Ecole Polytechnique | **Statistical Learning**
Master M2 Data Science : 2026/2027 | **Stephan Clémençon - Quentin Bouniot**

TRAVAUX DIRIGÉS N° 1 : Bayes Classifier

Stephan CLÉMENÇON <stephan.clemencon@telecom-paris.fr>
Quentin BOUNIOT <quentin.bouniot@telecom-paris.fr>

We work in the framework of binary classification. Throughout this tutorial, we consider a random descriptor X taking values in a measurable space $\mathcal{X} \subset \mathbb{R}^d$ ($d \in \mathbb{N}^*$) and a random label Y taking the value 0 or 1. The joint distribution of the vector (X, Y) is denoted by P, and the regression function (posterior probability) is

$$\eta : x \in \mathcal{X} \mapsto \mathbb{P}(Y = 1 \mid X = x) \in [0, 1].$$

EXERCICE 1. We consider $\mathcal{X} = [0, 1]$ and P such that :

- the conditional distribution of X given $Y = 0$ is $P_0 = \mathcal{U}([0, \theta])$, where $\theta \in]0, 1[$,
- the conditional distribution of X given $Y = 1$ is $P_1 = \mathcal{U}([0, 1])$,
- $p = \mathbb{P}(Y = 1) \in]0, 1[$.

For $x \in \mathcal{X}$, give $\eta(x)$ as a function of p and θ. Give the numerical expression for $\theta = 1/2$.

EXERCICE 2. We consider $\mathcal{X} = \mathbb{R}_+$ and P such that :

- the distribution of X is P_X on $\mathcal{X}$,
- the regression function is $\eta(x) = \dfrac{x}{x + \theta}$ for every $x \in \mathcal{X}$, where $\theta > 0$ is fixed.

Let h^* denote the Bayes classifier. Make the Bayes classifier explicit in this model. Then show that its 0-1 risk (its probability of error) is

$$L(h^*) = \int_{\mathcal{X}} \min\Big(\eta(x), 1 - \eta(x)\Big)\, dP_X(x).$$

Compute the Bayes risk when $P_X = \mathcal{U}([0, \alpha\theta])$ with $\alpha > 1$.

EXERCICE 3. Let $\omega(0), \omega(1) \geq 0$ be weights such that $\omega(0) + \omega(1) = 1$. We consider the weighted classification risk :

$$L_\omega(g) = \mathbb{E}\left(2\omega(Y) \cdot \mathbb{1}_{\{Y \neq g(X)\}}\right), \qquad g : \mathcal{X} \to \{0, 1\}.$$

Give the Bayes classifier and the Bayes risk for this criterion. What is the interest of considering such a criterion ?

EXERCICE 4. We consider $X = (T, U, V)$, where T, U, V are i.i.d. real-valued random variables with the standard exponential distribution. Set $Y = \mathbb{1}_{\{T+U+V<\theta\}}$, where $\theta \in \mathbb{R}_+$ is fixed. Compute the Bayes classifier $g^*(T, U)$ when V is unobserved, and then compute its 0-1 risk as a function of θ.

Bibliographic advice

Useful starting points for machine learning are listed below :

- Theoretical and focused on probabilistic aspects : [DGL97]
- Practical and utility-oriented : [HTF13]
- A recent book focused mainly on optimization : [SSBD14] (and, by the same author, on online learning : [SS12])
- Bayesian methods and graphical models : [MB12]

Références

[DGL97] L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition.* Stochastic Modelling and Applied Probability. Springer New York, 1997. 2

[HTF13] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning : Data Mining, Inference, and Prediction.* Springer Series in Statistics. Springer New York, 2013. 2

[MB12] K.P. Murphy and F. Bach. *Machine Learning : A Probabilistic Perspective.* Adaptive Computation and Machine Learning Series. MIT Press, 2012. 2

[SS12] S. Shalev-Shwartz. *Online Learning and Online Convex Optimization.* Foundations and Trends in Machine Learning Series. Now Publishers, 2012. 2

[SSBD14] S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning : From Theory to Algorithms.* Cambridge University Press, 2014. 2