Cours: MAPSI

Charles Vin

2022

Nouveau cours du 13/09

1 Introduction

- Exam final : 50%
- Partiel : 35%
- Participation : 15%
  - travail dans la séance
  - TME soumis en fin de séance omg

Deux grand type de modèle :

- Modèle paramétrique : connaissance sur la distribution stat des données. Puis on estime les paramètres de la loi.
- Modèle non paramétrique : l'inverse, on ne connait pas la loi. exemple : regression logistique

Echantillons :

- population
- ect

Définition 1.1. Vocabulaire :

- Voir diapo 9/51

Définition 1.2 (Mesure de proba)**.** Une fonction qui associe chaque événement à une valeur entre 0 et 1. Voir diapo 15, definition importante.

Définition 1.3.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Densité de proba

Retrouver la définition.

Fonction de répartition

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(x)dx.$$

Espérance :

$$E(X) = \sum x_k * p_k$$

$$E(X) = \int Xp(x)dx$$

$$E(aX + b) = aE(X) + b$$
$$E(X + Y) = E(X) + E(Y)$$

Le Mode

$$p(Mo) = \max_k p(x_k)p(Mo) = \max_x p(x)$$

Variance :

$$\sigma^2 = \sum (x_k - E(X))^2$$

$$\sigma^2 = \int (x - E(X))^2 p(x) dx$$

$$V(aX + b) = a^2 V(X)$$

$$V(X) = E(X^2) - E(X)$$

Médiane et quantile

idk diapo

Définition 1.4 (Loi marginale)**.** La marginalisation consiste à projeter une loi jointe sur l'une des variables aléatoires. Par exemple extraire $P(A)$ à partir de $P(A, B)$.

$$P(A) = \sum_i P(A, B = pb_i).$$

C'est la somme de la ligne ou de la colonne du tableau.

Définition 1.5. Probabilités conditionnelles

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

$$\Leftrightarrow P(A \cup B) = P(A|B)P(B)$$

Proposition 1.1.
— *Réversibilité :* $P(A, B) = P(A|B)$
— *Théorème de Bayes :*

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

— *Intégration des probabilités totale*
— *DIAPO 39*

Définition 1.6 (Indépendance probabiliste)**.** Deux événements A et B sont indépendants si

$$P(A, B) = P(A) * P(B).$$

Corollaire : $P(A|B) = P(A)$

Définition 1.7. La covariance

$$cov(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Définition 1.8 (Coefficient de corrélation linéaire)**.** Soit X,Y deux variables. Le coefficient de corrélation linéaire entre X et Y est :

$$r = \frac{cov(X, Y)}{\sigma_X \sigma_Y}.$$

CCL :
VOIR DIAPOProbabilité Marginalisation Conditionnement Indépendance : Si X_1 et X_2 sont indépendantes : $P(X_1, X_2) = P(X_1)P(X_)$

Nouveau cours du 20/09

Définition 1.9 (Indépendance de deux variables discrète)**.** Discrète : Continue :

Définition 1.10 (Indépendance mutuelle de n variable)**.** Soient n variables aléatoires $(X_1, ..., X_n)$. Elle sont **mutuellement indépendantes** si tout événement lié à une partie d'entre elles est indépendant de tout événement lié à toute autre partie disjointe de la précédente. Propriété :
— Indépendance mutuelle $\rightarrow$ Indépendance deux à deux. **Attention :** réciproque fausse
$\rightarrow$ Permet de réduire la taille du tableau des probabilité de chaque événement !

Définition 1.11 (Indépendance conditionnelles)**.** On reprend les formules de l'indépendances mais en sachant une variable, au final c'est dans un cas particulier.

$$X \perp Y | Z$$
$$\forall x, \forall y, \forall z P(X = x \cap Y = y | Z = z) = P(X = x | Z = z) * P(Y = y | Z = z)$$

$\rightarrow$

$$\Rightarrow P(X, Y | Z) = P(X | Z) * P(Y | Z).$$

Définition 1.12. Loi normale

Proposition 1.2. *- Moyenne linéaire et variance comme bilinéaire*

$$X \sim \mathcal{N}(\mu, \sigma^2) \text{ alors } Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2).$$

- Centrer et réduire

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Définition 1.13 (Convergence en loi)**.**

$$\forall x, \lim_{n \to \infty} F_n(x) = F(x).$$

On ne sais pas comment ça converge

Définition 1.14 (Convergence en probabilité)**.** (X_n) **converge en probabilité** vers X si, pour tout $\epsilon > 0$ la probabilité que l'écart absolu entre X_n et X dépasse ϵ tend vers 0 quand $n \to \infty$

$$\lim_{n \to \infty} P(|X_n - X| \geq \epsilon) = 0.$$

Définition 1.15 (convergence presque sur)**.** (X_n) **converge presque surement**vers X s'il y a yne proba 1 que la suite des réalisation des X_n tende vers X

Définition 1.16 (Loi faible des grands nombre)**.** Soit $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires :
- De même loi
- D'éspérance m
- Possédant une variance σ^2
- **Deux à deux** indépendante

Alors

$$\bar{X_n} = \frac{\sum_{k=1}^{n} X_k}{n} \to_{\mathbb{P}} m.$$

Rappel :

$$E(\bar{X_n}) = m$$
$$V(\bar{X_n}) = \frac{\sigma^2}{n}$$

Définition 1.17 (Loi forte des grands nombres)**.** Soit $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires :
- De même loi
- D'éspérance m
- Possédant une variance σ^2
- **mutuellement** indépendante

Alors

$$\bar{X_n} = \frac{\sum_{k=1}^{n} X_k}{n} \to_{p.s} m.$$

Définition 1.18 (Théorème centrale limite)**.** Soit $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires :
- De même loi
- D'éspérance μ
- Possédant une variance σ^2
- **mutuellement** indépendantes

Alors

$$\frac{\bar{X_n} - \mu}{\sigma / \sqrt{n}} \to_{loi} \mathcal{N}(0, 1).$$

Nouveau cours du 27/09

2 Maximum Vraissemblance

Définition 2.1 (Vraissemblance d'un échantillon)**.** Soit $x = (x_1, \ldots, x_n)$ réalisation de $(X_1, \ldots, X_n)$ **iid = Mutuellement indépendant** Alors on défini la vraisemblance dans le cas discret comem étant la proba d'obtenir **cet** échantillon sachant la loi P

$$L(x) = P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i).$$

Dans le cas continue :

Exemple 2.1 (avec des pièces de monnaies)**.** DIAPO 5

Exemple 2.2 (innondation)**.** 3 type de parcelles : Inondables (PI), partiellement inondables (PPI), non inondable (NI) On a deux loi caractérisant le niveau de gris par rapport à la catégorie d'inondation.

$$P(n|PI) = \mathcal{N}(\mu_1, \sigma_1^2), P(n|PPI) = \mathcal{N}(\mu_2, \sigma_2^2).$$

Avec n le niveau de gris.

Soit une image Z avec un niveau de gris $n = 80$; Deux hypothèses

- θ_1 Z PI
- θ_2 Z PPI

On vas calculer le max de vraisemblance d'obtenir la zone Z sous $\theta 1$ ou $\theta 2$

2.1 Maximum de vraissemblance

Exemple 2.3 (Pièce de monnaie)**.** On vas faire la même chose mais cette fois ci, on prend des paramètre $\theta_1 et \theta_2$

Définition 2.2 (Vraissemblance d'un échantillon)**.** On cherche à estimer un paramètre Θ Soit $x = (x_1, \ldots, x_n)$ réalisation de $(X_1, \ldots, X_n)$ **iid = Mutuellement indépendant** Alors on défini la vraisemblance dans le cas discret comme étant la proba d'obtenir **cet** échantillon sachant la loi P

$$L(x) = P(X_1 = x_1, \ldots, X_n = x_n | \Theta = \theta) = \prod_{i=1}^{n} P(X_i = x_i | \Theta = \theta).$$

On peut utiliser la fonction de densité.

Définition 2.3 (Maximum vraissemblance)**.** On cherche le maximum de la fonction $L(x, \theta), \forall \theta$. Donc on vas la dériver ! et utiliser le $\log$

Exemple 2.4. Plein d'exemple dans le diapo

Exemple 2.5 (problème d'ajustement)**.** On a des points un peu random, réparti comme un sinusoide qu'on vas approximé par un polynome. Problème : on a une erreur Normale.

il vas vite. Mais ça resemble à une regression.

2.2 Estimation par maximum a posteriori

Exemple 2.6 (Pièce de monnaie)**.** Imaginons qu'on a un tirage de 3 piles. Le maximum de vraisemblance vaut 1. Mais ça va à l'encontre du bon sens. $\rightarrow$ Solution : Maximum a posteriori. (A voir pourquoi ça fait 1)

Définition 2.4 (Maximum a posteriori)**.** On se base dans un modèle bayésien avec

- $\mathcal{X} =$ l'espace des observations x de taille n
- Θ

Formile de la vraissemblance diapo 32.

Estimateur du maximum a posteriori toujours égal à l'argmax de la vraissemblance

$$x \mapsto t = Argmax_{\theta \in \Theta} \pi(\theta | x).$$

Exemple 2.7 (pièce de monnaie)**.** En faite la grosse différence c'est la dernière ligne du diapo 34. On pose, on invente l'information a priori de la proba de chaque paramètre qu'on vas tester. Cette information vas permettre d'être utiliser dans le modèle bayésien. On l'a choisi en fonction d'une loi normale. Fin du cours sans qu'il ait fini.