autres décisions
Choisir l'estimateur pour lequel le risque (moyenne du coût sur les réalisations possible) est minimum $\Rightarrow$ Sélectionner l'estimateur qui correspond à la valeur de θ la plus probable

Estimateur de Bayes : Maximum de la loi *a posteriori* $f_{\theta|z}$

$\Rightarrow T(z) = argmax_{\theta \in \Theta} f_{\theta|z}$

Exemples de définitions d'estimateurs de Bayes

Risque minimax

Un estimateur de Bayes pour π a un risque qui minimise le risque bayésien pour π,
Ce risque est une moyenne des risques ponctuels en θ suivant la loi *a priori* π sur θ.
Cet estimateur minimise donc un risque en moyenne selon π.
Risque maximal d'un estimateur :
$\mathcal{R}_{max}(T) = sup_{\theta \in \Theta} R(\theta, T)$
Le *Risque minimax* est la plus petite valeur de ce risque maximal pour l'ensemble des estimateurs.
$\mathcal{R}_M = inf_T \mathcal{R}_{max}(T)$
Un estimateur T est *minimax* si le Risque maximal de T est égal au risque minimax $\mathcal{R}_M$
Cet estimateur est le *meilleur dans le pire des cas*
Remarque : Un estimateur admissible avec un risque constant est minimax

Choix du modèle des observations

Choix du modèle

Construction de la relation entre les observations et θ
Utilisation de Lois de probabilités associées à un type de réalisation
Exemples : Lois continues

- Loi uniforme continue : équiprobabilité de tous les points d'un intervalle
- Loi exponentielle : temps d'attente avant l'occurence d'un événement aléatoire dans un processus de Poisson
- Loi Normale (Gauss) : Somme d'un grand nombre de variables indépendantes, de moyenne et variance finies
- Loi Gamma : généralisation de la loi exponentielle à n événements

Choix du modèle des observations

Choix du modèle

Construction de la relation entre les observations et θ
Utilisation de Lois de probabilités associées à un type de réalisation
Exemples : lois discrètes

- Loi uniforme discrète : n tirages équiprobables ; lancers de dés, roulette
- Loi de Bernoulli : tirage aléatoire à deux résultats possible ; pile ou face , sain ou défectueux
- Loi Binomiale : nombre de succès d'une série de n tirages Bernoulli indépendants
- Loi Géométrique : nombre d'essais nécessaires, dans une suite de tirages de Bernoulli, avant d'obtenir un succès.
- Loi de Poisson : probabilité d'observer un certain nombre d'événements aléatoires dans un intervalle continu

Modèle de mélange

Modèle de mélange

Soit des observations $(z_1, \dots, z_n)$ provenant d'un mélange de deux gaussiennes :
$f_\theta(z) = \frac{1}{2}\mathcal{N}(m_1, \sigma_1)(z) + \frac{1}{2}\mathcal{N}(m_2, \sigma_2)(z)$
où $\mathcal{N}(m_i, \sigma_i)$ est la densité gaussienne d'espérance m_i et de variance σ_i^2 et $\theta = (m_1, \sigma_1, m_2, \sigma_2)$.

Dans ce cas, l'estimateur du maximum de vraisemblance n'est pas consistant :
La vraisemblance de $\theta = (Z_1, \sigma_1,\ 0,\ 1)$ tend vers l'infini quand σ_1 tend vers 0.
L'ensemble Θ n'est pas compact.

Problème : Déterminer un estimateur pour ce type de modèles

Modèle de mélange

Principe

Soit un ensemble de K lois de probabilités (continues ou discrètes) :

- $\mathbb{P}_i$, $i \in \{1, ..., K\}$.

On sait que la variable observée suit UNE de ces K lois de probabilités.
Mais on ne sait pas précisément laquelle !
Par contre, on sait que le choix de cette loi de probabilité s'effectue selon une loi de probabilité *discrète* connue :

- La probabilité que la variable observée suive la loi $\mathbb{P}_i$ est p_i. (avec $\sum_{i=1}^{K} p_i = 1$).

Nous considérons donc $K + 1$ lois de probabilités : K lois continues ou discrètes et 1 loi discrète.

Définition

Définition

Une densité mélange est une fonction de densité qui est une combinaison linéaire convexe de plusieurs fonctions de densités.
Autrement dit, f est une densité mélange s'il existe $K \in \mathbb{N}$, des densités f_1, ..., f_K, et des réels p_1, ..., p_K tels que $\sum_{i=1}^{K} p_i = 1$ et

$$f(x) = \sum_{i=1}^{K} p_i f_i(x). \tag{5}$$

En écrivant les p_i comme des $\mathbb{P}(Z = i)$, avec Z variable aléatoire discrète à valeurs dans $\{1, ..., K\}$, on a alors :

$$f(x) = \sum_{i=1}^{K} \mathbb{P}(Z = i) f_i(x). \tag{6}$$

Cela signifie que pour générer une réalisation d'une variable aléatoire de densité f, il faut d'abord :

- générer une réalisation de la variable aléatoire Z pour obtenir un nombre j dans $\{1, ..., K\}$,
- puis générer une réalisation de la variable aléatoire de densité f_j

Exemple 1

Exemple

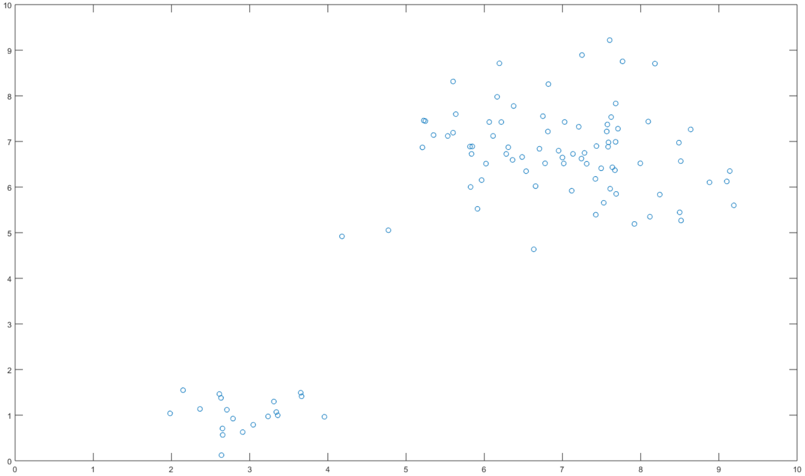
Considérons le mélange de deux lois gaussiennes bidimensionnelles :

- loi 1 : $\mathcal{N}(\begin{bmatrix} 3 \\ 1 \end{bmatrix}, 0.5 I_2)$
- loi 1 : $\mathcal{N}(\begin{bmatrix} 7 \\ 7 \end{bmatrix}, I_2)$

Le choix entre ces deux lois s'opère via une loi de Bernoulli $\mathcal{B}(0.3)$.

Exemple 1

Exemple de génération d'une telle loi pour $n = 100$ tirages.



Question : comment retrouver les paramètres des deux lois à partir des mesures ?

Exemple 2

Exemple

Un joueur disposes de deux pièces :

- l'une a une probabilité p_1 d'obtenir pile,
- l'autre a une probabilité p_2 d'obtenir pile.

Avant chaque tour, le joueur choisit la pièce qu'il lance selon une loi de Bernoulli de paramètre λ :

- il a une probabilité λ de choisir la pièce 1,
- il a une probabilité $1 - \lambda$ de choisir la pièce 2.

Un tour de jeu consiste en une série de trois lancers consécutifs avec la même pièce.
Par exemple, une série de 5 tours de jeu pourra donner lieu au relevé suivant :

$$\{P, P, F\}; \qquad \{F, P, F\}; \qquad \{F, F, P\}; \qquad \{F, P, F\}; \qquad \{P, P, P\}. \tag{7}$$

Exemple 2

Exemple

Formalisation de l'exemple précédent : mélange de deux lois binomiales :

- loi 1 : $X_1 \sim \mathcal{B}(p_1, 3)$ est le nombre de pile sur 3 lancers de la pièce 1
- loi 2 : $X_2 \sim \mathcal{B}(p_2, 3)$ est le nombre de pile sur 3 lancers de la pièce 2

Le choix entre ces deux lois s'opère via une loi de Bernoulli $Z \sim \mathcal{B}(\lambda)$ est le numéro de pièce utilisée (pièce 1 si le résultat du lancer est 1), pièce 2 sinon).
Le vecteur de paramètres est donc $\theta = (\lambda, p_1, p_2)$.
Question : comment estimer θ à partir des résultats d'une succession de N occurrences de 3 lancers ?

Problématique

Objectifs de l'algorithme EM

- Si on sait à quelle loi correspond chaque mesure (loi X_1 ou loi X_2), l'estimateur du maximum de vraisemblance permet de déterminer les paramètres respectifs de ces lois (en séparant les mesures selon la loi avec laquelle elles ont été générées).
- Problème : nous ne savons pas *a priori* quelle loi a été utilisée pour générer une mesure donnée !
  On dit qu'on est en présence de variables cachées (ou latentes) Z_i : c'est la v.a. *discrète* utilisée pour effectuer le choix de la loi à utiliser.
- L'algorithme EM résout ce type de problème.
- On considère le vecteur de paramètres incluant les paramètres suivant :
  - Paramètres inconnus de chacune des K lois, (θ des modèles paramétriques précédents)
  - Paramètres inconnus de la loi *discrète* globale effectuant le choix entre les différentes lois.

Dans l'exemple 2 (lancers de pièces), le vecteur de paramètres est $\theta = (\lambda, p_1, p_2)$.

Principe et Intuition de l'algorithme EM pour le modèle de mélange

Principes de l'algorithme EM

Soit un n-échantillon $(X_1, ..., X_n)$ correspondant à un densité de mélange f.

- Algorithme itératif : on part d'une valeur arbitraire θ_0, et on calcule $\theta_{m+1} = EM(\theta_m)$.
- A chaque itération, la vraisemblance de la loi mélange augmente : $f(\theta_{m+1}; x_1, ..., x_n) \geq f(\theta_m; x_1, ..., x_n)$.
- Mais, il n'y a pas forcément convergence de l'algorithme vers le maximum de vraisemblance : possibilité de maximum local => très dépendant du choix de la valeur θ_0.
- Pour outrepasser la limite précédente, on lance l'algorithme EM pour plusieurs valeurs du paramètre initial θ_0.

Principe et Intuition de l'algorithme EM

Intuition de l'algorithme

- Rappel : si l'on connaissait les réalisations des variables cachées Z_i, l'estimation serait facilitée.
- Algorithme fonctionnant en deux étapes :
  1. Etape E - Expectation :
     On suppose que θ_m est le vecteur de paramètres inconnus. On en déduit les probabilités $\tilde{p}_{ij,m}$ d'obtenir les variables cachées $Z_i = j = \{1, ..., K\}$, compte-tenu des valeurs des réalisations $x_1, ..., x_n$ et le vecteur de paramètres θ_m (estimé à l'issue de l'étape m).
  2. Etape M - Maximization :
     En partant de ces probabilités $\tilde{p}_{ij}$, on calcule la valeur de θ qui maximise la vraisemblance : θ_{m+1}.

Description de l'algorithme

Algorithme

Le passage de l'étape m à l'étape $m+1$ consiste à effectuer :

$$\theta_{m+1} = arg \max_{\theta \in \Theta} \left\{ \sum_{i=1}^{n} E_{Z_i|X_i,\theta_m} \left[\log \mathbb{P}(X_i = x_i, Z_i|\theta) \right] \right\}, \quad (8)$$

l'espérance étant considérée sur Z. Cette équation peut se réécrire :

$$\theta_{m+1} = arg \max_{\theta \in \Theta} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{K} \tilde{p}_{ij} \left[\log \mathbb{P}(X_i = x_i, Z_i = j|\theta) \right] \right\}, \quad (9)$$

avec $\tilde{p}_{ij,m} = \mathbb{P}(Z_i = j|X_i = x_i, \theta_m)$.

Description de l'algorithme

Algorithme

En utilisant la formule de Bayes et la formule des probabilités totales, on peut écrire les équations de l'algorithme :

- Etape (E) - Expectation - Calcul des probabilités *a posteriori* (permettant d'écrire l'espérance en Z) :

$$\tilde{p}_{ij,m} = \frac{\mathbb{P}(X_i = x_i | Z_i = j, \theta_m)\mathbb{P}(Z_i = j | \theta_m)}{\sum_{k=1}^{K} \mathbb{P}(X_i = x_i | Z_i = k, \theta_m)\mathbb{P}(Z_i = k | \theta_m)} = \frac{f_j(x_i; \theta_m) p_j(\theta_m)}{\sum_{k=1}^{K} f_k(x_i; \theta_m) p_k(\theta_m)}. \tag{10}$$

- Etape (M) - Maximization - Maximisation de la vraisemblance :

$$\begin{aligned} \theta_{m+1} &= arg \max_{\theta \in \Theta} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{K} \tilde{p}_{ij,m} \log \mathbb{P}(X_i = x_i | Z_i = j, \theta)\mathbb{P}(Z_i = j; \theta) \right\} \\ &= arg \max_{\theta \in \Theta} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{K} \tilde{p}_{ij,m} \log f_j(x_i; \theta) p_j(\theta) \right\}. \end{aligned} \tag{11}$$

Description de l'algorithme

Algorithme

En résumé, l'algorithme s'écrit donc :

- Etape (E) - Expectation - Calcul des probabilités *a posteriori* (permettant d'écrire l'espérance en Z) :

$$\tilde{p}_{ij,m} = \frac{f_j(x_i;\theta_m)p_j(\theta_m)}{\sum_{k=1}^{K} f_k(x_i;\theta_m)p_k(\theta_m)}. \tag{12}$$

- Etape (M) - Maximization - Maximisation de la vraisemblance :

$$\theta_{m+1} = arg \max_{\theta\in\Theta} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{K} \tilde{p}_{ij} \log f_j(x_i;\theta)p_j(\theta) \right\}. \tag{13}$$

Exemple

Reprenons l'exemple des lancers de pièce de monnaie

Exemple 1

Rappel du problème :

- Soit un n-échantillon de couples $(X_1, Z_1), ..., (X_n, Z_n)$, où :
  - chaque Z_i (variables cachées) suit une loi $\mathcal{B}(\lambda)$: si $z_i = 1$, le joueur utilisera la pièce 1, si $z_i = 0$, le joueur utilisera la pièce 2,
  - chaque X_i (variables observées) est le nombre de piles obtenus (sur trois lancers) au tour i, $i \in \{1, ..., n\}$.
- Si $Z_i = 1$, alors $X_i \sim \mathcal{B}(p_1, 3)$ est le nombre de pile sur 3 lancers de la pièce 1
- Si $Z_i = 2$, alors $X_i \sim \mathcal{B}(p_2, 3)$ est le nombre de pile sur 3 lancers de la pièce 2
- Le vecteur des paramètres inconnus est donc $\theta = (\lambda, p_1, p_2)$.

Notons $\theta_m = (\lambda_m, p_{1,m}, p_{2,m})$ le vecteur des paramètres estimées à l'issue de l'étape m de l'algorithme.

Exemple

Etape (E)

Pour $j = 1$, on a :

$$\tilde{p}_{i,1,m} = \frac{f_1(x_i;\theta_m)p_1(\theta_m)}{\sum_{k=1}^{K} f_k(x_i;\theta_m)p_k(\theta_m)} = \frac{p_{1,m}^{x_i}(1-p_{1,m})^{3-x_i}\lambda}{p_{1,m}^{x_i}(1-p_{1,m})^{3-x_i}\lambda + p_{2,m}^{x_i}(1-p_{2,m})^{3-x_i}(1-\lambda)} \tag{14}$$

Pour $j = 2$, on a :

$$\tilde{p}_{i,2,m} = \frac{f_2(x_i;\theta_m)p_2(\theta_m)}{\sum_{k=1}^{K} f_k(x_i;\theta_m)p_k(\theta_m)} \tag{15}$$

$$\tilde{p}_{i,2,m} = \frac{p_{2,m}^{x_i}(1-p_{2,m})^{3-x_i}(1-\lambda)}{p_{1,m}^{x_i}(1-p_{1,m})^{3-x_i}\lambda + p_{2,m}^{x_i}(1-p_{2,m})^{3-x_i}(1-\lambda)} = 1 - \tilde{p}_{i,1} \tag{16}$$

Exemple

Etape (M)

L'équation (M) s'écrit avec (rappel) : $\theta = (\lambda, p_1, p_2)$:

$$\theta_{m+1} = arg \max_{\theta \in \Theta} \left\{ \sum_{i=1}^{n} \tilde{p}_{i1} \log p_1^{x_i} (1 - p_1)^{3 - x_i} \lambda + \tilde{p}_{i2} \log p_2^{x_i} (1 - p_2)^{3 - x_i} (1 - \lambda) \right\}, \quad (17)$$

Il reste ensuite à trouver le θ maximisant l'expression ci-dessus. Les équations

$$\frac{\partial \{.\}}{\partial \lambda} = 0, \qquad \frac{\partial \{.\}}{\partial p_1} = 0, \qquad \frac{\partial \{.\}}{\partial p_2} = 0, \quad (18)$$

donnent

$$\hat{\lambda}_{m+1} = \frac{1}{n} \sum_{i=1}^{n} \tilde{p}_{i,1}, \qquad \hat{p}_{2,m+1} = \frac{\sum_{i=1}^{n} x_i \tilde{p}_{i,1}}{3 \sum_{i=1}^{n} \tilde{p}_{i,1}}, \qquad \hat{p}_{2,m+1} = \frac{\sum_{i=1}^{n} x_i \tilde{p}_{i,2}}{3 \sum_{i=1}^{n} \tilde{p}_{i,2}}. \quad (19)$$

D'où l'on déduit la valeur de $\theta_{m+1} = (\hat{\lambda}_{m+1}, \hat{p}_{1,m+1}, \hat{p}_{2,m+1})$.

Bibliographie

- Jean-François Delmas, *Introduction aux probabilités et à la statistique*, Les Presses de l'Ensta
- Stephan Morgenthaler, *Génétique statistique*, Springer, 2008.
- Stephan Morgenthaler, *Introduction à la statistique*, PPUR, 2013.
- Maxime Ossonce, *Statistique*, Les Presses de l'Ensta
- Christophe Pouet, *Probabilités et Statistique*, Les Presses de l'Ecole Centrale de Marseille
- Frédéric Santos, *L'algorithme EM : une courte présentation*, Les Presses du CNRS

Lois discrètes

Loi	Notation	Proba.
Uniforme discrète	$\mathcal{U}[a,\ b]$	$P(X) = \frac{1}{b-a+1}$ si $a \leq x \leq b$; $= 0$ sinon.
Bernoulli	$\mathcal{B}(\theta)$	$P(X=0) = 1-\theta$; $P(X=1) = \theta$
Binomiale	$\mathcal{B}(n,\theta)$	$P(X=k) = C_n^k \theta^k (1-\theta)^{n-k}$
Géométrique	$\mathcal{G}(\theta)$	$P(X=k) = (1-\theta)^{k-1}\theta$
Poisson	$\mathcal{P}(\theta)$	$P(X=k) = \exp(-\theta)\frac{\theta^k}{k!}$

Lois discrètes

Loi	Notation	Espérance	Variance
Uniforme discrète	$\mathcal{U}[a,\ b]$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
Bernoulli	$\mathcal{B}(\theta)$	θ	$\theta(1-\theta)$
Binomiale	$\mathcal{B}(n,\theta)$	$n\theta$	$n\theta(1-\theta)$
Géométrique	$\mathcal{G}(\theta)$	$\frac{1}{\theta}$	$\frac{1-\theta}{\theta^2}$
Poisson	$\mathcal{P}(\theta)$	θ	θ

Lois discrètes

Propriétés

- La somme de n v. a. indépendantes suivant la loi de Bernoulli $\mathcal{B}(p)$ suit une loi binomiale $\mathcal{B}(n, p)$
- La somme de deux v. a. indépendantes suivant les lois binomiales $\mathcal{B}(n, p)$ et $\mathcal{B}(m, p)$ suit une loi binomiale $\mathcal{B}(m + n, p)$
- La somme de deux v. a. indépendantes suivant les lois de Poisson $\mathcal{P}(\lambda_1)$ et $\mathcal{P}(\lambda_2)$ suit une loi de Poisson $\mathcal{P}(\lambda_1 + \lambda_2)$

Lois continues

Loi	Notation	Densité
Uniforme continue	$\mathcal{U}[a,\ b]$	$f_X(x) = \frac{1}{b-a}\mathbb{1}_{[a,\ b]}(x)$
Exponentielle	$\mathcal{E}(\lambda)$	$f_X(x) = \lambda \exp(-\lambda x)\mathbb{1}_{\mathbb{R}^+}(x)$
Normale	$\mathcal{N}(m, \sigma^2)$	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x-m)^2}{2\sigma^2})$
Gamma	$\Gamma(a, \lambda)$	$f_X(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1}\exp(-\lambda x)\mathbb{1}_{\mathbb{R}^{+*}}(x)$

NB : Fonction Gamma : $\Gamma(a) = \int_0^{+\infty} x^{a-1}\exp(-x)dx$

Lois continues

Loi	Notation	Espérance	Variance
Uniforme continue	$\mathcal{U}[a,\ b]$	$\frac{a+b}{2}$	$\frac{(b+a)^2}{12}$
Exponentielle	$\mathcal{E}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normale	$\mathcal{N}(m,\sigma^2)$	m	σ^2
Gamma	$\Gamma(a,\lambda)$	$\frac{a}{\lambda}$	$\frac{a}{\lambda^2}$

Lois continues

Propriétés

- La somme de n v. a. indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$ suit une loi Gamma $\Gamma(n, \lambda)$
- La somme de deux v. a. indépendantes suivant les lois Gamma $\Gamma(a, \lambda)$ et $\Gamma(b, \lambda)$ suit une loi Gamma $\Gamma(a + b, \lambda)$
- La somme de deux v. a. indépendantes suivant les lois normales $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ suit une loi normale $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$
- La somme de n v. a. indépendantes suivant la loi normale $\mathcal{N}(0, 1)$ suit la loi normale $\mathcal{N}(0, n)$
- La somme des carrés de n v. a. indépendantes suivant la loi normale $\mathcal{N}(0, 1)$ suit la loi du Khi-deux $\chi^2(n) = \Gamma(\frac{n}{2}, \frac{1}{2})$

Ce qu'il faut retenir

- Exemples d'estimateur de risque bayésien pour différents risques
- Choix de la structure du modèle d'obesrvations : sélection des lois de probabilité en fonction du problème abordé
- Estimateur pour des modèles de mélange : démarche et algorithme
- Rappels sur les lois de probabilités