

CREST - GENES
Cours de Formation par la Recherche
2010 – 2011

Uniform Probability Inequalities with Application in High-dimensional Statistical Models

Sara van de GEER

(Seminar for Statistics, ETH, Zurich, et Invitée CREST, LS)

We review some uniform probability and moment inequalities for sums of independent random variables indexed by a parameter. Concentration inequalities for three types of situations will be discussed: the case of sub-exponential, of sub-Gaussian, and of bounded random variables.

As an elaborated example, we study the case where the parameter space is the convex hull of a set with small covering number, extending some entropy results of Ball and Pajor (1990).

The results will be applied in high-dimensional statistical problems, where l_1 regularization is employed. One of our focus points here will be dealing with situations with highly correlated design. We moreover discuss high-dimensional nonlinear models.

Most of the results will be presented with a full proof. The course aims at providing insight into the main ideas and a background for further development of the theory for high-dimensional statistical models.

Bibliographie

- o Ball, K. et A. Pajor (1990), "The Entropy of Convex Bodies with Few Extreme Points", *London Math. Soc., Lecture Note Ser.*, 158, 25-32.
- o Bousquet, O. (2002), "A Bennett Concentration Inequality and its Application to Suprema of Empirical Processes", *C.R. Acad. Sci. Paris*, 334, 495-550.
- o Dümbgen, L., van de Geer, S.A., Veraar, M.C. et J.A. Wellner (2010), "Nemirovskis Inequalities Revisited", *Amer. Math. Monthly*, 117, 138-160.
- o Massart, P. (2000), "About the Constants in Talagrand's Concentration Inequalities for Empirical Processes", *Ann. Probab.*, 28, 863-884.
- o Talagrand, M. (1995), "Concentration of Measure and Isoperimetric Inequalities in Product Spaces", *Publications Mathématiques de IHES*, 81, 73-205.
- o Talagrand, M. (2005), *The Generic Chaining : Upper and Lower Bounds of Stochastic Processes*, Springer Verlag.
- o van de Geer, S. (2008), "High-dimensional Generalized Linear Models and the Lasso", *Annals of Statistics*, 36, 614-645

Cours	Lundi	28	Février 2011	De 14h à 16h 10	Amphi 2
Les :	Jeudi	3	Mars 2011	De 14h à 17h 15	Amphi 2
	Lundi	7	Mars 2011	De 14h à 16h 10	Amphi 2
	Jeudi	10	Mars 2011	De 14h à 17h 15	Amphi 2

à l'ENSAE 3, Avenue Pierre Larousse, Malakoff (Métro : Malakoff/Plateau de Vanves)

Ces cours sont proposés aux étudiants de 3^{ème} année de l'ENSAE, de l'ENSAI se préparant à la recherche et ouverts aux étudiants de M2 ou inscrits en thèse. **Une inscription préalable est demandée impérativement** pour tous les étudiants de l'ENSAE, de l'ENSAI, ou extérieurs, par courriel à guedj@ensae.fr ou par tél. au **01 41 17 35 50, afin de pouvoir être admis dans les locaux de l'ENSAE**. Les renseignements supplémentaires sur le contenu et les dates de ces cours peuvent être obtenus au 01 41 17 35 50.