Moderate deviations for quadratic forms in Gaussian stationary processes

YOSHIHIDE KAKIZAWA (Hokkaido University)†

Abstract

Moderate deviations limit theorem is proved for quadratic forms in zero-mean Gaussian stationary processes. Two particular cases are the cumulative periodogram and the kernel spectral density estimator. We derive the exponential decay of the logarithm of moderate deviation probabilities of goodness-of-fit tests for the spectral density. Using them we also discuss intermediate asymptotic efficiencies of tests, as in Kallenberg (1983; AS).

Large deviation theorem is a big issue in probability theory. We now recall the standard definition, as follows (cf. Dembo and Zeitouni (1998; Springer)). We say that a family of probability measures \( \mu_\varepsilon \) on a Polish space \( \mathcal{X} \) satisfies a Large Deviation Principle (LDP) with rate function \( I \) if \( I \) is a lower semicontinuous function \( I : \mathcal{X} \to [0, \infty] \) such that for any closed set \( F \) of \( \mathcal{X} \), \( \limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon (F) \leq -\inf_{x \in F} I(x) \), and for any open set \( G \) of \( \mathcal{X} \), \( -\inf_{x \in G} I(x) \leq \liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon (G) \). In many cases a countable family of measures \( \mu_T \) is considered, where \( \mu_T \) is the law of some random variable \( Z_T \in \mathbb{R}^d \), \( d \geq 1 \). For this situation, we say that \( Z_T \) satisfies the LDP with a speed \( a_T \), if the corresponding LDP statements are given by \( \mathcal{X} = \mathbb{R}^d \) and \( a_T^{-1} \log \mu_T (\cdot) = a_T^{-1} \log P(Z_T \in \cdot) (T \to \infty) \) for some sequence \( \{a_T\} \) of positive real numbers tending to \( \infty \), instead of \( \varepsilon \log \mu_\varepsilon (\cdot) (\varepsilon \to 0) \). The Cramér theorem states that the empirical mean of iid random variables taking values in \( \mathbb{R}^d \) satisfies the LDP with the speed \( a_T = T \) (sample size). An extension of the Cramér theorem has been known in the literature as the Gärtner and Ellis theorem: Assuming that (i) \( \lim_{T \to \infty} c_T (s) = c(s) \) exists as an extended real number, where \( c_T (s) = a_T^{-1} \log E[\exp(a_T s' Z_T)] \) is the normalized cumulant generating function of \( a_T Z_T \in \mathbb{R}^d \), (ii) \( 0 \in \text{Int}(D) \), where \( D = \{s \in \mathbb{R}^d : c(s) < \infty\} \), (iii) \( c \) is lower semicontinuous on \( \mathbb{R}^d \), (iv) \( c \) is differentiable on \( \text{Int}(D) \), and (v) either \( D = \mathbb{R}^d \) or \( c \) is steep at \( \partial D \), then \( Z_T \) satisfies the LDP with the speed \( a_T \), whose rate function is the Fenchel-Legendre transform of \( c \); \( c^*(x) = \sup_{s \in \mathbb{R}^d} \{x' s - c(s)\} \), \( x \in \mathbb{R}^d \). Although their theorem is quite general in its scope, it does not cover all \( \mathbb{R}^d \) cases in which the LDP holds. For the case where \( c \) is not steep, Dembo and Zeitouni (1995; In: Progress in Probability, Volume 36, Bolthausen et al.) presented a refinement of the technique.

†E-mail: kakizawa@econ.hokudai.ac.jp
of measure tilting in lower bound. See also Bercu et al. (1997; SPA), Bryc and Dembo (1997; JTP) and Zani (2002; JMA) for recent developments in time series (in many cases, set \( d = 1 \)). Our result is a consequence of the Gärtner and Ellis theorem.

Let \( \{Y_t : t \in \mathbb{Z}\} \) be a real-valued zero-mean stationary process with autocovariance function \( R(h) = E(Y_tY_{t+h}) = R(-h) \) satisfying \( \sum_{h=-\infty}^{\infty} |R(h)| < \infty \). The spectral density is then defined by

\[
f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} R(h)e^{-ih\lambda} = \frac{1}{2\pi} R(0) + \frac{1}{\pi} \sum_{h=1}^{\infty} R(h) \cos(h\lambda) \geq 0
\]

(e.g. Brockwell and Davis (1991; Springer)). We denote by \( L(f, \phi) \) the spectral average with the weight function \( \phi \not\equiv 0 \),

\[
L(f, \phi) = \int_{-\pi}^{\pi} \tilde{\phi}(\lambda)f(\lambda) d\lambda = \int_{-\pi}^{\pi} \frac{1}{2} \{\phi(\lambda) + \phi(-\lambda)\} f(\lambda) d\lambda = \int_{-\pi}^{\pi} \tilde{\phi}(\lambda)f(\lambda) d\lambda \quad \text{(say)}.
\]

Well-known estimates for the quantities \( L(f, \phi) \) and \( f(\alpha) \), based on the data \( Y_1, \ldots, Y_T \), are

\[
L(I_T, \phi) = \int_{-\pi}^{\pi} \tilde{\phi}(\lambda)I_T(\lambda) d\lambda = \frac{1}{2\pi T} \sum_{s,t=1}^{T} Y_sY_t \int_{-\pi}^{\pi} \tilde{\phi}(\lambda)e^{i(s-t)\lambda} d\lambda
\]

and

\[
\hat{f}_T(\alpha) = \int_{-\pi}^{\pi} K_M(\alpha - \lambda)I_T(\lambda) d\lambda = \int_{-\pi}^{\pi} \frac{1}{2} \{K_M(\alpha - \lambda) + K_M(\alpha + \lambda)\}I_T(\lambda) d\lambda.
\]

Here and below, the periodogram is defined by

\[
I_T(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} Y_te^{it\lambda} \right|^2 = \frac{1}{2\pi T} \sum_{s,t=1}^{T} Y_sY_t e^{i(s-t)\lambda}.
\]

As in the standard textbook of time series analysis (e.g. Hannan (1970; Wiley), Anderson (1971; Wiley) and Priestley (1981; Academic Press)), we consider two types of \( 2\pi \)-periodic function \( K_M(\cdot) \), called the spectral window. One is given by

\[
K_M(\lambda) = M \sum_{\nu=-\infty}^{\infty} K\{M(\lambda + 2\pi\nu)\}
\]

and the other is given by

\[
K_M(\lambda) = \frac{1}{2\pi} \sum_{h=-[M]}^{[M]} g\left(\frac{h}{M}\right) e^{-ih\lambda},
\]

where \( K(x), x \in \mathbb{R} \), is an even and integrable function with \( K(x) = 0 \) for \( |x| > \pi \) that integrates to one, \( g(x), x \in [0, 1] \), is an even and continuous function with \( g(0) = 1 \),
$M = M_T$ is a positive real number increasing to $\infty$ at a suitable rate depending on $T$, and
$[x]$ denotes the greatest integer not greater than $x$; $[x] \leq x < [x] + 1$. Since the spectral
density $f$ and its kernel spectral estimator $\hat{f}_T$ which we consider are even and $2\pi$-periodic
functions, it is sufficient to confine ourselves to the study of the interval $0 \leq \alpha \leq \pi$.

Asymptotic normality of $T^{1/2}\{L(I_T, \phi) - L(f, \phi)\}$ and $(T/M)^{1/2}\{\hat{f}_T(\alpha) - f(\alpha)\}$ has been well-established under various weaker regularity conditions (e.g. Brillinger (1981; Holden-Day) and Hosoya and Taniguchi (1982; AS)). Bentkus and Rudzkis (1982; LMJ) and Janas (1994; AISM) investigated the Edgeworth expansions for the distributions of $T^{1/2}\{L(I_T, \phi) - L(f, \phi)\}$ and $(T/M)^{1/2}\{\hat{f}_T(\alpha) - f(\alpha)\}$. See Velasco and Robinson (2001; ET) for the case of the studentized sample mean $\{2\pi T \hat{f}_T(0)\}^{-1/2} \sum_{t=1}^{T} Y_t$ and Taniguchi et al. (2003; ET) for the higher-order asymptotics of minimum contrast estimators based on $\hat{f}_T(\cdot)$. Bentkus (1978; LMJ) and Bentkus and Rudzkis (1982; TPA) derived upper bounds for tail probabilities $P[|L(I_T, \phi) - L(f, \phi)| \geq x]$ and $P[\sup_{\alpha \in [0, \pi]} |\hat{f}_T(\alpha) - f(\alpha)| \geq x]$ $(x > 0)$ as more precise versions of Bentkus et al. (1975; LMJ) and Rudzkis (1977; LMJ). Under the Gaussianity, Bercu et al. (1997; SPA), Bryc and Dembo (1997; JTP) and Taniguchi and Kakizawa (2000; Springer) established the LDP for quadratic forms, especially, for $L(I_T, \phi)$. Taking account of these results, it is natural to investigate a LDP
for $(T/a_T)^{1/2}\{L(I_T, \phi) - L(f, \phi)\}$ or $(T/(Ma_T))^{1/2}\{\hat{f}_T(\alpha) - f(\alpha)\}$ with a speed $a_T \to \infty$ such that $a_T = o(T)$ or $o(T/M)$. Such a context is often called a Moderate Deviation
Principle (MDP), since it is an intermediate estimate between the central limit theorem
and the LDP. As seen in Kallenberg (1983; AS), our MDP for quadratic forms is linked
with asymptotic efficiency in Gaussian stationary times series. Asymptotic efficiencies via
the LDP are found in Taniguchi and Kakizawa (2000; Springer) and Kakizawa (2005; JNS)
for time series. Nikitin (1995; Cambridge University Press) discussed several asymptotic
efficiencies of nonparametric tests in the IID model. We cite also Louani (1998; SJS),
Worms (2001; MMS) and Gao (2003; JTP) for the probability density estimation.

Our results based on the periodogram $I_T$ are easily extended to the case of the so-called
tapered periodogram

$$\tilde{I}_T(\lambda) = \frac{1}{2\pi H_{T,2}} \sum_{t=1}^{T} h\left(\frac{t}{T}\right) h\left(\frac{t}{T}\right) e^{i\lambda t},$$

with $H_{T,2} = \sum_{t=1}^{T} h^2(t/T)$, where $h(x)$ is bounded, is of bounded variation and vanishes
for $|x| > 1$ (see Brillinger (1981; Holden-Day) and Dahlhaus (1983; JTSA)).