

- [16] J.D. Gibson and J.L. Melsa, *Introduction to Nonparametric Decision with Applications*. New York: Academic Press, 1975 (Sec.2.7).
 [17] J.W. Carlyle, "Nonparametric methods in detection theory," in *Communication Theory*, A.V. Balakrishnan, Ed. New York: McGraw-Hill, 1986, pp. 293-319.

A Contribution to the Theory of Chernoff Bounds

James A. Bucklew, *Member, IEEE*, and
 John S. Sadowsky, *Member, IEEE*

Abstract—Some extensions that allow new estimating and bounding techniques for certain sequences of random variables controlled by a large deviation principle are given. These results can be thought of as generalizations and extensions of the Chernoff Bound used in communications theory.

Index Terms—Chernoff bounds, large deviation theory.

I. INTRODUCTION

The problem of calculating the probability of a small probability event is a common one in signal detection. In hypothesis testing problems with i.i.d. or Markov alternatives, the problem of calculating the type I and II errors becomes the probability that an average of functionals of the observations exceeds a fixed threshold. For large numbers of observations, this probability will typically be small. Direct calculation of this test statistic's probability law is usually hopelessly complex thereby engendering the need of various bounding/estimating techniques for such situations.

Chernoff's Bound is the most frequently resorted to communications theory technique for these sorts of situations. Chernoff's original paper [3] was concerned with the asymptotic discernibility of two i.i.d. sequences of random variables. He showed that the logarithmic rate of probability of error to zero was given by the now-called Chernoff entropy:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P(\text{error given first } n \text{ observations}) = \inf_{\alpha} \log(H_{\alpha}),$$

where $H(\alpha) = \int (dp/dw)^{\alpha} (dq/dw)^{1-\alpha} dw$, where p, q are the distributions of one element of the sequence under the two hypotheses and w is any measure dominating both of them. This result is considered to be one of the first large deviation theory theorems. As a consequence of the proof (in fact by application of Markov's inequality), one can show

$$P(\text{error given first } n \text{ observations}) \leq (\inf_{\alpha} H(\alpha))^n.$$

It is this last result that is typically known as Chernoff's Bound. However, this result is known to be somewhat pessimistic. In fact, it is known in the i.i.d. setting (under some mild technical conditions

Manuscript received March 26, 1992. This work was supported in part by NSF Grant NCR-9104823.

J. A. Bucklew is with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, WI 53706.

J. S. Sadowsky is with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47906.

IEEE Log Number 9202999.

[1]), that the following limit holds:

$$\lim_{n \rightarrow \infty} P(\text{error given first } n \text{ observations}) \sqrt{n} / (\inf_{\alpha} H(\alpha))^n = c > 0.$$

Chernoff's result, with hindsight, can rightly be regarded as a particular application to the case of likelihood ratio tests of a large deviations theorem (due to Cramér). Cramér's theorem dealt with tail probabilities of sums of i.i.d. random variables. (The likelihood ratio in the i.i.d. setting can be expressed as such a sum and hence establish Chernoff's result.) Cramér's work was generalized to the Markov case by Miller [10] and others.

More recently a very general large deviation theorem was proved for sequences of random variables by considering only properties of the moment generating function sequence [5], [8]. In this correspondence, we consider the Chernoff/Cramér Bound and its "square root n " refinements in the setting of these new theorems. In keeping with these new methodologies, all our technical conditions are specified in terms of the moment generating (or characteristic) functions of the sequence of interest. We present some novel examples that illustrate the use of these new techniques.

II. PRELIMINARIES

Lemma 1: Suppose we have a sequence of probability measures μ_n converging weakly to a probability measure μ and a sequence of real valued functions f_n converging uniformly to f , a bounded continuous function. Then,

$$\int f_n d\mu_n \rightarrow \int f d\mu.$$

Proof: By the uniform convergence and the finiteness of the probability measures we have that $\int |f_n(x) - f(x)| d\mu_n(x) \rightarrow 0$. But this implies that $|\int f_n(x) d\mu_n(x) - \int f(x) d\mu_n(x)| \rightarrow 0$. Since $f(x)$ is bounded and continuous and $\mu_n \rightarrow \mu$ weakly, we have by the definition of weak convergence (see e.g., [2]) $\int f(x) d\mu_n(x) \rightarrow \int f(x) d\mu$. \square

Let Y_n be a sequence of \mathbf{R} -valued random variables and define $M_n(\theta) = E\{\exp(\theta Y_n)\}$. Assume the following.

A1) The limit $\lim_{n \rightarrow \infty} \frac{1}{n} \log M_n(\theta) = c(\theta)$ for all $\theta \in \mathbf{R}$ (allowing $+\infty$ as the limit in the case of divergence to $+\infty$ or if $M_n(\theta) = +\infty$ for all large n),

A2) $C = \{\theta : c(\theta) < \infty\}^0 \neq \emptyset$ (where E^0 denotes the interior of a set E).

It is easy to see that $c(\theta)$ is a convex function because it is the limit of convex functions. It follows that the derivative $c'(\theta)$ exists for all but at most countably many points in C . The following is a version of the Gärtner-Ellis theorem.¹

Theorem 1 (Gärtner-Ellis): Assume A1)-A2) and that there exists a $\theta_a \in C \cap (0, \infty)$ such that $c'(\theta_a)$ exists and $c'(\theta_a) = a$. Then,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P\left(\frac{Y_n}{n} \geq a\right) = -\{a\theta_a - c(\theta_a)\}.$$

In the i.i.d. setting, the theorem and its conditions are somewhat less mysterious. Let $\{X_i\}_{i=1}^{\infty}$ be a collection of i.i.d. random variables with moment generating function $M(\theta)$. Define $Y_n = \sum_{i=1}^n X_i$. The $M_n(\theta) = M(\theta)^n$, or $c(\theta) = \log(M(\theta))$ which is, in fact, analytic on the open domain C .

¹This particular version follows from the remark about display (31) in [5].

By convexity, notice that the assumption that $\theta_n > 0$ is equivalent to $c'(0) < a$. In the i.i.d. sum case, this is just $E\{X_i\} = c'(0) < a$ which by the law of large numbers is necessary for $P\left(\frac{Y_n}{n} \geq a\right) \rightarrow 0$.

We will hereafter consider only the case $a = 0$, which by the transformation $Y'_n - an = Y_n$ is no loss of generality and somewhat simplifies the notation in the following proofs.

III. DEVELOPMENT

Let $Y_n, M_n(\theta)$ be as previously stated. We will always assume conditions A1)–A2) and $c'(0) < 0$. We further assume that $c(\theta_0) = 0$ for some $\theta_0 > 0$, and that $c''(\theta_0) = \sigma^2 > 0$. We are interested in precise estimates of the probability $P(Y_n/n \geq 0)$. Let $F_n(x)$ denote the distribution of Y_n . Define

$$dF_n^{(\theta)}(x) = \frac{\exp(\theta x) dF_n(x)}{\int \exp(\theta y) dF_n(y)} = \frac{\exp(\theta x) dF_n(x)}{M_n(\theta)}$$

as the “twisted” or “tilted” distribution. Let $Y_n^{(\theta)}$ be a random variable having $F_n^{(\theta)}$ as its probability distribution. Let $H_n(x)$ denote the distribution function of $Y_n^{(\theta_0)}/\sqrt{n}\sigma$.

Then,

$$\begin{aligned} P(Y_n/n \geq 0) &= \int_0^\infty dF_n(x) \\ &= M_n(\theta_0) \int_0^\infty \exp(-\theta_0 x) dF_n^{(\theta_0)}(x) \\ &= M_n(\theta_0) \int_0^\infty \exp(-\sqrt{n}\theta_0 x \sigma) dH_n(x). \end{aligned} \quad (1)$$

We now need to consider the asymptotics of the integral

$$\int_0^\infty \exp(-\sqrt{n}\alpha x) dH_n(x) \quad (2)$$

(for $\alpha > 0$) which may be transformed by integration by parts to

$$\sqrt{n}\alpha \int_0^\infty \exp(-\sqrt{n}\alpha x) [H_n(x) - H_n(0)] dx. \quad (3)$$

Remark 1: In (2), we may upper bound $\exp(-\sqrt{n}\alpha x)$ by 1 to find that an upper bound to the integral is 1. This is the analog to the Chernoff Bound for i.i.d. sums.

Let us define the sequence of complex functions

$$\phi_n(\omega) = \frac{M_n\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0\right)}{M_n(\theta_0)}.$$

We may now state one of our principal lemmas.

Lemma 2: Suppose

$$\phi_n(\omega) \xrightarrow{\mathcal{L}_1} \exp\left(-\frac{1}{2}\omega^2\right).$$

Then,

$$\sqrt{n} \int_0^\infty \exp(-\sqrt{n}\alpha x) dH_n(x) \rightarrow \frac{1}{\sqrt{2\pi}\alpha}. \quad (4)$$

Proof: We note that $\phi_n(\omega)$ is the characteristic function of $Y_n^{(\theta_0)}/\sqrt{n}\sigma$. The assumed \mathcal{L}_1 convergence implies that $\phi_n(\omega)$ is \mathcal{L}_1 (for large enough n) and hence that $Y_n^{(\theta_0)}/\sqrt{n}\sigma$ has a density

function $h_n(x)$ ([4, p. 155]). The Fourier inversion formula then implies

$$\left| h_n(x) - \frac{\exp\left(-\frac{x^2}{2}\right)}{\sqrt{2}} \right| \leq \frac{1}{2\pi} \int_{-\infty}^\infty \left| \phi_n(\omega) - \exp\left(-\frac{\omega^2}{2}\right) \right| d\omega.$$

The \mathcal{L}_1 convergence of $\phi_n(\omega)$ to $\exp(-\omega^2/2)$ immediately implies uniform convergence of $h_n(x)$ to $\exp(-x^2/2)/\sqrt{2\pi}$.

We now define the sequence of probability measures $\mu_n(dx) = \sqrt{n}\alpha \exp(-\sqrt{n}\alpha x) dx$. It is easy to check that μ_n is a probability measure for every n and that the sequence converges weakly to $\mu(dx) = \delta(x) dx$, point mass at the origin. Therefore, in the integral, we may replace $dH_n(x)$ by $h_n(x) dx$ and by invoking Lemma 1, we have completed the proof of the lemma. \square

Theorem 2: If $\phi_n(\omega) \xrightarrow{\mathcal{L}_1} \exp(-\omega^2/2)$ then,

$$\lim_{n \rightarrow \infty} P\left(\frac{Y_n}{n} \geq 0\right) \sqrt{2\pi n} \sigma \theta_0 / M_n(\theta_0) = 1.$$

Proof: The theorem follows immediately from (1) and the previous lemma. (Note that with the stated assumptions $\theta_0 \sigma \neq 0$.) \square

Corollary 1: Assume $\phi_n(\omega) \rightarrow \exp(-\omega^2/2)$ where the convergence is pointwise. Suppose $\phi_n(\omega)$ is \mathcal{L}_1 (for all n sufficiently large) or alternatively suppose $Y_n^{(\theta_0)}/\sqrt{n}\sigma$ has bounded density functions $h_n(\cdot)$ (for all n sufficiently large). Then,

$$\lim_{n \rightarrow \infty} P\left(\frac{Y_n}{n} \geq 0\right) \sqrt{2\pi n} \sigma \theta_0 / M_n(\theta_0) = 1.$$

Proof: For the first alternative, note that characteristic functions are bounded. Pointwise convergence and the dominated convergence theorem then imply the \mathcal{L}_1 convergence needed as a condition of Theorem 2.

For the second alternative, note that if a probability density h_n is bounded, then its characteristic function ϕ_n is \mathcal{L}_1 . \square

Remark 2: It is known from the proof of the Gärtner–Ellis theorem that with the stated assumptions, $Y_n^{(\theta_0)}/n \rightarrow 0$, almost surely. This fact and the fact that $M_n(\theta) \approx c(\theta)^n$ means that $Y_n^{(\theta_0)}/n$ is “behaving” as if it were the sample average of some random variables converging to a mean value. Hence the idea that by scaling with $1/\sqrt{n}$ instead of $1/n$, we may be able to have central limit theorem type behavior. The proof of Lemma 2 hinges on the fact that $Y_n^{(\theta_0)}/\sqrt{n}\sigma$ converges in distribution to a standard normal.

Remark 3: As a consequence of the hypothesis of \mathcal{L}_1 convergence, we rule out “lattice type” random variables. From previous results for the i.i.d. [1] and Markov [10] cases, we know that these must be treated separately and have (in general) a limit dependent upon the lattice spacing. We consider this case in more detail in the following theorem.

Remark 4: It may be suspected that \mathcal{L}_1 convergence of the characteristic functions is a rather stringent condition. Perhaps so, but some further assumption on the moment generating sequence is required other than the two conditions used to invoke the Gärtner–Ellis theorem. For example, suppose $Y_n = \sum_{i=1}^n X_i$, where the $\{X_i\}$ are independent Gaussian random variables with variance one. Suppose that $E\{X_i\} = m_i$, where the $\{m_i\}$ sequence is chosen so that $\sum_{i=1}^n m_i/n \rightarrow 0$ and $\sum_{i=1}^n m_i/\sqrt{n} \rightarrow \infty$. It is easy to check that $\lim_{n \rightarrow \infty} 1/n \log M_n(\theta) = c(\theta) = \theta^2/2$, which satisfies A1)–A2) and the other conditions for the Gärtner–Ellis theorem hold. However, any fixed exponential shift merely corresponds to a fixed mean shift in this Gaussian setting. Hence, since $\sum m_i/\sqrt{n} \rightarrow \infty$, $Y_n^{(\theta_0)}/\sqrt{n}\sigma$ does not converge in distribution for any θ_0 .

We now consider the lattice random variable case. Suppose the $\{Y_n\}$ have support on the set $\{b + jd; j \in \mathbb{N}, d \in (0, \infty), 0 \leq b < d\}$. Let d be the largest number with the property. Then d is called the *span* or *lattice spacing*.

Lemma 3: Suppose

$$\phi_n(\omega) 1_{\{-\sqrt{n}\pi\sigma/d, \sqrt{n}\pi\sigma/d\}}(\omega) \xrightarrow{\mathcal{L}_1} \exp\left(-\frac{1}{2}\omega^2\right).$$

Then, for $\alpha > 0$,

$$\sqrt{n} \int_{0^-}^{\infty} \exp(-\sqrt{n}\alpha x) dH_n(x) \rightarrow \frac{d \exp(-\frac{\alpha b}{\sigma})}{\sqrt{2\pi}\sigma(1 - \exp(-\frac{\alpha d}{\sigma}))}.$$

Proof: Note that the $\{Y_n^{(\theta)}\}$ have support on the same lattice as the $\{Y_n\}$. Thus, $\phi_n(\omega)$ is periodic with fundamental period $2\pi\sigma\sqrt{n}/d$. Now, let us consider the expression in (3) multiplied by \sqrt{n} :

$$\begin{aligned} & n\alpha \int_{0^-}^{\infty} \exp(-\sqrt{n}\alpha x) [H_n(x) - H_n(0^-)] dx \\ &= n\alpha \sum_{k=0}^{\infty} \int_{[kd+b]/\sqrt{n}\sigma}^{[(k+1)d+b]/\sqrt{n}\sigma} \exp(-\sqrt{n}\alpha x) dx \\ & \quad \cdot \left[H_n\left(\left[\frac{(k+1)d}{\sqrt{n}\sigma}\right]^{-}\right) - H_n(0^-) \right] \\ &= \sum_{k=0}^{\infty} \sqrt{n} \left[H_n\left(\left[\frac{(k+1)d+b}{\sqrt{n}\sigma}\right]^{-}\right) - H_n(0^-) \right] \\ & \quad \left[\exp\left(-\frac{\alpha}{\sigma}[kd+b]\right) - \exp\left(-\frac{\alpha}{\sigma}[(k+1)d+b]\right) \right]. \end{aligned}$$

Define $t_n(x) = P(Y_n^{(\theta_0)} / (\sqrt{n}\sigma) = x)$. Then, from properties of characteristic functions (see e.g., [6, p. 511]),

$$t_n\left(\frac{x}{\sqrt{n}}\right) \frac{\sqrt{n}\sigma}{d} = \frac{1}{2\pi} \int_{-\sqrt{n}\pi\sigma/d}^{\sqrt{n}\pi\sigma/d} \phi_n(\omega) \exp\left(-i\omega \frac{x}{\sqrt{n}}\right) d\omega.$$

Let $n(\cdot)$ denote the density function of a zero-mean, unit variance Gaussian random variable. Hence,

$$\begin{aligned} & \left| t_n\left(\frac{x}{\sqrt{n}}\right) \frac{\sqrt{n}\sigma}{d} - n\left(\frac{x}{\sqrt{n}}\right) \right| \\ & \leq \int_{-\sqrt{n}\pi\sigma/d}^{\sqrt{n}\pi\sigma/d} \left| \phi_n(\omega) - \exp\left(-\frac{1}{2}\omega^2\right) \right| d\omega \\ & \quad + \int_{|w| > \sqrt{n}\pi/d} \exp\left(-\frac{1}{2}\omega^2\right) d\omega. \end{aligned}$$

The assumed \mathcal{L}_1 convergence then implies

$$\begin{aligned} \lim_{n \rightarrow \infty} t_n\left(\frac{x}{\sqrt{n}}\right) \frac{\sqrt{n}\sigma}{d} &= \\ \lim_{n \rightarrow \infty} n\left(\frac{x}{\sqrt{n}}\right) &= n(0) \\ &= \frac{1}{\sqrt{2\pi}} \end{aligned}$$

uniformly in x .

We then note that

$$\begin{aligned} & \sqrt{n} \left[H_n\left(\left[\frac{[(k+1)d+b]}{\sqrt{n}\sigma}\right]^{-}\right) - H_n(0^-) \right] \\ &= \sum_{j=0}^k \sqrt{n} t_n\left(\frac{jd+b}{\sqrt{n}\sigma}\right) \xrightarrow{n \rightarrow \infty} \frac{d(k+1)}{\sigma\sqrt{2\pi}}. \end{aligned}$$

The uniform convergence implies that we may interchange limit with sum in the following expression:

$$\begin{aligned} & \lim_{n \rightarrow \infty} n\alpha \int_{0^-}^{\infty} \exp(-\sqrt{n}\alpha x) [H_n(x) - H_n(0^-)] dx \\ &= \lim_{n \rightarrow \infty} \sqrt{n} \sum_{k=0}^{\infty} \left[H_n\left(\left[\frac{[(k+1)d+b]}{\sqrt{n}\sigma}\right]^{-}\right) - H_n(0^-) \right] \\ & \quad \times \left[\exp\left(-\frac{\alpha}{\sigma}[kd+b]\right) - \exp\left(-\frac{\alpha}{\sigma}[(k+1)d+b]\right) \right] \\ &= \exp\left(-\frac{\alpha b}{\sigma}\right) \sum_{k=0}^{\infty} \frac{d(k+1)}{\sigma\sqrt{2\pi}} \\ & \quad \times \left[\exp\left(-\frac{\alpha}{\sigma}kd\right) - \exp\left(-\frac{\alpha}{\sigma}(k+1)d\right) \right] \\ &= \frac{d \exp(-\frac{\alpha b}{\sigma})}{\sqrt{2\pi}\sigma(1 - \exp(-\frac{\alpha d}{\sigma}))}. \quad \square \quad (5) \end{aligned}$$

Theorem 3: If $\phi_n(\omega) 1_{\{-\sqrt{n}\pi\sigma/d, \sqrt{n}\pi\sigma/d\}} \xrightarrow{\mathcal{L}_1} \exp(-\frac{1}{2}\omega^2)$, then (if $c'(0) < 0$),

$$\lim_{n \rightarrow \infty} P\left(\frac{Y_n}{n} \geq 0\right) \frac{\sqrt{2\pi n}\sigma(1 - \exp(-\theta_0 d))}{d \exp(-\theta_0 b) M_n(\theta_0)} = 1.$$

Proof: The corollary follows immediately from (1) and the preceding lemma. \square

Remark 5: One can deduce from the above derivation that (for $b = 0$)

$$\lim_{n \rightarrow \infty} P\left(\frac{Y_n}{n} > 0\right) \frac{\sqrt{2\pi n}\sigma(1 - \exp(-\theta_0 d))}{d \exp(-\theta_0 b) M_n(\theta_0)} = 1.$$

Remark 6: Following an argument similar to that of ([6, p. 518]), we can show that the previous \mathcal{L}_1 convergence holds for sums of i.i.d. lattice type random variables. In this setting, $Y_n^{(\theta_0)}$ appears as a sum of zero-mean, unit variance, i.i.d. lattice type random variables. Let $\phi(\cdot)$ be the characteristic function of one of the summands. Then,

$$\phi\left(\frac{\omega}{\sqrt{n}}\right) = 1 - \frac{\omega^2}{2n} + o\left(\frac{1}{n}\right). \quad (6)$$

It is easy to verify that $\phi^n\left(\frac{\omega}{\sqrt{n}}\right) \rightarrow \exp(-\omega^2/2)$. Consider the integral,

$$\int_{-\sqrt{n}\pi/d}^{\sqrt{n}\pi/d} \left| \phi^n\left(\frac{\omega}{\sqrt{n}}\right) - \exp\left(-\frac{1}{2}\omega^2\right) \right| d\omega.$$

We split the integration over three regions, $|w| \leq a$, $a \leq |w| \leq \delta\sqrt{n}$, and $\delta\sqrt{n} \leq |w| \leq \sqrt{n}\pi/d$. The integral over the first region goes to zero by the convergence of $\phi^n(\omega)$ and the dominated convergence theorem. Because of (6), it is possible to choose $\delta > 0$ so that $\phi(\omega) \leq \exp(-\omega^2/4)$ for $|\omega| < \delta$, which implies that the integrand over the second region can be made negligible by making a large. In the third region, $\sup\{\phi(\omega/\sqrt{n}) : \delta\sqrt{n} \leq |\omega| \leq \sqrt{n}\pi/d\} = n < 1$.

The contribution of the third integral is then less than

$$\begin{aligned} \eta^{n-1} \int_{\sqrt{n}\delta}^{\sqrt{n}\pi/d} \left| \phi\left(\frac{\omega}{\sqrt{n}}\right) \right| d\omega + \int_{|\omega|>\delta\sqrt{n}} \exp\left(-\frac{1}{2}\omega^2\right) d\omega \\ \leq \eta^{n-1} \sqrt{n} \pi/d \\ + \int_{|\omega|>\delta\sqrt{n}} \exp\left(-\frac{1}{2}\omega^2\right) d\omega \rightarrow 0. \end{aligned}$$

One should be cautioned that in the i.i.d. lattice variable case, our result holds only for $b = 0$. The general lattice case has an n -varying coefficient ([1]) or ([7, p. 192]).

A. Examples

1) *Sum of Squares from a Gaussian Random Process*: Let $\{N_i\}$ be a wide sense stationary Gaussian random process with continuous power spectrum $f(\omega)$, $0 \leq \omega \leq 2\pi$. We suppose $f(\omega) \geq \delta > 0$ for some $\delta \in \mathbf{R}^+$. We also require as a technical condition that $\lim_{\gamma \rightarrow 1/\|f\|_\infty} \int_0^{2\pi} \log(1 - \gamma f(\omega)) d\omega = \infty$. We are interested in the asymptotic of the "sum of squares" of this process, i.e.,

$$P\left(\frac{1}{n} \sum_{i=1}^n N_i^2 \geq \gamma\right),$$

where we suppose $\gamma > E\{N_1^2\}$. Define $Y_n = \sum_{i=1}^n (N_i^2 - \gamma)$. Then,

$$\begin{aligned} M_n(\theta) &= E\{\exp(\theta Y_n)\} \\ &= \exp(-\theta \gamma n) \prod_{i=1}^n \left(\frac{1}{1 - 2\theta \lambda_i^{(n)}} \right)^{\frac{1}{2}}, \end{aligned} \quad (7)$$

where $\{\lambda_i^{(n)}\}$ are the eigenvalues of the covariance matrix of $(N_1, N_2, \dots, N_n)^T$. Therefore, for $\theta < 1/(2\|f\|_\infty)$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log M_n(\theta) &= \\ &= -\theta \gamma - \frac{1}{4\pi} \int_0^{2\pi} \log(1 - 2\theta f(\omega)) d\omega = c(\theta) \end{aligned}$$

and $c(\theta) = \infty$ for θ outside this range. The convergence follows by the Toeplitz distribution theorem [9]. The technical condition implies that $c(\theta) \rightarrow \infty$ as $\theta \rightarrow 1/(2\|f\|_\infty)$. By convexity, it is simple to verify that a solution to the equation $c'(\theta_0) = 0$ exists or

$$\gamma = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(\omega)}{1 - 2\theta_0 f(\omega)} d\omega.$$

(If $f(\omega) = 1$, then $\gamma = 1/(1 - 2\theta_0)$ or $\theta_0 = (1 - 1/\gamma)/2$.) However, in general, this is a transcendental equation and such closed form solutions will not always be available. With some effort, we can check

$$\begin{aligned} \phi_n(\omega) &= \frac{M_n\left(\frac{i\omega}{\sqrt{n}} + \theta_0\right)}{M_n(\theta_0)} \\ &= \exp(-i\omega \gamma \sqrt{n}) \prod_{i=1}^n \left(1 - \frac{2i\omega \lambda_i^{(n)}/\sqrt{n}}{1 - 2\theta_0 \lambda_i^{(n)}} \right)^{-\frac{1}{2}} \\ &\xrightarrow{n \rightarrow \infty} \exp\left(-\frac{c''(\theta_0)\omega^2}{2}\right), \end{aligned} \quad (8)$$

where

$$c''(\theta) = \frac{1}{\pi} \int_0^{2\pi} \frac{f(\omega)^2}{(1 - 2\theta f(\omega))^2} d\omega.$$

$\phi_n(\omega)$ is \mathcal{L}_1 for $n > 2$ certainly. The pointwise convergence, Remark 4, and Theorem 2 imply the following limit:

$$\lim_{n \rightarrow \infty} P\left(\frac{1}{n} \sum_{i=1}^n N_i^2 \geq \gamma\right) \sqrt{2\pi c''(\theta_0)n} / M_n(\theta_0) = 1.$$

2) *Random Sums of i.i.d. Random Variables*: Consider a random sum

$$Y_\lambda = \sum_{i=1}^{N_\lambda} X_i,$$

where $\{X_i\}$ are i.i.d. mean $m < 0$ random variables with moment generating function $M_X(\cdot)$. Take $0 < p < 1$, λ a positive integer and N_λ to be a "shifted" geometric random variable with $P(N_\lambda = n) = (1-p)p^{n-\lambda}$ ($n \geq \lambda$) and independent of the $\{X_i\}$. The characteristic function of Y_λ is

$$M_\lambda(i\omega) = \frac{(1-p)M_X(i\omega)^\lambda}{1 - M_X(i\omega)p}.$$

We then obtain (for values of θ such that $M_X(\theta) < 1/p$),

$$c(\theta) = \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \log(M_\lambda(\theta)) = M_X(\theta),$$

and $c(\theta) = \infty$, otherwise. Hence, $c'(\theta_0) = M_X'(\theta_0) = 0$ (as long as $M_X(\theta_0) < 1/p$) defines θ_0 and $c''(\theta_0) = M_X''(\theta_0) = \sigma^2$ defines σ . The characteristic function of interest is

$$\begin{aligned} \phi_\lambda(\omega) &= \frac{M_\lambda\left(\frac{i\omega}{\sqrt{\lambda}\sigma} + \theta_0\right)}{M_\lambda(\theta_0)} \\ &= \frac{M_X\left(\frac{i\omega}{\sqrt{\lambda}\sigma} + \theta_0\right)^\lambda (1 - M_X(\theta_0)p)}{M_X(\theta_0)^\lambda (1 - M_X\left(\frac{i\omega}{\sqrt{\lambda}\sigma} + \theta_0\right)p)} \\ &\xrightarrow{\lambda \rightarrow \infty} \exp\left(-\frac{1}{2}\omega^2\right). \end{aligned}$$

By the arguments given in Remark 6, one can easily verify that if $M_X\left(\frac{i\omega}{\sqrt{\lambda}\sigma} + \theta_0\right)/(1 - M_X\left(\frac{i\omega}{\sqrt{\lambda}\sigma} + \theta_0\right)p)$ is \mathcal{L}_1 , then the conditions of Theorem 2 hold. If X_1 is a lattice type random variable with span d , then the conditions of Theorem 3 automatically hold.

3) *Conditionally Independent Sums*: We now give an example where $P(Y_n/n \geq 0) \sim O(n^{-3/2}) \exp(c(\theta_0)n)$, instead of the usual situation where the leading coefficient is $O(n^{-1/2})$.

Let Ψ be a continuous random variable with probability density function $p_\Psi(\cdot)$ on $[0,1]$. Take $Y_n = \sum_{i=1}^n X_i$ where the sequence $\{X_i\}$ is a conditionally i.i.d. given $\{\Psi = \psi\}$. Let $M_X(\theta/\psi) = E[\exp(\theta X_i) | \Psi = \psi]$ denote the conditional moment generating function of the summands. Then,

$$M_n(\theta) = \int_0^1 M_X(\theta/\psi)^n p_\Psi(\psi) d\psi.$$

For a fixed θ , the previous integral behaves geometrically like $[\sup_{\psi \in [0,1]} M_X(\theta/\psi)]^n$. The conditions next ensure that the dominant value of ψ is $\psi = 0$ for $\theta \geq 0$, and it will follow that $c(\theta) = \log(M_X(\theta))$. We make the following assumptions:

- $p_\Psi(\cdot)$ is bounded and $\lim_{\psi \rightarrow 0} p_\Psi(\psi) = p_\Psi(0) > 0$.
- There is a unique $\theta_0 \in C \cap (0, \infty)$ such that $c'(\theta_0) = 0$ (where $c(\theta) = \log(M_X(\theta/0))$ for $\theta \geq 0$ and $C = \{\theta : c(\theta) < \infty\}$).
- For each $\theta \in C$, $M_X(i\nu + \theta_0/\psi)$ is continuously jointly differentiable in ν and ψ up to order 3 and the third order derivative in ν is uniformly bounded.
- For each $\theta \in C \cap (0, \infty)$, there exists a constant $\bar{\tau}(\theta) > 0$ such that $\frac{M_X(\theta/\psi)}{M_X(\theta/0)} \leq 1 - \bar{\tau}(\theta)\psi$ for all $\psi \in [0, 1]$.

- e) Define $\mu(\psi) = \frac{\partial}{\partial \theta} \log(M_X(\theta_0|\psi))$, $\sigma^2(\psi) = \frac{\partial^2}{\partial \theta^2} \log(M_X(\theta_0|\psi))$ (define $\sigma^2 \doteq \sigma^2(0)$) and $\phi_X(\nu/\psi) = \frac{M_X(i\nu + \theta_0|\psi)}{M_X(\theta_0|\psi)}$, and assume that

$$\phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \psi\right)^n \mathbf{1}_{\{-\sqrt{n}\pi\sigma/d, \sqrt{n}\pi\sigma/d\}}(\omega) - \exp\left(-\frac{i\sqrt{n}\mu(\psi)\omega}{\sigma} - \frac{\sigma^2(\psi)\omega^2}{2\sigma^2}\right) \rightarrow 0$$

in \mathcal{L}_1 uniformly over $\psi \in [0, 1]$ (omit the indicator function in the nonlattice case).

Now, consider

$$\begin{aligned} \frac{nM_n(\theta)}{M_X(\theta|0)^n} &= n \int_0^1 \left(\frac{M_X(\theta|\psi)}{M_X(\theta|0)}\right)^n p_\Psi(\psi) d\psi \\ &= \int_0^n \left(\frac{M_X(\theta|\frac{\zeta}{n})}{M_X(\theta|0)}\right)^n p_\Psi\left(\frac{\zeta}{n}\right) d\zeta. \end{aligned}$$

Define

$$\tau(\theta) = -\frac{\frac{\partial}{\partial \psi} M_X(\theta|0)}{M_X(\theta|0)}$$

and note that $\tau(\theta) \geq \bar{\tau}(\theta) > 0$ by assumption d). For fixed $\zeta \geq 0$, and $n \geq \zeta$, we have

$$\left(\frac{M_X(\theta|\frac{\zeta}{n})}{M_X(\theta|0)}\right)^n = \left(1 - \frac{\tau(\theta)\zeta}{n} + o\left(\frac{\zeta}{n}\right)\right)^n \rightarrow \exp(-\tau(\theta)\zeta).$$

Using $\bar{p}_\Psi = \sup_{\psi \in [0,1]} p_\Psi(\psi)$ and assumption d), we have the integrable dominating function

$$\begin{aligned} \left(\frac{M_X(\theta|\frac{\zeta}{n})}{M_X(\theta|0)}\right)^n p_\Psi\left(\frac{\zeta}{n}\right) &\leq \bar{p}_\Psi \left(1 - \frac{\bar{\tau}(\theta)\zeta}{n}\right)^n \\ &\leq \bar{p}_\Psi \exp(-\bar{\tau}(\theta)\zeta). \end{aligned}$$

Thus, by the dominated convergence theorem,

$$\frac{nM_n(\theta)}{M_X(\theta|0)^n} \rightarrow p_\Psi(0) \int_0^\infty \exp(-\tau(\theta)\zeta) d\zeta = \frac{p_\Psi(0)}{\tau(\theta)},$$

and hence, $M_n(\theta) \sim \frac{p_\Psi(0)}{\tau(\theta)n} M_X(\theta|0)^n$ for $\theta \in C \cap (0, \infty)$. Clearly now, $c(\theta) = \log(M_X(\theta|0))$ for $\theta \geq 0$ as claimed. So, if we can verify the \mathcal{L}_1 convergence condition of Theorem 2 we will have

$$P(Y_n/n \geq 0) \sim \frac{p_\Psi(0) d e^{\theta_0 b}}{\sqrt{2\pi} \sigma \tau(\theta_0) (1 - e^{-\theta_0 d})} n^{-3/2} \exp(c(\theta_0)n), \quad (9)$$

for the lattice case.

Now, consider what is needed for the \mathcal{L}_1 convergence condition. We have

$$\phi_n(\omega) = \frac{M_n\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0\right)}{M_n(\theta_0)} = \frac{nM_n\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0\right)/M_X(\theta_0|0)^n}{nM_n(\theta_0)/M_X(\theta_0|0)^n}.$$

We have just shown that the denominator in the last expression converges to $p_\Psi(0)/\tau(\theta_0)$. Hence, we need to show that the numerator $nM_n\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0\right)/M_X(\theta_0|0)^n \rightarrow (p_\Psi(0)/\tau(\theta_0)) \exp(-\omega^2/2)$ in

\mathcal{L}_1 . We have

$$\begin{aligned} &\frac{nM_n\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0\right)}{M_X(\theta_0|0)^n} \\ &= \frac{n \int_0^1 M_X\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0|\psi\right)^n p_\Psi(\psi) d\psi}{M_X(\theta_0|0)^n} \\ &= n \int_0^1 \left(\frac{M_X(\theta_0|\psi)}{M_X(\theta_0|0)}\right)^n \left(\frac{M_X\left(\frac{i\omega}{\sqrt{n}\sigma} + \theta_0|\psi\right)}{M_X(\theta_0|\psi)}\right)^n p_\Psi(\psi) d\psi \\ &= \int_0^n \left(\frac{M_X(\theta_0|\frac{\zeta}{n})}{M_X(\theta_0|0)}\right)^n \phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \frac{\zeta}{n}\right)^n p_\Psi\left(\frac{\zeta}{n}\right) d\zeta \\ &= \int_0^n \left(\frac{M_X(\theta_0|\frac{\zeta}{n})}{M_X(\theta_0|0)}\right)^n p_\Psi\left(\frac{\zeta}{n}\right) d\zeta e^{-\omega^2/2} \\ &\quad + \int_0^n \left(\frac{M_X(\theta_0|\frac{\zeta}{n})}{M_X(\theta_0|0)}\right)^n \left[\phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \frac{\zeta}{n}\right)^n - e^{-\omega^2/2}\right] \\ &\quad \times p_\Psi\left(\frac{\zeta}{n}\right) d\zeta. \end{aligned}$$

From a previous argument, we note that the integral in the first term of the last expression converges to $p_\Psi(0)/\tau(\theta_0)$. It is now sufficient to show that the second term vanishes in \mathcal{L}_1 . Using $(M_X(\theta_0|\frac{\zeta}{n})/M_X(\theta_0|0))^n \leq \exp(-\bar{\tau}(\theta_0)\zeta)$, we have

$$\begin{aligned} &\int_0^n \left(\frac{M_X(\theta_0|\frac{\zeta}{n})}{M_X(\theta_0|0)}\right)^n \left[\phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \frac{\zeta}{n}\right)^n - e^{-\omega^2/2}\right] p_\Psi\left(\frac{\zeta}{n}\right) d\zeta \\ &\leq \bar{p}_\Psi \int_0^n \exp(-\bar{\tau}(\theta_0)\zeta) \left|\phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \frac{\zeta}{n}\right)^n - e^{-\omega^2/2}\right| d\zeta. \end{aligned}$$

Finally, we use the triangle inequality to write

$$\begin{aligned} &\left|\phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \frac{\zeta}{n}\right)^n - e^{-\omega^2/2}\right| \leq \\ &\left|\phi_X\left(\frac{\omega}{\sqrt{n}\sigma} \mid \frac{\zeta}{n}\right)^n - \exp\left(-\frac{i\sqrt{n}\mu\left(\frac{\zeta}{n}\right)\omega}{\sigma} - \frac{\sigma^2\left(\frac{\zeta}{n}\right)\omega^2}{2\sigma^2}\right)\right| \\ &\quad + \left|\exp\left(-\frac{i\sqrt{n}\mu\left(\frac{\zeta}{n}\right)\omega}{\sigma} - \frac{\sigma^2\left(\frac{\zeta}{n}\right)\omega^2}{2\sigma^2}\right) - e^{-\omega^2/2}\right|. \end{aligned}$$

By assumption e), the first term vanishes in \mathcal{L}_1 uniformly in ζ . (Evidently, the uniform convergence assumption can be weakened due to the $\exp(-\bar{\tau}(\theta_0)\zeta)$ weighting factor.) Using assumption c), we can show that $|\mu(\psi)| \leq K_1\psi$ and $|\sigma^2(\psi) - \sigma^2| \leq K_2\psi$ for some constants $K_1, K_2 < \infty$, and from this it follows that the integral of the second absolute difference against $\exp(-\bar{\tau}(\theta_0)\zeta)$ also vanishes in \mathcal{L}_1 .

We remark that an interesting feature of this example is that $c(\theta)$ is actually not differentiable at $\theta = 0$. One finds that $c(\theta) = \log(M_X(\theta|1))$ for $\theta \leq 0$, $c^-(0) =$ the left derivative $= E[X_i|\Psi = 1]$ and $c^+(0) =$ the right derivative $= E[X_i|\Psi = 0] < 0$.

We also point out that, while some of our assumptions might be weakened, Ψ must have a continuous component. If the distribution of Ψ has a discrete probability mass at $\psi = 0$, then $P((Y_n/n \geq 0) | \Psi = 0) \sim O(n^{-1/2}) \exp(c(\theta_0)n)$ so we lose the $O(n^{-3/2})$ behavior.

The case that $\tau(\theta_0) = 0$ could also be analyzed using the second-order Taylor series of $M_X(\theta|\frac{\zeta}{n})$. The conclusion is that $M_n(\theta) = O(n^{-1/2}) \exp(c(\theta)n)$, and hence, $P((Y_n/n \geq 0) | \Psi = 0) \sim O(n^{-1}) \exp(c(\theta_0)n)$.

TABLE I

p_0	b	n	$P(Y_n/n \geq 0)$	Equation (9)	$\exp(c(\theta_0)n)$
0.1	0.1	5	2.86×10^{-2}	1.41×10^{-2}	7.78×10^{-2}
		25	1.26×10^{-7}	4.59×10^{-8}	2.84×10^{-6}
		50	4.33×10^{-14}	4.62×10^{-14}	8.08×10^{-12}
0.3	0.2	25	7.52×10^{-3}	6.63×10^{-3}	1.13×10^{-1}
		50	2.11×10^{-4}	2.65×10^{-4}	1.28×10^{-2}
		100	1.06×10^{-6}	1.21×10^{-6}	1.64×10^{-4}
0.01	0.005	4	3.46×10^{-4}	6.38×10^{-4}	1.57×10^{-3}
		8	2.64×10^{-7}	3.54×10^{-7}	2.46×10^{-6}
		16	2.68×10^{-13}	3.08×10^{-13}	6.05×10^{-12}

For a numerical example, suppose that $X_i = \pm 1$ with $P(X_i = +1) = p_0 - b\psi$ with $p_0 < 1/2$ and $b \leq p_0$, and Ψ is a uniform random variable on $[0,1]$. After a little work it can be shown that

$$P(Y_n/n \geq 0) = \frac{1}{b} \sum_{k=\lfloor n/2 \rfloor}^n \binom{n}{k} [\beta_{p_0}(k+1, n-k+1) - \beta_{p_0-b}(k+1, n-k+1)],$$

where $\beta_x(a, b)$ is the incomplete beta function. This formula can be numerically evaluated using the BETDF subroutine from the IMSL library. We also find that $\theta_0 = \frac{1}{2} \log((1-p_0)/p_0)$, $c(\theta_0) = \log(2\sqrt{p_0(1-p_0)})$, $\sigma^2 = 1$ and $\tau(\theta_0) = b(1-2p_0)/[2p_0(1-p_0)]$. Table I compares some numerical values of the exact value $P(Y_n/n \geq 0)$, the asymptotically sharp approximation given in (9), and the crude exponential approximation $\exp(c(\theta_0)n)$. A similar comparison is carried out in the i.i.d. setting in ([7, pp. 129–131]).

A practical situation where conditional i.i.d. sums arise is in the analysis of the correlator receiver for direct sequence spread spectrum, multiple access communications systems. In this application, the random phases and timing delays of interfering spread spectrum signals play the role of the "nuisance variable" Ψ . A more detailed large deviations analysis of this receiver is given in [11].

IV. DISCUSSION

We note that finding the asymptotics of $M_n(\theta_0)$ can in of itself be a nontrivial problem. Our philosophy has been to assume that knowledge of the moment generating function sequence is complete. In the setting of the first example and in more general cases of the third, this can be a nontrivial task, even though the logarithmic behavior is known.

REFERENCES

- [1] R. Bahadur and R. R. Rao, "On deviations of the sample mean," *Ann. Math. Statist.*, vol. 31, pp. 1015–1027, 1960.
- [2] P. Billingsley, *Convergence of Probability Measures*. Wiley, New York, 1968.
- [3] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Statist.*, vol. 23, pp. 493–507, 1952.
- [4] K. L. Chung, *A Course in Probability Theory*, 2nd ed. New York: Academic Press, 1974.
- [5] R. Ellis, "Large deviations for a general class of random vectors," *Ann. Probab.*, vol. 12, no. 1, pp. 1–12, 1984.
- [6] W. Feller, *An Introduction to Probability Theory and Its Applications—Vol. II*, 2nd ed. New York: Wiley, 1971.
- [7] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
- [8] J. Gärtner, "On large deviations from the invariant measure," *Theory Probab. Appl.*, vol. 22, pp. 24–39, 1977.
- [9] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 767–800, 1972.
- [10] H. Miller, "A convexity property in the theory of random variables on a finite Markov chain," *Ann. Math. Statist.*, vol. 32, pp. 1260–1270, 1961.
- [11] J. S. Sadowsky and R. K. Bahr, "Direct sequence spread spectrum multiple access communications with random signature sequences: A large deviations analysis," *IEEE Trans. Inform. Theory*, vol. 37, pt. 1, pp. 514–527, Mar. 1991.

The Asymptotic Risk in a Signal Parameter Estimation Problem

Lawrence D. Brown and Richard C. Liu

Abstract—In estimating the unknown location of a rectangular signal observed with white noise, the asymptotic risks of three important estimators are compared under L_1/L_2 losses. A different numerical scheme is used to improve the accuracy of Ibragimov/Hasminskii's result, which also leads to further information and numerical comparisons about the problem.

Index Terms—Rectangular signal, Bayes/minimax risks, squared/absolute error losses, MLE.

I. INTRODUCTION

Consider an observed signal of the form

$$dr(t) = s(t - \theta)dt + \sigma dB(t), \quad t \in (0, T + 1), \quad (1)$$

where $B(\cdot)$ is Brownian motion, $\theta \in (0, T)$ is an unknown shift parameter, and s is the rectangular signal with $s(\tau) = X_{[0,1]}(\tau)$. The objective is to estimate θ under normalized squared error loss, $L_2 = \sigma^{-4}(d - \theta)^2$, or under normalized absolute loss, $L_1 = \sigma^{-2}|d - \theta|$.

Manuscript received February 3, 1992; revised June 3, 1992. L. D. Brown was supported by NSF Grant DMS 8809016. R. C. Liu was supported by the NSF Postdoctoral Research Fellowships and in part by an Army Research Grant through Cornell MSI.

The authors are with the Department of Mathematics, White Hall, Cornell University, Ithaca, NY 14853.
IEEE Log Number 9203640.