MODERATE DEVIATIONS FOR DEGENERATE U-PROCESSES

PETER EICHELSBACHER

ABSTRACT. Sufficient conditions for a rank-dependent moderate deviations principle (MDP) for degenerate U-processes are presented. The MDP for VC classes of functions is obtained under exponential moments of the envelope. Among other techniques, randomization, decoupling inequalities and integrability of Gaussian and Rademacher chaos are used to present new Bernstein-type inequalities for U-processes which are the basis of our proofs of the MDP. We present a complete rank-dependent picture. The advantage of our approach is that we obtain in the degenerate case moderate deviations in non-Gaussian situations.

1. INTRODUCTION

Let $(S, \mathcal{S}, \mu)$ be a probability space and let $X_i : S^{\mathbb{N}} \to S$ be the coordinate functions ($\{X_i\}_{i \in \mathbb{N}}$ is thus an i.i.d. sequence with $\mathcal{L}(X_i) = \mu$). The *U-empirical measure of order m* is defined by

$$L_n^m := \frac{1}{n_{(m)}} \sum_{(i_1,\ldots,i_m) \in I_{m,n}} \delta_{(X_{i_1},\ldots,X_{i_m})},$$

where $n_{(m)} := \prod_{k=0}^{m-1}(n-k)$ and $I_{m,n} \subset \{1,\ldots,n\}^m$ contains all m-tuples with pairwise different components and δ_x denotes the probability measure degenerate at $x \in S$. Let $\mathcal{H}$ be a collection of measurable functions $h\colon S^m \to \mathbb{R}$. The *$U$-process* of order m indexed by $\mathcal{H}$ is defined for every integer $n \geq m$ as

$$U_n^m(h,\mu) := \int_{S^m} h \, dL_n^m = \frac{1}{n_{(m)}} \sum_{(i_1,\ldots,i_m) \in I_{m,n}} h(X_{i_1},\ldots,X_{i_m}), \quad h \in \mathcal{H}.$$

For a fixed h, the expression $U_n^m(h,\mu)$ is called a *U-statistic of order m* with kernel h based on the probability measure μ. U-processes appear in statistics frequently as unbiased estimators of the functional $\{\mu^{\otimes m} h\colon h \in \mathcal{H}\}$. For instance, Liu's simplicial depth process (Liu 1990) is a U-process. Noland and Pollard (1987), (1988) studied the law of large numbers and the central limit theorem for U-processes of order $m = 2$. Arcones and Giné (1993) developed the theory for an arbitrary $m > 1$. Regarding the law of large numbers they presented a necessary and sufficient condition for its validity in complete analogy with the results for empirical processes. The central limit theorem (CTL) and the law of iterated logarithm (LIL) are developed only for Vapnic-Červonenkis (VC) classes of sets and functions, because such conditions are unknown at present.

Wu (1994) proved necessary and sufficient conditions for the large deviation and moderate deviation estimations and the LIL of the empirical process $L_n(h) =$

1991 *Mathematics Subject Classification.* 60F10 (primary), 60B12, 60G17.

Key words and phrases. Rank-dependent moderate deviations, U-statistics, U-processes, VC classes, Bernstein-type inequality, metric entropy.

$1/n \sum_{i=1}^n h(X_i)$ with h varying in a class of uniformly bounded functions. The principles are proved for laws in the Banach space of bounded functionals on the class $\mathcal{H}$. Sufficient conditions for the large deviation principle (LDP) as well as for the moderate deviations principle (MDP) in the so-called non-degenerate case are proved in Eichelsbacher (1998). Precise definitions of LDP and MDP are given below.
The case of completely degenerate or canonical kernels is the crucial one for the MDP and in this paper the goal is to study the MDP for degenerate U-processes under certain conditions on $\mathcal{H}$. We develop the principle for VC classes. We are able to present a complete MDP ,,picture" for non-Gaussian cases.
Next we present some notation. Let us recall the definition of the MDP. A sequence of probability measures $\{\mu_n\}_{n\in\mathbb{N}}$ on a topological space $\mathcal{X}$ equipped with σ-field $\mathcal{B}$ is said to satisfy the LDP with speed $a_n \downarrow 0$ and good rate function $I(\cdot)$ if the level sets $\{x : I(x) \leq \alpha\}$ are compact for all $\alpha < \infty$ and for all $\Gamma \in \mathcal{B}$ the lower bound

$$\liminf_{n\to\infty} a_n \log \mu_n(\Gamma) \geq - \inf_{x\in \mathrm{int}(\Gamma)} I(x),$$

and the upper bound

$$\limsup_{n\to\infty} a_n \log \mu_n(\Gamma) \leq - \inf_{x\in \mathrm{cl}(\Gamma)} I(x)$$

hold, where $\mathrm{int}(\Gamma)$ and $\mathrm{cl}(\Gamma)$ denote the interior and closure of Γ, respectively. We say that a sequence of random variables satisfies the LDP provided the sequence of measures induced by these variables satisfies the LDP. Let $\{b_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ be a sequence which satisfies

$$\lim_{n\to\infty} \frac{b_n}{n} = 0 \qquad \text{and} \qquad \lim_{n\to\infty} \frac{n}{b_n^2} = 0. \tag{1.1}$$

If $\mathcal{X}$ is a topological vector space, then a sequence of random variables $\{Z_n\}_{n\in\mathbb{N}}$ satisfies the MDP with speed $\frac{n}{b_n^2}$ and with good rate function $I(\cdot)$ in case the sequence $\big\{\Phi\big(\frac{n}{b_n}\big)Z_n\big\}_{n\in\mathbb{N}}$ satisfies the LDP in $\mathcal{X}$ with the good rate function $I(\cdot)$ and with speed $\frac{n}{b_n^2}$. Here $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a convex function. In our applications $\Phi(x) = x^r$ for some $r \in \mathbb{N}$.
Since we are using *Hoeffding's decomposition* of a U-statistic we state it here together with some notation. The operator $\pi_{k,m}^{\mu} = \pi_{k,m}$ (sometimes called Hoeffding projections) acts on $\mu^{\otimes m}$-integrable symmetric functions $h : S^m \to \mathbb{R}$ (symmetric in the sense that for all $x_1, \dots, x_m \in S$ and all permutations s of $\{1, \dots, m\}$ $h(x_1, \dots, x_m) = h(x_{s_1}, \dots, x_{s_m})$) as follows:

$$\pi_{k,m} h(x_1, \dots, x_k) = (\delta_{x_1} - \mu) \otimes \cdots \otimes (\delta_{x_k} - \mu) \otimes \mu^{\otimes(m-k)} h,$$

where $\nu_1 \otimes \cdots \otimes \nu_m h = \int \cdots \int h(x_1, \dots, x_m) d\nu_1(x_1) \cdots d\nu_m(x_m)$ and

$$\mu^{\otimes(m-1)} h(x) = \int \cdots \int h(x_1, \dots, x_{m-1}, x) d\mu(x_1) \cdots d\mu(x_{m-1}),$$

for $0 \leq k \leq m$. Note that $\pi_{0,m} h = \mu^{\otimes m} h$. A symmetric function h is called μ-*canonical* or *completely degenerate* if $\int h(x_1, \dots, x_m) d\mu(x_1) = 0$ for all $x_2, \dots, x_m \in S$. Note that $\pi_{k,m} h$ is a μ-canonical function of k variables. A $\mu^{\otimes m}$-integrable symmetric function $h : S^m \to \mathbb{R}$ is μ-degenerate of order $r-1$, $1 \leq r \leq m$, if

$$\int h(x_1, \dots, x_m)\, d\mu^{\otimes m-r+1}(x_r, \dots, x_m) = \int h\, d\mu^{\otimes m}$$