Some Aspects in Large Deviations

Peter Eichelsbacher
Sonderforschungsbereich 343
Diskrete Strukturen der Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld, Germany

Abstract

Let V be a topological space and $\mathcal{B}(V)$ the Borel σ–field on V. A sequence of probability measures $(P_n)_n$ satisfies a large deviation principle (LDP) with rate function $I(\cdot)$ if $P_n(A)$ can be approximated by $\exp(-n \inf_{x \in A} I(x))$ for appropriate events A in $\mathcal{B}(V)$. In the proof of the LDP for m–variate von Mises–statistics and U–statistics combinatorial factorization arguments can be dropped. In general, the governing rate function is not convex. Nevertheless, we are able to give a representation analogously to the convex conjugate concept. Moreover, we analyse a convex minorant of the rate function.

Short running title: Large deviations

1 Introduction

Let V be a topological space and $\mathcal{B}(V)$ the Borel σ–field on V. A sequence of probability measures $(P_n)_n$ satisfies a large deviation principle (LDP) with rate function $I(\cdot)$ if $P_n(A)$ can be approximated by $\exp(-n \inf_{x\in A} I(x))$ for appropriate events A in $\mathcal{B}(V)$. The following definition makes this more precise:

A sequence of probability measures $(P_n)_n$ on $(V, \mathcal{B}(V))$ obeys a large deviation principle (LDP) with speed a_n and rate function I if there exists a sequence $(a_n)_n$ of positive numbers tending to infinity and a function $I : V \to [0, \infty]$ such that the following conditions hold:

1. $I(x)$ is a lower semicontinuous function. $I \not\equiv 0$.
2. $\{x \;:\; I(x) \le C\} \subseteq V$ is compact for all $C < +\infty$ (compact level sets).
3. For every open set $O \subseteq V$
$$\liminf_{n\to\infty} \frac{1}{a_n} \log P_n(O) \ge -\inf_{x\in O} I(x). \tag{1}$$
4. For every closed set $A \subseteq V$
$$\limsup_{n\to\infty} \frac{1}{a_n} \log P_n(A) \le -\inf_{x\in A} I(x). \tag{2}$$

A sequence $(X_n)_n$ of random variables with values in V is said to obey a LDP with rate function I satisfying 1. and 2. if the corresponding distributions satisfy 3. and 4.. In this case we will shortly write
$$P(X_n \sim a) \approx e^{-a_n I(a)}.$$

Throughout this paper $(X, \mathcal{B}(X), \pi)$ denotes a probability space and, except otherwise specified, $(\Omega, \mathcal{F}, P)$ is its $\mathbb{N}$-fold product. Define $(X_n)_{n\ge1}$ to be a sequence of i.i.d. random variables with values in X and distribution (law) π, i.e. for every i $P(X_i \in A) = \pi(A)$. Let $h : X^m \to \mathbb{R}^d$ be a measurable and (without loss of generality) symmetric function. The function
$$U_n := U_n(h, X_1, \ldots, X_n) := \frac{1}{\binom{n}{m}} \sum_{1\le i_1<\ldots<i_m\le n} h(X_{i_1}, \ldots, X_{i_m}), \quad m \le n \tag{3}$$
is called U-statistic of degree m with kernel h.
$$V_n := V_n(h, X_1, \ldots, X_n) := \frac{1}{n^m} \sum_{1\le i_1,\ldots,i_m\le n} h(X_{i_1}, \ldots, X_{i_m}) \tag{4}$$