Extremal Large Deviations in Controlled I.I.D. Processes with Applications to Hypothesis Testing

NAHUM SHIMKIN

Institute for Mathematics and its Applications
University of Minnesota
Minneapolis, MN 55455, USA

January 1993

Abstract

We consider a controlled i.i.d. process, where several i.i.d. sources are sampled sequentially. Any causal sampling policy, possibly history-dependent, may be employed. The purpose is to characterize the extremal large deviations of the sample mean, namely to obtain asymptotic rate bounds (similar to and extending Cramér's theorem) which hold uniformly over all sampling policies. Lower and upper bounds are obtained, and it is shown that in many (but not all) cases stationary sampling policies are sufficient to obtain the extremal large deviations rates. These results are applied to a hypothesis testing problem, where data samples may be sequentially chosen from several i.i.d. sources (representing different types of experiments). The analysis provides asymptotic estimates for the error probabilities, corresponding both to optimal and worst-case sampling policies.

1

1   Introduction

A basic result in large deviations (LD) theory is Cramér's theorem, which provides the exponential decay rates of the probabilities $P\{\overline{X}_n \in \Gamma\}$, where Γ is a fixed set and $\overline{X}_n$ is the sample mean of an i.i.d. (independent, identically distributed) sequence. For an exposition of this and other aspects of the theory and its application, see, e.g., the recent texts Deuschel and Stroock [6], Bucklew [2], Dembo and Zeitouni [7], Shwartz and Weiss [11]. This paper concerns a generalization of the Cramér-type theory to *controlled* i.i.d. processes. The main goal is to characterize the extremal decay rates, namely the minimal and maximal rates achievable by any sampling policy.

The model consists of several different sources of i.i.d. sequences, which are being sequentially sampled by a controller. At each time instant $n = 1, 2 \ldots$ the controller may choose the source to be sampled next (hence choose the probability distribution of the next sample X_n). While the value of the next sample is not known in advance, the controller does recall the values of past samples $\{X_1, \ldots, X_{n-1}\}$ and may use this information to direct his choice. A *sampling policy* is any causal sampling rule that the controller may choose to employ.

The process described above may be viewed as a degenerate (one-state) version of a controlled Markov process (e.g. [10], [1]). We shall use here some terminology from this field. In particular, a *stationary* policy is a sampling policy in which choices do not depend on the time or on the process history. The use of randomized decisions is however allowed, so that stationary policies refer here both to randomized and non-randomized ones.

The exponential decay rate of $P\{\overline{X}_n \in \Gamma\}$ obviously depends on the sampling policy. The problem addressed is to bound the extremal rates, i.e. the minimal and maximal rates which may be attained by any sampling policy, and furthermore find sampling policies which attain the extremal rates.

It is important to note the following facts concerning stationary and non-stationary policies. Under a stationary sampling policy, the induced sequence $\{X_n\}$ is an i.i.d. sequence, so the standard Cramér-type theory applies. However, a large deviations principle need not hold for a general (non-stationary, history dependent) sampling policy. Such policies are an important part of our problem; indeed, for fixed n and Γ, the sampling policy which minimizes or maximizes $P\{\overline{X}_n \in \Gamma\}$ (and which may be computed by dynamic programming) is generally history-dependent (i.e., the optimal decision at the second stage depends on the value of X_1, and so forth). Nonetheless, it will be shown that stationary policies do achieve the extremal decay rates in many (but not all) cases of interest.

The main results may be summarized as follows, starting with the upper bounds on $P\{\overline{X}_n \in \Gamma\}$. Here we are interested in the *smallest* decay rate achievable by any sampling policy. A tight upper bound is established, which is exactly the same as the one supplied by Cramér's theorem over the *stationary* policies. Thus, the smallest decay rate may always be attained by stationary policies. Deterministic sampling schemes which may replace *randomized* stationary policies in achieving the bound are also considered. As we turn to the lower bound (corresponding to maximal rates of decay), the situation is somewhat more involved. For the case of a *convex* set

Γ the results are similar to ones just described, namely a tight lower bound is established which may be attained by a stationary policy. However, if Γ is non-convex, non-stationary policies may perform significantly better than any stationary one. For this general case a lower bound is provided which, although it recovers the Cramér lower bound for the uncontrolled case, may be non-tight in general.

The problem considered in this paper corresponds to the simplest large deviations problem, namely large deviations of the sample mean for real (although vector-valued) i.i.d. random variables. Consideration of vector-values (rather than scalar) random variables serves to bring out the role of convexity in the proof of the lower bound. Further generalizations have been avoided here in the purpose of emphasizing the main ideas involved in the controlled problem. By the same token, finite moment generating functions are assumed throughout. This assumption may be weakened, using similar methods to the ones employed in the uncontrolled case (cf. [7]).

As an application of these results, we shall consider a problem of (Bayesian) hypothesis testing, where the data is sequentially sampled from several available i.i.d sources. This conforms to the situation where several types of experiments may be performed sequentially in order to test a common property. For the classical case of i.i.d. data sequence, a well known result of Chernoff provides the exponential rate at which the optimal error probability decays as the sample size increases. Bounds of similar type will be developed here for the controlled problem, in the two-hypotheses case. Decay rates are provided for the minimal and maximal (worst case) error probabilities over all sampling policies, and are shown to be achieved by stationary sampling policy. The multiple-hypothesis case is briefly considered, and it is argued that only the results for the maximal error probabilities (but *not* for the minimal ones) can be extended.

Regarding notations, $\mathbb{R}^d$ is the d-dimensional Euclidean space, $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ denote the corresponding norm and inner product. Γ^o and $\overline{\Gamma}$ are the interior and closure of the set $\Gamma \in \mathbb{R}^d$, which is assumed to be Borel measurable throughout. $\mathbf{1}\{A\}$ denotes the indicator function of an event A.

2 The Controlled Model

Let $\{P_u, u \in \mathcal{U}\}$ be a finite collection of probability distributions on $\mathbb{R}^d$, $d \geq 1$, indexed by a finite set of *actions* $u \in \mathcal{U}$. At each time instant $n = 1, 2, \ldots$ the controller chooses an action $u_n \in \mathcal{U}$, and consequently a sample X_n is obtained according to the distribution P_{u_n}. Given u_n, the random variable X_n is independent of prior actions and samples. Let $\overline{X}_n = n^{-1} \sum_{k=1}^{n} X_k$ denote the sample mean.

A sampling policy is a causal rule for choosing actions, which may depend on the entire observed history of the process and may include randomized decisions. More formally, a policy π is a sequence of maps $\pi_n : \mathcal{H}_n \to \Delta(\mathcal{U})$, $n \geq 1$, where $\mathcal{H}_n$ is the set of possible histories $h_n = (u_1, X_1, \ldots, u_{n-1}, X_{n-1})$, and $\Delta(\mathcal{U})$ is the set of probability vectors over $\mathcal{U}$. The action u_n is chosen according to the probability vector $\pi_n(h_n)$ over $\mathcal{U}$. The maps π_n are assumed measurable

so that the above description induces a well defined stochastic process, i.e. a probability measure $\mathbb{P}^\pi$ is induced on the canonical sample space $\mathcal{H}_\infty$ endowed with the product σ–field. Let Π denote the set of all such sampling policies, and let E^π denote the expectation operator corresponding to $\mathbb{P}^\pi$.

A sampling policy is said to be *stationary* if $\pi_n \equiv q$ for every $n \geq 1$, where q is a fixed probability vector. We shall use the notation q both for this probability vector and for the stationary policy it induces. Each $q = (q_u) \in \Delta(\mathcal{U})$ induces also a probability distribution over $\mathbb{R}^d$, according to

$$P_q(\cdot) = \sum_u q_u P_u(\cdot) \, .$$

P_q is just the probability distribution of the sample X_n if the corresponding action u_n is chosen randomly according to q. Let E_u and E_q denote the expectations corresponding to P_u and P_q. With some abuse of notations, X_1 is also used to denote a canonical random variable with respect to the distribution P_q. More precisely, $P_q\{\Gamma\}$ and $P_q\{X_1 \in \Gamma\}$ will be used interchangeably, and $E_q(f(X_1)) \stackrel{\triangle}{=} \int f(x) P_u(dx)$. The logarithmic moment generating function (m.g.f.) associated with P_q is

$$\Lambda_q(\lambda) = \log E_q(e^{\langle \lambda, X_1 \rangle}) \, , \quad \lambda \in \mathbb{R}^d \, ,$$

and its Legendre transform is defined by

$$\Lambda_q^*(x) = \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x \rangle - \Lambda_q(\lambda)\} \, , \quad x \in \mathbb{R}^d \, . \tag{2.1}$$

The logarithmic m.g.f. and Legendre transform associated with P_u are define similarly and denoted by $\Lambda_u(\lambda)$ and $\Lambda_u^*(x)$.

The main technical assumption, assumed to hold throughout this paper, is the following.

Assumption A. Λ_u is finite for every $u \in \mathcal{U}$, i.e. $\Lambda_u(\lambda) < \infty$ for all λ.

This is easily seen to imply that Λ_q is finite for every $q \in \Delta(\mathcal{U})$. Note however that Λ_u^* and Λ_q^* may assume infinite values.

3 Preliminaries

This section presents the basic large deviations results for i.i.d. sequences, in the form relevant to this paper, as well as some properties of the associated rate functions.

Consider a fixed probability vector $q \in \Delta(\mathcal{U})$, which corresponds to a stationary sampling policy $\pi \equiv q$. The sampled sequence $\{X_k\}$ induced by this policy is clearly i.i.d., with marginal distribution $P_q = \sum_u q_u P_u$. Recall that Λ_q^* is the Legendre transform of the logarithmic m.g.f. associated with P_q, and that Assumption A is in effect. Applying the multi-dimensional version of Cramér's theorem yields the following LD bounds:

Theorem 3.1 *For any set $\Gamma \in \mathbb{R}^d$, the following upper and lower bounds hold:*

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}^q \{\overline{X}_n \in \Gamma\} \leq -\inf_{x \in \overline{\Gamma}} \Lambda_q^*(x),$$

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}^q \{\overline{X}_n \in \Gamma\} \geq -\inf_{x \in \Gamma^\circ} \Lambda_q^*(x).$$

In particular, if the two bounds coincide then an exact decay rate is obtained:

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}^q \{\overline{X}_n \in \Lambda\} = -\inf_{\Gamma} \Lambda_q^*(x).$$

Some useful properties of the functions Λ and Λ^* are next summarized. The following are well known (e.g. [6],[7]).

Lemma 3.1 *For every $q \in \Delta(\mathcal{U})$,*

(i) $\Lambda_q(\cdot)$ *is non-negative and convex.*

(ii) $\Lambda_q^*(\cdot)$ *is non-negative, convex, and lower semi-continuous.*

(iii) *Assume $d = 1$ (scalar case). Then $\Lambda_q^*(x) = \sup_{\lambda \geq 0} \{\lambda x - \Lambda_q(\lambda)\}$ for all $x \geq E_q(X_1)$, and $\Lambda_q^*(x)$ is non-decreasing in that range.*

The following properties, where q is considered a variable, will also be required:

Lemma 3.2

(i) *For every $\lambda \in \mathbb{R}^d$, $\Lambda_q(\lambda)$ is concave in $q \in \Delta(\mathcal{U})$.*

(ii) $\Lambda_q^*(x)$ *is jointly convex in $(x, q) \in \mathbb{R}^d \times \Delta(\mathcal{U})$.*

Proof: Recall that $\Lambda_q(x) = \log E_q(e^{\langle \lambda, X_1 \rangle})$, where the expectation operator E_q is linear in $q \in \Delta(\mathcal{U})$, i.e. $E_q = \sum_j \alpha_j E_{q_j}$ if $q = \sum_j \alpha_j q_j$. Since $\log(\cdot)$ is a concave function then (i) follows by Jensen's inequality. It now follows that the function $f_\lambda(x, q) := [\langle \lambda, x \rangle - \Lambda_q(\lambda)]$ is jointly convex in (x, q) for each fixed λ, so that $\Lambda_q^*(x)$ is a point-wise supremum of a convex function, hence convex. $\qquad\square$

4 Extremal Large Deviations – Upper Bound

We return now the the controlled i.i.d. model of Section 2, with the general set Π of sampling policies. In this section we establish the upper bound, which corresponds to the minimal exponential decay rate of $\mathbb{P}^\pi \{\overline{X}_n \in \Gamma\}$ over all policies in Π.

Let Γ be a measurable set in $\mathbb{R}^d$. Note first that for a stationary policy $q \in \Delta(\mathcal{U})$ the upper bound of Theorem 3.1 applies. This immediately implies the following uniform bound over all *stationary* policies:

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}^q \{\overline{X}_n \in \Gamma\} \leq -\inf_{q \in \Delta(\mathcal{U})} \inf_{x \in \overline{\Gamma}} \Lambda_q^*(x). \tag{4.1}$$

It will next be established that the same bound holds for all sampling policies, stationary or not.

Theorem 4.1 *For every set* $\Gamma \in \mathbb{R}^d$,

$$\limsup_{n\to\infty} \sup_{\pi\in\Pi} \frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} \leq - \inf_{q\in\Delta(\mathcal{U})} \inf_{x\in\overline{\Gamma}} \Lambda_q^*(x)\,.$$

Before proving this theorem, we state and prove a slightly stronger version for the scalar case with $\Gamma = [a, \infty)$. This result will be used in the proof of Theorem 4.1, is of some independent interest. and admits a simple direct proof which already contains the main ideas relevant to the controlled case.

Proposition 4.1 *Assume $d = 1$ (scalar case), and let $a \geq \max_u E_u(X_1)$. Then for every $n \geq 1$ and $\pi \in \Pi$,*

$$\frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \geq a\} \leq - \inf_{q\in\Delta(\mathcal{U})} \Lambda_q^*(a)\,.$$

Remark: Only $a \geq \max_u E_u(X_1)$ need be considered, since otherwise the event $\{\overline{X}_n \geq a\}$ becomes "non-rare" under appropriate policies. Indeed, the repeated choice of any action u for which $E_u(X_1) > a$ obtains $P\{\overline{X}_n \geq a\} \to 1$.

Proof: Fix a in the required range. For every $\pi \in \Pi$, $n \geq 1$ and $\lambda \geq 0$ Chebycheff's inequality implies

$$\mathbb{P}^\pi\{\overline{X}_n \geq a\} \leq E^\pi(e^{\lambda n(\overline{X}_n - a)}) = e^{-\lambda n a} E^\pi(e^{\lambda S_n})\,, \tag{4.2}$$

where $S_n = n\overline{X}_n = \sum_{k=1}^n X_k$. Now, a standard dynamic programming argument establishes that

$$E^\pi(e^{\lambda S_n}) \leq [\max_{q\in\Delta(\mathcal{U})} E_q(e^{\lambda X_1})]^n\,. \tag{4.3}$$

Indeed, this follows by iterating the inequality:

$$\begin{aligned}
E^\pi(e^{\lambda S_n}) &= E^\pi\{E^\pi(e^{\lambda S_n}|h_n, u_n)\} \\
&= E^\pi\{e^{\lambda S_{n-1}} E^\pi(e^{\lambda X_n}|u_n)\} \\
&\leq E^\pi(e^{\lambda S_{n-1}}) \max_{q\in\Delta(\mathcal{U})} E_q(e^{\lambda X_1})\,. \tag{4.4}
\end{aligned}$$

(Note that $\max_q E_q(e^{\lambda X_1})$ may be simplified to $\max_u E_u(e^{\lambda X_1})$, since $E_q = \sum_u q_u E_u$. The former expression will however be useful for the minimax equality below.) Since (4.2) and (4.3) hold for every $\lambda \geq 0$, then

$$\begin{aligned}
\frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \geq a\} &\leq \inf_{\lambda\geq 0}\{-\lambda a + \log(\max_{q\in\Delta(\mathcal{U})} E_q(e^{\lambda X_1}))\} \\
&= -\sup_{\lambda\geq 0} \min_{q\in\Delta(\mathcal{U})}\{\lambda a - \Lambda_q(\lambda)\}\,. \tag{4.5}
\end{aligned}$$

We now use the minimax equality in the last expression. From Lemma 3.1(*i*) and Lemma 3.2(*i*) it follows that the function $f(q, \lambda) := \{\lambda a - \Lambda_q(\lambda)\}$ is convex in q and concave in λ. The set $\Delta(\mathcal{U})$

is obviously convex and compact, while $[0, \infty)$ is closed convex. Thus, by a standard minimax theorem ([9], Corollary 37.3.2),

$$\sup_{\lambda \geq 0} \min_{q \in \Delta(\mathcal{U})} \{\lambda a - \Lambda_q(\lambda)\} = \inf_{q \in \Delta(\mathcal{U})} \sup_{\lambda \geq 0} \{\lambda a - \Lambda_q(\lambda)\} = \inf_{q \in \Delta(\mathcal{U})} \Lambda_q^*(a), \tag{4.6}$$

where the last equality follows by Lemma 3.1(iii). The required upper bound follows by substituting (4.6) in (4.5). $\qquad\square$

Proof of Theorem 4.1: We shall use here a result from de Acosta [5], which is a somewhat strengthened version of the Gärtnet-Ellis type upper bound. It is readily verified by a dynamic programming argument, similarly to (4.3), that

$$\frac{1}{n} \log \sup_{\pi \in \Pi} E^{\pi} e^{\langle S_n, \lambda \rangle} \leq \sup_{q \in \Delta(\mathcal{U})} \Lambda_q(x), \quad n \geq 1. \tag{4.7}$$

By Theorem 2.1(i) in [5], this implies that for every compact Γ,

$$\limsup_{n \to \infty} \frac{1}{n} \log \sup_{\pi \in \Pi} \mathbb{P}^{\pi}\{\overline{X}_n \in \Gamma\} \leq -\inf_{x \in \Gamma} I(x), \tag{4.8}$$

where

$$I(x) = \sup_{\lambda \in \mathbb{R}^d} \left\{ \langle \lambda, x \rangle - \sup_{q \in \Delta(\mathcal{U})} \Lambda_q(x) \right\}. \tag{4.9}$$

Using the minimax theorem it follows as in (4.6) that

$$I(x) = \sup_{\lambda \in \mathbb{R}^d} \inf_{q \in \Delta(\mathcal{U})} \{ \langle \lambda, x \rangle - \Lambda_q(x) \} = \inf_{q \in \Delta(\mathcal{U})} \Lambda_q^*(x), \tag{4.10}$$

which establishes the required upper bound for compact sets. The extension to any closed set follows readily from the following uniform exponential tightness property (cf. [5], Theorem 2.1(ii)):

$$\lim_{\rho \to \infty} \limsup_{n \to \infty} \sup_{\pi \in \Pi} \frac{1}{n} \log \mathbb{P}^{\pi}\{\|\overline{X}_n\| > \rho\} = -\infty. \tag{4.11}$$

The latter may be deduced from Proposition 4.1. Indeed, for every π and n,

$$\mathbb{P}^{\pi}\{\|\overline{X}_n\| > \rho\} \leq \mathbb{P}^{\pi}\left\{\frac{1}{n} \sum_{k=1}^{n} \|X_k\| > \rho\right\} \tag{4.12}$$

Now $\{\|X_k\|\}$ may be considered a (scalar-valued) controlled i.i.d. sequence. It is readily verified that Assumption A is satisfied, and the upper bound of Proposition 4.1 applies. It is also easily verified that the corresponding rate function $\Lambda^*(\rho) \to \infty$ as $\rho \to \infty$ ([7], Lemma 2.2.20), and (4.11) follows. $\qquad\square$

Remarks:

1. The results just established provide bound on the decay rate of $\mathbb{P}^{\pi}\{\overline{X}_n \in \Gamma\}$ for any fixed policy π. Moreover, noting that the supremum over Π is taken before the limit in n, these results provide bounds on the maximized probabilities $\max_{\pi} \mathbb{P}^{\pi}\{\overline{X}_n \in \Gamma\}$, where the maximizing policy is generally different for each time horizon n.

2. Note that $\Lambda_q^*(x)$ is convex in (q, x) (Lemma 3.2), which may facilitate the computation of the bound.

3. The upper bound is tight, up to certain continuity gaps which are inherent in most large deviations results. Indeed, assume that the supremum in the upper bound is achieved at some $q_o \in \Delta(\mathcal{U})$, and that Γ is a lower continuity set of $\Lambda_{q_o}^*(\cdot)$ (i.e. the infima over Γ° and $\overline{\Gamma}$ are equal). Then Theorem 3.1 implies that the upper bound is attained by the stationary policy q_0.

4. If the infimum over $\Delta(\mathcal{U})$ is not attained, then only ϵ–optimal stationary policies may exist. Noting however that $\Lambda_q(x)$ is a convex function, hence smooth (except for possible jumps to infinity), the latter case seems the exception rather than the rule.

5. The stationary policy q_o which attains the upper bound may be randomized, implying that (random) mixing of two or more actions can perform better than any single one. The following example illustrates this fact. Let the action set $\mathcal{U} = \{1, 2\}$ correspond to distributions $P_1 = \mathcal{N}(0, 1)$ and $P_2 = \mathcal{N}(-40, 81)$ (the Normal distribution with mean $m = -40$ and variance $\sigma^2 = 81$). Let us examine the exponential decay rate of $P\{X_n \geq 4\}$. The rate function for a Normal random variable is $\Lambda^*(x) = (x - m)^2 / 2\sigma^2$. If $u = 1$ is chosen at all times (corresponding to the stationary strategy $q = (1, 0)$), the resulting rate from Cramèr's theorem is $\Lambda_1^*(4) = 4^2/2 = 8$, while always choosing $u = 2$ results in $\Lambda_2^*(4) = 44^2/162 = 11.95$. However, some calculation shows that the upper bound in Proposition 4.1 is $\min_q \Lambda_q^*(4) = 3.5$, with the minimum uniquely achieved at $q_o = \left(\frac{39}{40}, \frac{1}{40}\right)$. Thus, the policy of choosing P_1 with probability $39/40$ and P_2 with probability $1/40$ at each step results in a rate which is significantly lower than that obtained by either P_1 or P_2 alone.

As the last example demonstrates, randomized decisions may be necessary in order to attain the upper bound via a *stationary* policy. Since randomizations may sometimes be undesirable in an actual decision rule, it is of interest to find simple non-randomized policies which also achieve the bound. We conclude this section by briefly considering such sampling schemes.

Given a probability vector $q \in \Delta(\mathcal{U})$, let $\pi(q)$ denote a policy which pre-specifies deterministically the action sequence $\{u_n\}$, so that the relative frequencies of these actions converge to q. For example, for $q = (\frac{1}{4}, \frac{3}{4})$ one can choose $u = 1$ every fourth time. Observe first that the sample means $\{\overline{X}_n\}$ obey a large deviations principle under this policy. Indeed, the following limit exists:

$$
\begin{aligned}
\hat{\Lambda}_q(\lambda) \quad &\triangleq \quad \lim_{n \to \infty} \frac{1}{n} \log E^{\pi(q)} \exp(\langle \lambda, \sum_{k=1}^n X_k \rangle) \\
&= \quad \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n \log E_{u_k}(e^{\langle \lambda, X_1 \rangle}) \\
&= \quad \sum_u q_u \Lambda_u(\lambda) \,, \quad\quad\quad\quad\quad\quad\quad (4.13)
\end{aligned}
$$

and it follows by the Gärtner-Ellis Theorem ([8], Theorem II.2) that a so-called large deviations principle is satisfied with rate function $\hat{\Lambda}_q^*$ (the Legendre transform of $\hat{\Lambda}_q$), which means that the

8

bounds in Theorem 3.1 are satisfied with $\hat{\Lambda}_q^*$ replacing Λ_q^*. The following Lemma compares the two rate functions Λ_q^* (corresponding to randomized stationary policies) and $\hat{\Lambda}_q^*$ (corresponding to deterministic sampling schemes).

Lemma 4.1 *For every* $x \in \mathbb{R}^d$,

  (i) $\Lambda_q^*(x) \leq \hat{\Lambda}_q^*(x)$ *for every* $q \in \Delta(\mathcal{U})$.

  (ii) $\inf_q \Lambda_q^*(x) = \inf_q \hat{\Lambda}_q^*(x)$.

  (iii) *Any* q_o *which minimizes* $\Lambda_q^*(x)$ *also minimizes* $\hat{\Lambda}_q^*(x)$.

Proof: (*i*) By concavity of $\Lambda_q(\lambda)$ in q we have $\Lambda_q(\lambda) \geq \sum_u q_u \Lambda_u(\lambda) = \hat{\Lambda}_q(\lambda)$, which implies that the opposite inequality holds for the respective Legendre transforms.

(*ii*) By the minimax equality (4.10) and its counterpart for $\hat{\Lambda}_q^*$, and by equation (4.13) and the concavity of $\Lambda_q(\lambda)$,

$$
\begin{aligned}
\inf_q \Lambda_q^*(x) &= \sup_\lambda \min_q \{\langle \lambda, x \rangle - \Lambda_q(\lambda)\} = \sup_\lambda \min_u \{\langle \lambda, x \rangle - \Lambda_u(\lambda)\} \\
&= \sup_\lambda \min_q \{\langle \lambda, x \rangle - \hat{\Lambda}_q(\lambda)\} = \inf_q \sup_\lambda \{\langle \lambda, x \rangle - \hat{\Lambda}_q(\lambda)\} \\
&= \inf_q \hat{\Lambda}_q^*(x) \, . \tag{4.14}
\end{aligned}
$$

(iii) Assume that q_o minimizes $\Lambda_q^*(x)$. Then

$$
\begin{aligned}
\hat{\Lambda}_{q_o}^*(x) &= \sup_\lambda \{\langle \lambda, x \rangle - \hat{\Lambda}_{q_o}(\lambda)\} \\
&\geq \sup_\lambda \min_q \{\langle \lambda, x \rangle - \hat{\Lambda}_q(\lambda)\} = \inf_q \Lambda_q^*(x) = \Lambda_{q_o}^*(x) \, , \tag{4.15}
\end{aligned}
$$

where the next-to-last equality is taken from the proof of (*ii*). Combined with part (*i*) of this lemma, this implies that q_o minimizes $\hat{\Lambda}_{q_o}^*(x)$. □

Lemma (4.1) has the following consequences. By part (*i*), the decay rate associated with stationary policies dominates from below the rate associated with deterministic sampling policies. Nonetheless, parts (*ii*) and (*iii*) imply that the minimal rates achievable by either class of policies are equal, namely $\inf_q \inf_{\overline{\Gamma}} \Lambda_q^* = \inf_q \inf_{\overline{\Gamma}} \hat{\Lambda}_q^*$; moreover, if a stationary policy q_o achieves the minimal rate then so does $\pi(q_o)$.

5 Extremal Large Deviations – Lower Bounds

Our purpose in this section is to provide asymptotic lower bounds for $\mathbb{P}\{\overline{X}_n \in \Gamma\}$, which delimit the maximal rate of decay achievable by any sampling policy. The lower-bound part of Theorem 3.1 implies that for any *stationary* policy q the following holds:

$$
\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}^q \{\overline{X}_n \in \Gamma\} \geq - \sup_{q \in \Delta(\mathcal{U})} \inf_{x \in \Gamma^o} \Lambda_q^*(x) \, .
$$

A reasonable conjecture, motivated by the results of the previous section, would be that this lower bound is valid for any policy π. The next example shows that this is false.

Example 5.1. Let $\mathcal{U} = \{1, 2\}$ with P_1 and P_2 the following Bernoulli distributions:

$$P_1 = 0.9\delta_{-1} + 0.1\delta_9 , \quad P_2 = 0.1\delta_{-9} + 0.9\delta_1 ,$$

where δ_a is the probability measure concentrated at a. Note that the expected value equals 0 in both cases. Suppose we wish to minimize the probabilities $\mathbb{P}^\pi\{|\overline{X}_n| > 2\}$, which correspond to the set $\Gamma = (-\infty, -2) \cup (2, \infty)$. Restrict attention first to stationary policies. Since $P_q = q_1 P_1 + q_2 P_2$, then for every $q \in \Delta(\mathcal{U})$ either $P_q(\{-9\}) \geq 0.05$ or $P_q(\{9\}) \geq 0.05$, implying that $\mathbb{P}\{|\overline{X}_n| > 2\} \geq (0.05)^n$. Consider now the following non-stationary 'steering' policy π': let $u_n = 1$ if $\overline{X}_n < 0$, and $u_n = 2$ if $\overline{X}_n \geq 0$. Simple algebraic calculation shows that now $\mathbb{P}^{\pi'}\{|\overline{X}_n| > 2\} = 0$ for every $n \geq 8$.

Although false in general, the above-mentioned conjecture does hold for *convex* sets.

Theorem 5.1 *Let Γ be a convex set in $\mathbb{R}^d$. Then*

$$\liminf_{n \to \infty} \inf_{\pi \in \Pi} \frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} \geq - \max_{q \in \Delta(\mathcal{U})} \inf_{x \in \Gamma^o} \Lambda_q^*(x)$$

$$= - \max_{u \in \mathcal{U}} \inf_{x \in \Gamma^o} \Lambda_u^*(x) .$$

Proof: We first establish the last equality. Since $\Lambda_q^*(x)$ is jointly convex in (q, x) (Lemma 3.2) and Γ^o is a convex set, it follows that $\inf_{\Gamma^o} \Lambda_q^*$ is a convex function of q. Hence it achieves its maximal value at one of the vertices of (the convex set) $\Delta(\mathcal{U})$, which correspond to the action set $\mathcal{U}$.

To prove the bound, let $\{y_u \in \Gamma^o, u \in \mathcal{U}\}$ be an arbitrary set of points in the interior of Γ. We proceed to establish that

$$\liminf_{n \to \infty} \inf_{\pi \in \Pi} \frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} \geq - \max_{u \in \mathcal{U}} \Lambda_u^*(y_u) . \tag{5.1}$$

Since the points $\{y_u\}$ are arbitrary, this clearly implies the bound of the theorem. The inequality (5.1) will first be established under the following assumption, to be removed later.

Assumption B: For every $u \in \mathcal{U}$ there exists $\lambda_u \in \mathbb{R}^d$ such that $y_u = \nabla \Lambda_u(\lambda_u)$.

Here $\nabla \Lambda_u$ denotes the gradient of Λ_u. It follows that λ_u is a maximizer in the definition of $\Lambda_u^*(y_u)$, namely

$$\Lambda_u^*(y_u) = \langle \lambda_u, y_u \rangle - \Lambda_u(\lambda_u) . \tag{5.2}$$

Given $\{\lambda_u\}$ as specified in Assumption B, define for each $u \in \mathcal{U}$ a probability measure $\hat{P}_u$ via the exponential change of measure:

$$\hat{P}_u(dx) \triangleq \frac{e^{\langle \lambda_u, x \rangle}}{E_u(e^{\langle \lambda_u, X_1 \rangle})} P_u(dx) = \exp\{\langle \lambda_u, x \rangle - \Lambda_u(\lambda_u)\} P_u(dx) . \tag{5.3}$$

10

By (5.2),
$$\hat{P}_u(dx) = \exp\{\langle \lambda_u, x - y_u \rangle + \Lambda_u^*(y_u)\} P_u(dx) . \tag{5.4}$$

Also, the definition of λ_u (and the dominated convergence theorem) imply that

$$\hat{E}_u(X_1) \stackrel{\triangle}{=} \int x \hat{P}_u(dx) = \nabla \Lambda_u(\lambda_u) = y_u . \tag{5.5}$$

For any policy $\pi \in \Pi$, let $\hat{\mathbb{P}}^\pi$ denote the measure which is defined similarly to $\mathbb{P}^\pi$, except that the modified measures $\{\hat{P}_u\}$ replace $\{P_u\}$ as the conditional marginal distributions of the samples. Let $\hat{E}^\pi$ denote the corresponding expectation. Then for any $n \geq 1$ and any random variable $Z_n = f(X_1, \ldots, X_n)$,

$$
\begin{aligned}
E^\pi(Z_n) &= \hat{E}^\pi(Z_n \prod_{k=1}^{n} \exp\{-\Lambda_{u_k}^*(y_{u_k}) - \langle \lambda_{u_k}, X_k - y_{u_k} \rangle\}) \\
&\geq \exp\{-n \max_u \Lambda_u^*(y_u)\} \hat{E}^\pi(Z_n \exp\{-\sum_{k=1}^{n} \langle \lambda_{u_k}, X_k - y_{u_k} \rangle\}) .
\end{aligned}
\tag{5.6}
$$

In particular, for $Z_n = \mathbf{1}\{\overline{X}_n \in \Gamma\}$ this yields

$$\frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} \geq -\max_u \Lambda_u^*(y_u) + \frac{1}{n} \log \hat{E}^\pi(V_n) , \tag{5.7}$$

where

$$V_n = \mathbf{1}\{\overline{X}_n \in \Gamma\} \exp(-\sum_{k=1}^{n} \langle \lambda_{u_k}, X_k - y_{u_k} \rangle) . \tag{5.8}$$

Thus, to prove (5.1) it remains to show that

$$\liminf_{n \to \infty} \inf_{\pi \in \Pi} \frac{1}{n} \log \hat{E}^\pi(V_n) \geq 0 . \tag{5.9}$$

To this end, define the random variables

$$W_k = \langle \lambda_{u_k}, X_k - y_{u_k} \rangle , \quad \overline{W}_n = \frac{1}{n} \sum_{k=1}^{n} W_k . \tag{5.10}$$

Note that for any $\delta > 0$,

$$V_n = \mathbf{1}\{\overline{X}_n \in \Gamma\} e^{-n\overline{W}_n} \geq \mathbf{1}\{\overline{X}_n \in \Gamma, |\overline{W}_n| < \delta\} e^{-\delta n} , \tag{5.11}$$

so that

$$\hat{E}^\pi(V_n) \geq e^{-\delta n} \hat{\mathbb{P}}^\pi\{\overline{X}_n \in \Gamma, |\overline{W}_n| < \delta\} . \tag{5.12}$$

We proceed to show that the last probability converges to 1 uniformly over $\pi \in \Pi$. (Convergence may actually be deduced from the weak law of large numbers for Martingales. A direct proof will be provided here to establish uniform convergence.) Let $\mathcal{F}_k$, $k \geq 1$, denote the σ-algebra generated

by the random variables $\{u_1, X_1, \ldots, u_k, X_k\}$. By (5.5) and (5.10) it follows that $\hat{E}^\pi(W_k|\mathcal{F}_{k-1}) = 0$, which in turn implies that $\hat{E}^\pi(W_k W_\ell) = 0$ for $k \neq \ell$. Furthermore,

$$\hat{E}^\pi(W_k^2) \leq \max_u \|\lambda_u\|^2 \, \hat{E}_u(\|X_1 - y_u\|^2) \triangleq C^2 < \infty \; . \tag{5.13}$$

(Finiteness of the moment generating function for P_u implies the same for $\hat{P}_u$, so that the second moments on the right-hand side are indeed finite.) It can now be concluded that

$$\hat{E}^\pi(\overline{W}_n^2) = \frac{1}{n^2} \sum_{k=1}^n \hat{E}^\pi(W_k^2) \leq \frac{C^2}{n} \; , \tag{5.14}$$

and by Chebycheff's inequality

$$\hat{\mathbb{P}}^\pi\{|\overline{W}_n| < \delta\} \geq 1 - \delta^{-2} \hat{E}^\pi(\overline{W}_n^2) \geq 1 - \frac{1}{n} \frac{C^2}{\delta^2} \; . \tag{5.15}$$

Consider next the part of (5.12) which involves $\overline{X}_n$. Recall that y_u is interior to Γ, so there exists a $\delta' > 0$ such that, for every u, the ball $B(y_u, \delta')$ of radius δ' around y_u is contained in Γ. Since Γ is convex by assumption, it follows that $B(y_q, \delta') \subset \Gamma$ for *any* convex combination $y_q = \sum_u q_u y_u$ of $\{y_u\}$. Thus,

$$\hat{\mathbb{P}}^\pi\{\overline{X}_n \in \Gamma\} \geq \hat{\mathbb{P}}^\pi\{\|\frac{1}{n} \sum_{k=1}^n (X_k - y_{u_k})\| < \delta'\} \geq 1 - \frac{1}{n}\left(\frac{C'}{\delta'}\right)^2 \; , \tag{5.16}$$

where the last inequality follows similarly to (5.15), with $C' = \max_u \hat{E}_u \|X_1 - y_u\|^2$. Now (5.15) and (5.16) together imply that

$$\liminf_{n \to \infty} \inf_{\pi \in \Pi} \hat{\mathbb{P}}^\pi\{\overline{X}_n \in \Gamma, |\overline{W}_n| < \delta\} = 1 \; , \tag{5.17}$$

and by (5.12)

$$\liminf_{n \to \infty} \inf_{\pi \in \Pi} \frac{1}{n} \log \hat{E}^\pi(V_n) \geq -\delta \; . \tag{5.18}$$

Since this holds for any $\delta > 0$, then (5.9) follows. This concludes the proof of the bound under Assumption B.

The extension to the case where Assumption B does not hold may be performed exactly as in the standard i.i.d. case. For example, it is possible to employ the regularization procedure in [7] (proof of Theorem 2.2.30), where small Gaussian random variables are added to the original ones to ensure that Assumption B is satisfied. This method applies here without modification, so that further details may be omitted. $\qquad\square$

Remarks: The following remarks apply to the last theorem and its proof.

1. Contrary to the usual proof method for Cramér's theorem, the proof above could not proceed by first establishing the bound for small balls. Such an approach would lead to the bound of Corollary 5.1 below, which is weaker than the one just established.

2. Convexity of Γ enters critically in the convergence $\hat{\mathbb{P}}^\pi\{\overline{X}_n \in \Gamma\} \to 1$ (cf. equation (5.16)).

3. The bound is tight, in the sense discussed in relation with the upper bound, and implies that stationary policies are sufficient to obtain the maximal decay rate.

4. Moreover, as the last equality in the bound implies, the maximal rate is alway achieved on the vertices of $\Delta(\mathcal{U})$. . This means that *non-randomized* stationary policies (choosing the same action u at all times) are sufficient to achieve the maximal rate.

If Γ is a non-convex set the results above do not apply, as demonstrated in Example 5.1. The problem of establishing a tight lower bound and "optimal" policies is not resolved in the present paper. Obviously, some lower bounds may be inferred from the convex result simply by applying it to convex subsets of Γ. In particular, if applied to small balls in Γ, the following bound is obtained.

Corollary 5.1 *For any $\Gamma \in \mathbb{R}^d$,*

$$\liminf_{n\to\infty} \inf_{\pi\in\Pi} \frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} \geq - \inf_{x\in\Gamma^o} \max_{u\in\mathcal{U}} \Lambda_u^*(x) . \qquad (5.19)$$

Proof: For every $x \in \Gamma^o$, let B_x be an open ball around x contained in Γ. By Theorem 5.1,

$$\liminf_{n\to\infty} \inf_\pi \frac{1}{n} \log \mathbb{P}^\pi\{\overline{X}_n \in B_x\} \geq - \max_u \inf_{x'\in B_x} \Lambda_u^*(x') \geq - \max_u \Lambda_u^*(x) , \qquad (5.20)$$

and since $\mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} \geq \mathbb{P}^\pi\{\overline{X}_n \in B_x\}$ the corollary follows. $\qquad\qquad \square$

We remark that the bound of Corollary 5.1 is tight if the inf and the max on the right-hand side can be interchanged, i.e., if $\inf_{x\in\Gamma^o} \max_u \Lambda_u^*(x) = \inf_{x\in\Gamma^o} \Lambda_{u'}^*(x)$ for some $u' \in \mathcal{U}$. In particular, this will be satisfied if the same u' maximizes $\Lambda_u(x)$ for all $x \in \Gamma^o$. The bound can then be attained by always choosing u'. Note also that the bound of Corollary 5.1 reduces to the standard Cramér lower bound when $\mathcal{U}$ is a singleton. However, the bound is *not* tight in general, and does not even coincide with the (tight) lower bound of Theorem 5.1 for the convex case. This is explicitly illustrated by the following example.

Example 5.2. Let $\mathcal{U} = \{1, 2\}$, $P_1 = \mathcal{N}(-1, \sigma^2)$, $P_2 = \mathcal{N}(1, \sigma^2)$ and let Γ be and open interval which contains the interval $[-1, 1]$. Obviously $\lim_{n\to\infty} \mathbb{P}^\pi\{\overline{X}_n \in \Gamma\} = 1$ for every $\pi \in \Pi$, so that the maximal exponential decay rate is 0. Indeed, since $\Lambda_1(x) = (x+1)^2/2\sigma^2$ and $\Lambda_2(x) = (x-1)^2/2\sigma^2$, the rate bound of Theorem 5.1 equals $\max_{u=1,2}\{\inf_{x\in\Gamma} \Lambda_u^*(x)\} = 0$. However, the rate bound of Corollary 5.1 in this case is

$$\inf_{x\in\Gamma} \max_{u=1,2} \Lambda_u^*(x) = \inf_{x\in\Gamma} \frac{(|x|+1)^2}{2\sigma^2} = \frac{1}{2\sigma^2} ,$$

which can be arbitrarily large.

It should be noted that the results of the last two sections remain true in the following slightly more general case (which may be useful in applications, as the next section shows). Consider a

controlled i.i.d. process $\{Y_k\}$, defined similarly to $\{X_k\}$ above, and define a new random sequence $X_k = f(Y_k, u_k)$. We consider the probabilities $P\{\overline{X}_n \in \Gamma\}$. Obviously $\{X_n\}$ may be considered as a controlled i.i.d. process by itself, and the only difference compared to the case studied so far is that the sampling policies are allowed to depend on past values of Y_k and not only on their functions X_k. However, it is easy to show this added information is not relevant in the problem of extremizing $P\{\overline{X}_n \in \Gamma\}$ (cf. [1], Section 4.2), so the bounds above hold in this case as well.

6   Controlled Hypothesis Testing

One of the classical results related to large deviations theory is Chernoff's theorem, concerning the error probabilities in statistical hypothesis testing using an i.i.d. data sequence. The preceding results will now be applied to the generalized situation where samples may be chosen sequentially from several i.i.d. sources. Tight asymptotic bounds will be obtained for the minimal and maximal error probabilities, which may be achieved by appropriate sampling policies. The case of two (simple) hypotheses will be considered, followed by a discussion of the multiple hypothesis case. As it turns out, some of the main results do not carry over to the latter case.

The following is a standard Bayesian hypothesis testing problem. Based on n measurements $\{Y_1 \ldots Y_n\}$, assumed to originate from an i.i.d. source with marginal distribution $P_Y \in \{\mu_0, \mu_1\}$, we have to decide between the two hypotheses $H_1 : P_Y = \mu_1$ and $H_2 : P_Y = \mu_2$ (to which some prior probabilities are assigned). It is well known that the probability of error is minimized by a likelihood ratio test. It was established in Chernoff [3] that these minimal probabilities converge to zero exponentially as the sample size increases, at the following rate. Let $P_n(e|H_j)$ denotes the error probability based on n samples, given that H_j is true and using the optimal likelihood ratio test. Then

$$\lim_{n \to \infty} \frac{1}{n} \log P_n(e|H_j) = -I, \quad j = 1, 2. \tag{6.1}$$

Here I is the *Chernoff information*:

$$I = I(\mu_1, \mu_2) = -\log \inf_{0 \le \lambda \le 1} \int f_1(y)^\lambda f_2(y)^{1-\lambda} \mu(dy), \tag{6.2}$$

where μ is any measure which dominates both μ_1 and μ_2, and f_j is the density of μ_j with respect to μ. Note that the same rate applies to the two conditional errors, hence also to the unconditional error probability $P_n(e)$.

Consider now the generalized problem, where the data sequence $\{Y_1 \ldots Y_n\}$ is obtained by sampling sequentially from several i.i.d. sources. Assume that two hypotheses H_1 and H_2 need to be tested. A set $\mathcal{U}$ of different types of experiments is available, where the probability distribution for (the result of) experiment u under hypothesis H_j is μ_j^u, with corresponding density function $f_j(\cdot|u)$ (all densities are with respect to some measure μ). It is assumed that $\mu_1^u \ne \mu_2^u$. At each stage $k = 1 \ldots n$ exactly one of these experiments u_k is chosen, and Y_k is its result. Independent repetitions of the same type of experiment are allowed.

Since experiments are performed sequentially, choice of the next one may depend on the previous results. Indeed, if experiments are chosen with the purpose of minimizing the error probability, then as new data arrives our estimation of the true state of nature (the true hypothesis) improves, which puts us in a better position to make further decisions. The model considered here and other related models have been studied within the framework of statistical sequential design (see, e.g., [4] and references therein), where the emphasis is on obtaining an optimal balance between the total experiment cost (length) and the error probability. Here we concentrate on choice of experiments, with the sample size not being a decision variable. Thus, either n is fixed at some large value, or the experiment continues indefinitely with the best current estimates of the true hypothesis needed at each stage.

The overall decision problem has two components. The first is the sampling policy π. The set of possible policies Π is defined as in Section 2: u_k may depend on the whole history $\{u_1, Y_1 \ldots u_{k-1}, Y_{k-1}\}$, and randomizations may be used. The second component is the rule for choosing between the hypotheses H_1 and H_2 based on available data. Given the data $\{u_1, Y_1 \ldots u_n, Y_n\}$, the decision rule which minimizes the error probability is still a likelihood ration test. It can be written as

$$\text{accept } H_1 \text{ if } L_n < \alpha_n, \text{ and accept } H_2 \text{ otherwise} \qquad (6.3)$$

where L_n is the normalized log-likelihood ratio:

$$L_n = \frac{1}{n} \sum_{k=1}^{n} \log \frac{f_2(Y_k|u_k)}{f_1(Y_k|u_k)} \, ,$$

and $\{\alpha_n\}$ is a sequence converging to 0. (In the optimal test this sequence is determined by the prior probabilities, but for our asymptotic analysis it is enough that $\alpha_n \to 0$.) This rule is assumed fixed throughout the following, leaving only the sampling policy yet undetermined.

6.1   Minimal Error Probabilities

We shall first consider sampling policies which minimize the error probability. Some definitions and technical assumptions will be required.

Let $u \in \mathcal{U}$ be fixed. Let X be a random variable with distribution P_u specified by

$$X = \log \frac{f_2(Y|u)}{f_1(Y|u)} \, , \text{ where } Y \sim \mu_1^u \, . \qquad (6.4)$$

We identify this distribution with P_u from Section 2, and apply the related notations. In particular, Λ_u denotes the logarithmic m.g.f.:

$$\Lambda_u(\lambda) \overset{\triangle}{=} \log E_u(e^{\lambda X}) = \log \int f_2(y|u)^\lambda f_1(y|u)^{1-\lambda} d\mu \, . \qquad (6.5)$$

The following will be assumed:

Assumption C: $\Lambda_u(\lambda)$ is finite for all $\lambda \in \mathbb{R}$ and $u \in \mathcal{U}$.

This assumption implies that, for each u, the probability measures μ_1^u and μ_2^u are equivalent (assign nonzero probabilities to the same events), and is then satisfied if the ratio $f_2(y|u)/f_1(y|u)$ as well a its inverse are bounded in y. It should be emphasized that assumption C is used only to streamline the proofs, and its removal need not affect the results.

Theorem 6.1 *Assume C. Let $P(e|n,\pi)$ denote the n-stage error probability given the sampling policy π (and the likelihood ratio test (6.3)). Then*

$$\liminf_{n\to\infty} \inf_\pi \frac{1}{n} \log P(e|n,\pi) \geq -\max_{u\in\mathcal{U}} I_u$$

where $I_u = I(\mu_1^u, \mu_2^u)$, as defined in (6.2). Furthermore, equality may be obtained by choosing at each stage the same action (experiment) u^*, a maximizer of I_u.

Proof: By the Bayes rule, the error probability is the weighed sum of the two conditional error probabilities:

$$P(e|n,\pi) = p(H_1)P(L_n \geq \alpha_n|n,\pi,H_1) + p(H_2)P(L_n < \alpha_n|n,\pi,H_2) . \qquad (6.6)$$

We proceed to bound the first term on the right-hand side. Define

$$X_k = \log \frac{f_2(Y_k|u_k)}{f_1(Y_k|u_k)} ,$$

so that $\overline{X}_n$ (the corresponding sample mean) coincides with L_n. Λ_u, the logarithmic m.g.f. corresponding to X_k conditioned on $u_k = u$, was specified in (6.5). Before applying the lower bounds of Section 5, we note that $E_u(X) \leq 0$ (by Jensen's inequality), and further that $\Lambda_u^*(x)$ (the Legendre transform of Λ_u) is finite and continuous at $x = 0$. To establish the latter it is enough to show that (the convex function) $\Lambda_u^*(x)$ is finite in some neighborhood of zero. Indeed, if $\mu_1^u \neq \mu_2^u$ (as assumed) then X obtains both negative and positive values, namely $P_u(X > 0) \geq \epsilon$ and $P_u(X < 0) \geq \epsilon$ for some $\epsilon > 0$. This implies that $\Lambda_u(\lambda) \leq \log(\epsilon e^{|\lambda|\epsilon})$, so that for every $x \in [-\epsilon, \epsilon]$:

$$\Lambda_u^*(x) \leq \sup_\lambda \{|x||\lambda| - \log \epsilon - |\lambda|\epsilon\} \leq -\log \epsilon < \infty .$$

By Assumption C and the remark at the end of the last section, the lower bound of Theorem 5.1 is applicable. Recalling that $\alpha_n \to 0$ and that $\Lambda_u^*(x)$ is non-decreasing for $x \geq 0$ (Lemma 3.1) and continuous at $x = 0$ (as just established), this lower bound implies

$$\liminf_{n\to\infty} \inf_\pi \frac{1}{n} \log P(L_n \geq \alpha_n|n,\pi,H_1) = -\max_u \Lambda_u^*(0) = -I(\mu_1^u, \mu_2^u) , \qquad (6.7)$$

where the last equality follows by direct calculation or by comparison with (6.1)–(6.2). Turning to the second term in (6.6), a symmetric argument yields

$$\liminf_{n\to\infty} \inf_\pi \frac{1}{n} \log P(L_n < \alpha_n|n,\pi,H_1) = -\max_u I(\mu_1^u, \mu_2^u) ,$$

which together with (6.6) and (6.7) yields the the lower bound part of the theorem. Furthermore, observe that equality in the above lower bounds for the two conditional errors may be attained by the *same* sampling policy, e.g. the stationary policy $u_n \equiv u^*$. It follows that this policy attains the lower bound also for the unconditional error, which completes the proof. $\square$

The result has two implications. First, it supplies an asymptotic estimate for the error probability under the optimal sampling policy (for fixed n). It also implies that the simple policy of repeating a single "best" experiment is asymptotically comparable to any other, in terms of the exponential decay rates.

As noted in the proof, a key fact is that two conditional errors are simultaneously minimized by the same sampling policy. This is a fortunate consequence of using the optimal likelihood ratio test, whose optimality depends on balancing the two conditional errors.

6.2 Maximal Error Probabilities

We next consider sampling policies which *maximize* the probability of error. This problem may be of interest when the choice of experiments is not controlled by the statistical decision maker (who is still interested in minimizing errors, and continues to use the likelihood ratio test). Indeed, it may be "controlled" by Nature, in which case we are performing worst-case error analysis; or it may be controlled by another decision maker with opposing interests (e.g., an airplane seeking to avoid detection or identification by a radar). Depending on the situation, it may or may not be reasonable to allow here for dependence of the sampling policy on the true hypothesis; interestingly enough, the following result holds in both cases.

Theorem 6.2 *Assume C. Then*

$$\limsup_{n\to\infty} \sup_{\pi} P(\mathrm{e}|n,\pi) \leq - \min_{q\in\Delta(\mathcal{U})} \Lambda_q^*(0)\,,$$

and equality is obtained by the (possibly randomized) stationary sampling policy q, where q is a minimizer of $\Lambda_q^*(0)$.

The proof is similar to the previous one, and follows from the upper bound of Proposition 4.1. As before, this result implies that the extremal decay rate (corresponding here to maximized error probabilities) can be obtained by simple sampling policies which do not make use of observations from past experiments. However, the use of more than one type of experiment may now be required. This may be performed by a stationary randomized policy, or by the deterministic alternative outlined in Lemma 4.1.

6.3 Multiple Hypotheses

We consider briefly the possible generalization to the multiple-hypotheses case. The problem is the same, except that there are $J > 2$ (simple) hypotheses denoted $\{H_1, , \ldots, H_J\}$. Thus, each

hypothesis H_j corresponds to a collection of probability densities $\{f_j(\cdot|u), u \in \mathcal{U}\}$. Assume for simplicity that the prior probabilities are equal. Given the data sequence $\{u_1, Y_1, \ldots, u_n, Y_n\}$, the optimal decision rule is to choose the hypothesis j^* which maximizes the likelihood function. This may be written, similarly to the two-hypothesis case, as: $j^* = \arg \max_j L_n(j)$, where

$$L_n(j) = \frac{1}{n} \sum_{k=1}^{n} \log \frac{f_j(Y_k|u_k)}{f_1(Y_k|u_k)}.$$

(Note that $L_n(1) \equiv 0$.) Let us consider the probability of error given that H_1 is the true hypothesis, and the sampling policy is π. Denoting by $\mathbb{P}_1^\pi$ the corresponding probability measure, we have

$$P(e|n, \pi, H_1) = \mathbb{P}_1^\pi \{ \max_{j=2\ldots J} L_n(j) > 0 \} = \mathbb{P}_1^\pi \{ \vec{L}_n(j) \in \Gamma_0 \} ,$$

where $\vec{L}_n = (L_n(2), \ldots, L_n(J))$ and $\Gamma_0 = \{x \in \mathbb{R}^{J-1} : \max_j x_j > 0\}$. As in the two-hypotheses case, one may consider $\vec{L}_n$ as the sample mean of a (vector-valued) controlled i.i.d. process, and attempt to apply the general bounds from the previous sections. In the case of the upper bound (maximum error probabilities), this approach indeed yields the analogue of Theorem 6.2. However, when the lower bound is considered (minimal error probabilities), a crucial observation is that the set Γ_0 is no longer convex for $J > 2$. Thus Theorem 5.1 is not applicable, and the gap noted in Section 5 between stationary and non-stationary (history-dependent) sampling policies exists. In other words, in the multiple-hypotheses case, planning future experiments according to the results of previous ones may significantly improve the asymptotic error probabilities. Computation of the decay rate for the minimized error probabilities remains in this case an interesting open problem, which is closely related to the non-convex lower bound problem of Section 5.

7 Concluding Remarks

We mention here some problems for further research. Investigation of the lower bound for non-convex sets (Section 5) has not been exhausted here, and tight bounds for this case (as well as non-stationary policies which attain them) are yet to be derived. The problem considered in this paper, namely extremal large deviations for controlled processed, may be generalized in several directions. These include higher level large deviations (e.g., large deviations of the empirical distribution), as well as more general controlled processes (e.g. controlled Markov processes).

References

[1] Bertsekas, D.P. (1976). *Dynamic Programming and Stochastic Control,* Academic Press, New York.

[2] Bucklew, J.A. (1990). *Large Deviations Techniques in Decision, Simulation and Estimation,* Wiley, New York.

[3] Chernoff, H. (1952). "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Statist.* **23**, pp. 493–507.

[4] Chernoff, H. (1972). *Sequential Analysis and Optimal Design,* Regional Conference Series in Applied Mathematics 8, SIAM, Philadelphia.

[5] de Acosta, A. (1985). "Upper bounds for large deviations of dependent random vectors," *Z. Wharsch. verw. Gebiete* **69**, pp. 551–565.

[6] Deuschel, J.-D. and Stroock, D.W. (1989). *Large Deviations,* Academic Press, San Diego.

[7] Dembo, A. and Zeitouni, O. (1992). *Large Deviations and Applications,* Jones and Bartlett, Boston, MA.

[8] Ellis, R.S. (1984). "Large deviations for a general class of random vectors," *Ann. Prob.* **12**, pp. 1–12.

[9] RockaFellar, R.T. (1970). *Convex Analysis,* Princeton University Press, Princeton, NJ.

[10] Ross, S.M. (1983). *Introduction to Stochastic Dynamic Programming,* Academic Press, New York.

[11] Shwartz, A. and Weiss, A. (1992). *Large Deviations and their Applications to Computer and Communications Systems,* draft.