

Notes TC4

Adrien Pavao

September 2017

Contents

1 Définitions et formules	1
1.1 Notions générales	1
1.2 Différentes probabilités liées	2
1.3 Moyenne, variance, covariance et corrélation	2
1.4 Indépendance statistique	3
1.5 Théorème de Bayes	4
1.6 Lois de probabilités courantes	4
2 Classification bayésienne	4

1 Définitions et formules

1.1 Notions générales

- **Variable aléatoire :** Une fonction définie depuis l'ensemble des résultats possibles d'une expérience aléatoire, dont on doit pouvoir déterminer la probabilité qu'elle prenne une valeur donnée ou un ensemble donné de valeurs.

Cette variable peut être discrète ou continue.

Dans le cas d'une variable discrète, la fonction masse est la fonction qui donne la probabilité d'un résultat élémentaire d'une expérience.

Dans le cas d'une variable continue, la distribution de la masse de probabilité est caractérisée par la densité de probabilité $f(x)$:

$$P(a < X \leq b) = \int_a^b f(x) dx$$

- **Réalisations :** Les réalisations d'une variable aléatoire sont les résultats des valeurs choisies au hasard en fonction de la loi de probabilité de la variable. On les appelle également les variations aléatoires.

- **Distribution (loi de probabilité) :** Le concept de loi de probabilité se formalise mathématiquement à l'aide de la théorie de la mesure : une loi de probabilité est une mesure, souvent vue comme la loi décrivant le comportement d'une variable aléatoire, discrète ou continue. Une mesure est une loi de probabilité si sa masse totale vaut 1. L'étude d'une variable aléatoire suivant une loi de probabilité discrète fait apparaître des calculs de sommes et de séries, alors que si sa loi est absolument continue, l'étude de la variable aléatoire fait apparaître des calculs d'intégrales.
- **Inférence :** Trouver la valeur des v.a. à partir d'autres qui sont connues.
- **Vraisemblance :** $P(D; \Theta)$ avec D les données d'estimation et Θ l'ensemble des paramètres.
- **Estimation :** Retrouver les paramètres d'une distribution à partir de l'observation d'un ensemble de réalisations de celle-ci. Il s'agit du maximum de vraisemblance.
- **Sampling :** Générer des données.

1.2 Différentes probabilités liées

- **Probabilité jointe :** Probabilité d'une configuration donnée. On note $P(X, Y)$. Il s'agit de la probabilité de la réalisation de l'événement X et de l'événement Y.
- **Probabilité conditionnelle :** La probabilité de X sachant Y se note $P(X|Y)$. Pour faire simple, on fixe la variable connue.

$$P(X|Y) = \frac{P(X,Y)}{P(Y)}$$
- **Probabilité marginale :** On vire une variable en la sommant. A partir de $P(X = x, Y = y)$, on peut obtenir :

$$P(X = x) = \sum_y P(X = x, Y = y)$$

La probabilité jointe inclut les deux autres. On peut retrouver la distribution jointe à partir de la distribution conditionnelle et marginale.

1.3 Moyenne, variance, covariance et corrélation

Soit x_1, x_2, \dots, x_n un ensemble de valeurs générées par une distribution de probabilité inconnue. On peut caractériser cette distribution par :

- La **moyenne**, qui caractérise le **centre** de la distribution.
- La **variance**, qui mesure la **dispersion** de la distribution.

$$- \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

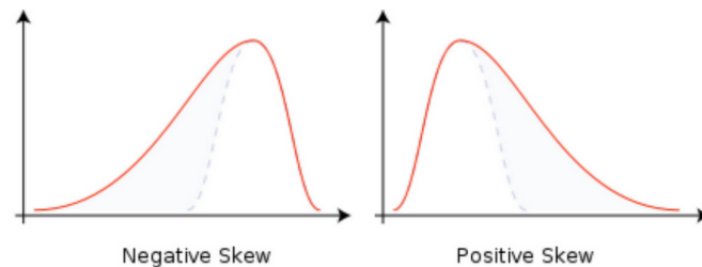
$$- \text{var}(x) = \frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})^2$$

- L'**écart-type**, qui est la racine carrée de la variance.

$$\sigma = \sqrt{\text{var}(x)}$$

- La symétrie ou l'asymétrie (skewness).

Figure 1: Schéma représentatif de la skewness



- La covariance, le lien entre les variations de deux variables.

$$\text{cov}(x, y) = \frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})(y_n - \bar{y})$$

- La corrélation, une covariance normalisée. Elle quantifie la qualité de l'approximation linéaire de x par y, et réciproquement.

$$\text{cor}(x, y) = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}} = \frac{\text{cov}(x, y)}{\sqrt{\text{cov}(x, x)\text{cov}(y, y)}}$$

Pour les V.A. multidimensionnelles, on a la **matrice de covariance**.

1.4 Indépendance statistique

- **Variables indépendantes** : Deux variables sont indépendantes si et seulement si :

$$P(X = x, Y = y) = P(X = x) \times P(Y = y)$$

- **Variables conditionnellement indépendantes** : Deux variables sont conditionnellement indépendantes si et seulement si :

$$P(X = x, Y = y | Z = z) = P(X = x | Z = z) \times P(Y = y | Z = z)$$

- **Indépendantes et identiquement distribuées (i.i.d)** : Ce dit de deux variables indépendantes suivant la même loi de probabilité. Il sera souvent nécessaire de poser cette hypothèse. On en déduit : $P(X_{1,N}) = \prod_{i=1}^N P(X_i)$

1.5 Théorème de Bayes

- $P(Y = y_j | X = x_i) = \frac{P(X=x_i|Y=y_j)P(Y=y_j)}{\sum_{y_j \in A_y} P(X=x_i|Y=y_j)P(Y=y_j)}$
- $P(Y = y_j | X = x_j) = \frac{P(X=x_i|Y=y_j)P(Y=y_j)}{P(X=x_i)}$

1.6 Lois de probabilités courantes

Loi	Type de la V.A.	Formule	Paramètres
Bernouilli	Binaire	$p^x(1-p)^{1-x}$	p
Binomiale	Discrete	$C_n^k \times p^k \times (1-p)^{n-k}$	n, p
Poisson	Discrete	$\frac{\lambda^k}{k!} \times e^{-\lambda}$	λ
Normale dimension 1	Continue	$\frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$	μ, σ
Normale dimension d	Continue	$\frac{1}{(2\pi)^{\frac{d}{2}} \Sigma ^{\frac{1}{2}}} \times e^{-\frac{1}{2}(x-\mu)^t \Sigma^{-1}(x-\mu)}$	μ, Σ

2 Classification bayésienne

Note : A completer.

Soit les points suivantes:

- Variable aléatoire X, une observation à classer.
- Variable aléatoire Y, désignant la classe à affecter à X.
- Décision α_i , affectation de x à la classe $y = i$.
- Fonction de perte $\lambda(\alpha_i|j)$, décision α_i alors que la classe correcte était j.

Sur l'observation x, l'espérance du risque de décider α_i est :

$$R(\alpha_i|x) = \sum_{j=1}^k \lambda(\alpha_i|j)P(Y = j|x)$$