Sharp Estimates on the Tail Behaviour of Some Random Integrals and Their Application in Statistics

Péter Major

Mathematical Institute of the Hungarian Academy of Sciences

Abstract

I met the problems discussed here in an investigation when I tried to adapt the methods of maximum likelihood estimates to some non-parametric problems. One can give a good and simple approximation for the error of the maximum likelihood estimate by means of an appropriate linearization in the maximum-likelihood equation. It can be shown that a really good approximation is obtained in such a way with the help of the observation that the coefficient of the second term in the Taylor expansion we apply in this linearization is bounded. This linearization argument can be adapted to the study of several interesting non-parametric problems, but in the non-parametric case some multiple random integrals have to be bounded instead of Taylor coefficients. This led to the following problem described below.

Let us have a sequence of iid. random variables $\xi_1, \dots, \xi_n$ on a space $(X, \mathcal{X})$ with distribution μ, and let μ_n denote their empirical distribution. Given a real valued function $f(x_1, \dots, x_k)$ of k variables on the space $(X, \mathcal{X})$ consider the k-fold random integral

$$J_{n,k}(f) = \frac{n^{k/2}}{k!} \int' f(u_1, \dots, u_k)(\mu_n(\,du_1) - \mu(du_1)) \dots (\mu_n(\,du_k) - \mu(\,du_k)),$$

where prime means that the diagonals are omitted from the domain of integration. We want to give a good bound on the probability $P(|J_{n,k}(f)| > x)$ for all $x > 0$. More generally, given a nice class of functions $f \in \mathcal{F}$ of k variables we are interested in the probability $P\left(\sup_{f \in \mathcal{F}} |J_{n,k}(f)| > x\right)$ for $x > 0$.

The tail-behaviour of $J_{n,k}(f)$ is similar to that of the k-th power of a Gaussian random variable with expectation zero and variance of the same

order as the k-th root of the variance of the random variable $J_{n,k}(f)$ provided that this variance is not too small. More explicitly, $P(|J_{n,k}(f)| > x) < Ce^{-B(x/\sigma)^{2/k}}$ with some universal constants $C > 0$ and $B > 0$ and $\sigma^2 = \int f^2(u_1, \dots, u_k)\mu(\,du_1)\dots\mu(\,du_k)$ if the absolute value of the function f is bounded by 1, and $0 < x < n^{k/2}\sigma^{k+1}$. Beside this, the variance of $J_{n,k}(f)$ has the same order as σ^2. The same estimate holds with possibly different universal constants $C > 0$ and $B > 0$ for the probability $P\left(\sup\limits_{f \in \mathcal{F}} |J_{n,k}(f)| > x\right)$ if $\mathcal{F}$ is a nice class of functions of k variables whose elements are such functions which are bounded in supremum norm by 1 and in L_2 norm by σ. Such a result holds for instance if $\mathcal{F}$ is a so-called Vapnik–Červonenkis class of functions. The only additional restriction we have to impose for the validity of such an estimate is the condition $\left(\frac{x}{\sigma}\right)^{2/k} > D \log n$ with some $D > 0$ to exclude the possibility that the supremum of relatively small random variables be large.

References:

Major, P. and Rejtő, L. (1998) A note on nonparametric estimations. In: The Conference Volume to the 65 Birthday of Miklós Csörgő. 759–774

Major, P. (2003) An estimate about multiple stochastic integrals with respect to a normalized empirical measure. Submitted to *Probability Theory and Related Fields*

Major, P. (2003) An estimate on the maximum of a nice class of stochastic integrals. Submitted to *Probability Theory and Related Fields*