

PROCÉDURES ADAPTATIVES DE TESTS NON PARAMÉTRIQUES

Béatrice LAURENT

*INSA de Toulouse, Institut de Mathématiques de Toulouse,
beatrice.laurent@insa-toulouse.fr*

Cet exposé sera une synthèse de plusieurs travaux écrits avec divers co-auteurs et relatifs à la construction de procédures de tests dans un cadre statistique non paramétrique. Nous présenterons les exemples des tests d'adéquation dans un modèle de densité, des tests d'homogénéité pour les processus de Poisson avant de considérer un problème de tests à deux échantillons. La première étape d'une telle construction repose sur la définition d'une statistique de test T dont le comportement diffère sous l'hypothèse nulle et sous l'alternative. Nous nous attacherons principalement à des statistiques de tests basées sur l'estimation du carré d'une distance \mathbb{L}^2 entre la loi des observations et cette loi sous l'hypothèse nulle. Les procédures de tests proposées sont non asymptotiques : le niveau est garanti quelque soit la taille de l'échantillon considéré. Plusieurs cas peuvent se présenter: le cas le plus simple est celui où la loi de la statistique de test T sous l'hypothèse nulle H_0 est connue; on peut alors déterminer les quantiles de cette loi et définir la règle de décision du test afin de garantir le niveau. Lorsque l'hypothèse nulle est composite, la loi de T sous H_0 n'est en général pas connue. On peut alors avoir recours à des tests conditionnels s'il est possible de déterminer une variable aléatoire Z , fonction des observations, telle que la loi de T conditionnellement à Z sous H_0 est connue. Pour les problèmes de tests à deux échantillons en densité, en régression ou pour des processus de Poisson, nous avons recours à des procédures à noyaux qui s'appuient sur des méthodes de bootstrap et de permutation. Pour ce qui concerne la puissance des tests proposés, nous verrons l'intérêt d'agréger diverses procédures de tests, ce qui permet de s'adapter à la structure inconnue de l'alternative. Nous obtenons ainsi des vitesses de séparation optimales sur diverses classes d'alternatives simultanément.

Mots-clés. Tests adaptatifs, Méthodes d'agrégation, Méthodes à noyaux, Vitesses de séparation.

Abstract. We will present a synthesis of several papers written with co-authors and related to the construction of adaptive testing procedures in non parametric statistics. We consider the examples of goodness-of-fit tests in a density model, homogeneity test for Poisson processes and the two sample testing problem. The first step in such a construction is to define a test statistic T which behavior under the null hypothesis and under the alternative differs. We consider test statistics based on the estimation of the \mathbb{L}^2 distance of the distribution of the observations under the null hypothesis and under

the alternative. The testing procedures that we propose are non asymptotic : the level is guarantied for all sample size. Several cases can occur : the most simple case is the case where the distribution of the test statistic T under the null hypothesis is known, one can then determine the quantiles of this distribution and define the testing procedure. When the null hypothesis is composite, this distribution is generally unknown. One can then consider conditional tests, if it is possible to determine a random variable Z such that the distribution of T conditionally on Z under the null is known. For the two sample testing problem in density, in regression or for Poisson processes, we use kernel procedures combined with bootstrap and permutation methods. Concerning the power of the testing procedures, we will see that the aggregation of several tests allows to adapt to the unknown structure of the alternative.

Keywords. Adaptive tests, Agregation methods, Kernel methods, Separation rates.

Bibliographie

- [1] M. Fromont et B. Laurent (2006), *Adaptive goodness-of-fit tests in a density model*, The Annals of Statistics, 2006, Vol. 34, No 2. p.1-45.
- [2] M. Fromont, B. Laurent., P. Reynaud-Bouret, (2011), *Adaptive tests of homogeneity for a Poisson process*, Annales de l'Institut Henri Poincaré (B), Vol. 47, No. 1, 176-213.
- [3] M. Fromont, B. Laurent., M. Lerasle, P. Reynaud-Bouret (2012), *Kernels Based Tests with Non-asymptotic Bootstrap Approaches for Two-sample Problems*, COLT 2012, JMLR W &CP 23: 1-23.
- [4] M. Fromont, B. Laurent., P. Reynaud-Bouret (2013), *The two-sample problem for Poisson processes: adaptive tests with a non-asymptotic wild bootstrap approach*, The Annals of Statistics, Vol. 41, No. 3, 1431-1461.
- [5] B. Laurent., C. Marteau, C. Maugis-Rabusseau (2016), *Non asymptotic detection of two component mixtures with unknown means*, Bernoulli 22(1), 2016, 242-274.