Bull. Korean Math. Soc. **32** (1995), No. 2, pp. 297–302

ON THE RATE FUNCTION OF CONVERGENCE IN LARGE DEVIATION PROBABILITY

Dae-Sik Hwang

In [1], Cramér introduced the first large deviation theorem at a probability symposium in 1938. Since 1938, there have been many developments in the large deviation theorems.

Donsker and Varadhan [2, 3] have developed a powerful machinery in a series of papers to deal with many old and new problems in estimates of large deviation probabilities.

Ellis [5] and Gärtner [6] obtained useful and surprising generalizations in the dependence of the random variables using the moment generating functions.

Large deviation questions have been successfully dealt with for various stochastic processes [4, 10, 11].

Let $\{X_n : n \geq 1\}$ be a sequence of real-valued independent random variables with a common distribution F and mean m, and we denote $S_n = \sum_{i=1}^{n} X_i$ for each integer n. The strong law of large numbers tells us that the sample mean $\frac{S_n}{n}$ of independent random variables converges to the true mean m almost surely. The first large deviation results were those of Cramér (1938) and Chernoff (1952). Cramér studied the problem of finding the rate at which the probability of $P\left\{\frac{S_n}{n} \in (a,b)\right\}$ tends to zero as $n \to \infty$ for $a < b$ and $m \notin [a,b]$. Also Chernoff's theorem says that the probability of the sample mean being away from the true mean vanishes exponentially fast in the number of samples. This is called a large deviation result because it provides an estimate for the probability that the sample mean will make a large deviation about its typical behavior.

We assume that the readers are familiar with the language of random measures. As a good reference for details, Kallenberg's book [8] may be

Received April 11, 1994.
1991 AMS Subject Classification: 60F10.
Key words: Rate of convergence, Rate function, Random measure.

consulted. Let β^d denote the collection of Borel subsets of R^d. M will denote the set of Randon Borel measures on R^d, so that if $\mu \in M$, then it is finite on bounded Borel sets. Let Ξ be the σ-algebra of subsets of M generated by sets of the form $\{\mu \in M \mid \mu(B) < k\}$ for a bouded Borel set B and a non-negative real number k. If B is a Borel set, then we let $\mu(B)$ be the random amount of mass the measure μ gives to B.

DEFINITION 1. A random measure X is a measurable function from $(\Omega, \Im, P)$ to (M, Ξ), where $(\Omega, \Im, P)$ is a fixed probability space. The induced measure $Px = P \circ X^{-1}$ on (M, Ξ) is the distribution of X.

Here, let N be a random measure on (R^d, β^d). All random measures will be assumed to be stationary, i.e., with a translation invariant distribution. The most well known random measures are the Poisson point process with intensity $\alpha > 0$ and the Poisson center cluster process.

Given $N \in M$ and $f \in C_0(R^d)$, we define the integral $N(f)$ of f with respect to N by $N(f) = \int_R^\alpha f(x)\, N\,(dx) = \int f dN$.

Suppose that $\frac{N_\gamma(f)}{\gamma^d}$ converges to some constant $E\left[\frac{N_\gamma(f)}{\gamma^d}\right]$ a.r. as $\gamma \to \infty$ for an ergodic $N \in M$ and $f \in C_0(R^d)$.

This paper is concerned with rate of convergence of this time average. That is to say, we shall check that $I_B(\mu)$ is a rate function.

Let $\tilde{N}_\gamma$ (denoted by μ_τ, later) be a random measure restricted to $M[B]$. We define f_r and $\tilde{N}_r$ by $f_r(x) = f\left(\frac{x}{r}\right)$ and $\tilde{N}_r(f) = \tilde{N}(f_r)$.

We also define the cumulant generating functional of μ_r as follows:

$$\Phi_{\mu_r}(f) = \frac{1}{r^d} \cdot \log E\left[e^{\mu_r(f)}\right] \quad \text{for} \quad f \in C[B].$$

We shall assume here that

(a) Each function $\Phi_{\mu_r}(f)$ is finite for every $f \in C[B]$, and

(b) $\Phi_\mu(f) = \lim_{r\to\infty} \Phi_{\mu_r}(f)$ (we denote $\Phi_\mu(f)$ by $\Phi(f)$) exists and is also finite for $f \in C[B]$.

The Cramér transform $I_B(\mu)$ of $\Phi(f)$ is defined as the convex conjugate of $\Phi(f)$, i.e.,

$$I_B(\mu) = \sup_{f \in C[B]} \{\mu(f) - \Phi(f)\} \quad \text{for a bounded Borel subset } B \text{ in } R^d.$$

From now on, let B be a compact set in R^d.

FACT1 $I_B(\mu)$ is non-negative on $M[B]$, i.e., $0 \leq I_B(\mu) \leq \infty$ for $\mu \in M[B]$:
We take a zero function f on B. Then $\mu(f) - \Phi(f) = 0$. So, $I_B(\mu) \geq 0$. This implies that $I_B(\mu) \in [0, \infty]$.

FACT2 $I_B(\mu)$ is lower semicontinuous and convex on $M[B]$:
Since $\mu_n \to \mu$ if and only if $\mu_n(f) \to \mu(f)$ for $f \in C[B], \mu(f)$ is continuous for μ. Thus $\mu(f) - \Phi(f)$ is continuous. Sup of continuous function is lower semicontinuous. $I_B(\mu)$ is clearly convex by the Hólder's inequality. Thus $I_B(\mu)$ is a lower semicontinuous and convex function.

From FACT 1 and 2, we can get the following theorem.

THEOREM 2. *A function $I_B(\cdot)$ is a rate function such that*

(i) *$I_B(\mu)$ is non-negative,*
(ii) *$I_B(\mu)$ is lower semicontinuous and*
(iii) *$I_B(\mu)$ has compact level sets, i.e., the set $\{\mu \in M[B] : I_B(\mu) < \alpha\}$ is compact in $M[B]$ for each real $\alpha < \infty$.*

Proof. From FACT 1, 2, it suffices to show that $I_B(\mu)$ has compact level sets. Let $K_\alpha = \{\mu \in M[B] : I_B(\mu) < \alpha\}$ for real $\alpha < \infty$. Since $C[B]$ is separable, let $f_0, f_1, \cdots, f_m, \cdots$ be countable dense subsets. Let us take $f_0 = 1_B$. We take any sequence $\{\mu_n\}$ in K_α and f_m in $C[B]$. Then, $\mu_n(f_m) \leq \alpha + \Phi(f_m)$. By the diagonal process, $\{\mu_n\}$ has a subsequence $\{\mu_{nk}\}$ such that $\mu_{n_k}(f_m)$ converges for $f_0, f_1, \cdots, f_m, \cdots$ as $k \to \infty$. Since $\lim_{k \to \infty} \int f_0 d\mu_{n_k}$ exists, let the valus denote by $\mu_{n_\infty}(B)$. Thus, there exists $\delta > 0$ such that, for n_∞ big enough, $|\mu_{n_\infty}(B) - \lim_{k \to \infty} \int 1_B d\mu_{n_k}| < \delta$. For every continuous function f_m in $C[B]$ and $\varepsilon > 0$, $\lim_{k \to \infty} \int f_m d\mu_{n_k}$ exists and $\|f_m - f\| \leq \varepsilon$.

$$\lim_{k \to \infty} \int f d\mu_{n_k} = \lim_{k \to \infty} \int (f - f_m) d\mu_{n_k} + \lim_{k \to \infty} \int f_m d\mu_{n_k}$$
$$\leq \varepsilon \cdot \lim_{k \to \infty} \int 1_B d\mu_{n_k} + \lim_{k \to \infty} \int f_m d\mu_{n_k}.$$

So,

$$(1) \qquad \lim_{k \to \infty} \int f d\mu_{n_k} \leq \varepsilon \cdot \mu_{n_\infty}(B) + \liminf_{m} \lim_{k \to \infty} \int f_m d\mu_{n_k}.$$

Similarly, replacing f_n by f,

$$\text{(2)} \qquad \limsup_m \lim_{k\to\infty} \int f_m d\mu_{n_k} \le \varepsilon \cdot \mu_{n_\infty}(B) + \lim_{k\to\infty} \int f d\mu_{n_k}.$$

Letting $\varepsilon \to 0$ from (1) and (2), we get

$$\limsup_m \lim_{k\to\infty} \int f_m d\mu_{n_k} \le \lim_{k\to\infty} \int f d\mu_{n_k} \le \liminf_m \lim_{k\to\infty} \int f_m d\mu_{n_k}.$$

So, the limit exists.

Define $\mu(f) := \lim_{k\to\infty} \int f d\mu_{n_k}$.

μ is obviously linear because of the linearity of an integral, i.e.,

$$\begin{aligned}\mu(f+g) &= \lim_{k\to\infty} \int (\alpha f + g) d\mu_{n_k} \\ &= \alpha \cdot \lim_{k\to\infty} \int f d\mu_{n_k} + \lim_{k\to\infty} \int g d\mu_{n_k} \\ &= \alpha\mu(f) + \mu(g) \quad \text{for} \quad \alpha \in R \quad \text{and} \quad f, g \in C[B].\end{aligned}$$

$\mu(f)$ defines a non-negative linear functional on $C[B]$. By the Markov-Riesz representaton theorem, the limiting linear function is represented as an integral and $\mu(f) = \int f d\mu$.

So,

$$\mu(f) = \int f d\mu = \lim_{k\to\infty} \int f d\mu_{n_k}.$$

This means that μ_{n_k} converges to μ weakly. Therefore, K_α is compact.

DEFINITION 3. X is called a Poisson point process with intensity $\alpha > 0$ if

(1) $X(A)$ is a Poisson random variable with parameter $\alpha|A|$ for every bounded Borel subset A in R^d and

(2) If $A_1, \cdots, A_n$ are disjoint bounded Borel subsets of R^d, then $X(A_1), \cdots, X(A_n)$ are independent Poisson random variables with respective parameters $\alpha|A_1|, \cdots, \alpha|A_n|$, where $|\cdot|$ denotes Lebesgue measure.

Definition 4. Let U be a stationary Poisson process on R^d with intensity $\alpha > 0$. V is a point process which has finite total expected mass, $\zeta = E[V(R^d)]$. Let x_i be the random occurrences of U and let $\{V_i\}$ be independent identically distributed (i.i.d) copies of V that are also independent of U. The resulting cluster process X is said to be a Poisson center cluster process, which is defined by superimposing i.i.d. copies of V centered at the occurrences of U.

In other words, If A is a bounded Borel subset of R^d, then X is defined by

$$X(A) = \sum_{x_i} V_i(A - x_i).$$

Note that $E[X(A)] = \alpha\zeta|A|$.

Example 5. (a) For the case of a Poisson random measure with $\alpha > 0$, we get the rate function $I_B(\mu)$ such that

$$I_B(\mu) = \sup_{f \in C[B]} \left\{ \mu(f) - \alpha \cdot \int_B (e^f - 1)dx \right\}.$$

In particular, if $f = 1_A(x)$ for a bounded Borel set A of R^d,

$$I_A(x) = x \cdot \log \left[\frac{x}{\alpha}|A|\right] - x + \alpha|A| \qquad \text{(see [7])}.$$

(b) In the case of a Poisson center cluster random measure, we get the rate function $I_B(\mu)$ such that

$$I_B(\mu) = \sup_{f \in C[B]} \left\{ \mu(f) - \alpha \cdot \int_B \left[e^{V[B](f)} - 1 \right] dx \right\}.$$

If $f = 1_A(x)$ for a bounded Borel set A of R^d,

$$I_A(x) = \sup_{t \in R} \{ tx + \alpha|A| - \alpha|A| M_{V(R^d)}(t) \} \qquad \text{(see [7])}.$$

References

1. Cramér, H., *Sur un nouveau théor éme-limite de la théorie des probabilités*, Actualités Sci. Indust. **736** (1938), 5-23.
2. Donsker, M. D. and Varadhan, S. R., *Asymptotic evaluation of certain Markov process expectations for large time. III*, Comm. Pure Appl. Math. **29** (1976), 389-461.
3. Donsker, M. D. and Varadhan, S. R., *Asymptotic evaluation of certain Markov process expectations for large time. II*, Comm. Pure Appl. Math. **28** (1975a), 1-47 (1975b), 279-301.
4. Ellis, R., *Entropy, Large Deviations and Statistical Mechanics*, Berlin-Heidel berg-New-York, Springer, 1985.
5. Ellis, R., *Large deviations for a general class of dependent random vectors*, Ann. Probab **12** (1984), 1-12.
6. Gártner, J., *On large deviations from the invariant measure*, Theor. Probab. Appl **22** (1977), 24-39.
7. Hwang, D. S., *Large deviation principles of random variables obtained by scaling random measure*, Comm. Korean Math. Soc. **8** (1993), 529-543.
8. Kallenberg, O., *Random Measures*, Academic Press, New York, 1983.
9. Rockafellar, R. T., *Convex analysis*, Princeton Univ. Press, Princeton, N. J., 1970.
10. Stroock, D. W., *An introduction to the theory of large deviations*, Berlin-Heidel berg-New York-Tokyo, Springer, 1984.
11. Varadhan, S. R. S., *Large deviations and applications*, SIAM CBMS-NSF Regional Conf., 1984.

DEPARTMENT OF MATHEMATICS, FACULTY BOARD, AIR FORCE ACADEMY, CHUNG-BUK 363-840, KOREA