Two approaches of analysis of efficiency in hypothesis testing

Pitman efficiency – CLT zone

Bahadur efficiency – Large deviation zone

Our results will be based on versions of Pitman efficiency in moderate deviation zone (MD and SMD efficiency). The local Bahadur efficiency is a particular case of MD - efficiency

Pitman efficiency

Let $X_1, \ldots, X_n$ be independent realizations of random variable X having pm P_θ, $\theta \in R^1$. Let $P_\theta >> \mu$ and let $f(x,\theta) = dP_\theta/d\mu(x)$. Suppose that the Fisher information

$$I(\theta) = \int \left(\frac{f_\theta(x,\theta)}{f(x,\theta)} \right)^2 dP_\theta < \infty$$

is finite.

The problem is to test the hypothesis

$$H_0 : \theta = \theta_0$$

versus

$$H_n : \theta = \theta_n = \theta_0 + \frac{u}{\sqrt{n}} = \theta_0 + u_n.$$

For any sequence of tests K_n denote $\alpha(K_n)$ and $\beta(K_n, \theta_n)$ their types I and types II error probabilities.

Lower bound. Let $\alpha(K_n) = \alpha$. Then

$$\beta(K_n, \theta_n) \geq \Phi(x_\alpha - uI^{1/2}(\theta_0))(1 + o(1)) \qquad (1)$$

Here x_α satisfies $\alpha = 1 - \Phi(x_\alpha)$.

In Ermakov (2003) a version of (1) was proved in moderate deviation zone with $nu_n^3 \to 0, nu_n^2 \to \infty$ as $n \to \infty$. If the corresponding lower bound is attained, we say that a sequence of tests is efficient in the sense of Strong (Sharp) Moderate Deviations (SMD -efficient).

For logarithmic asymptotic ($u_n \to 0, nu_n^2 \to \infty$) the following version of (1) hold

$$\ln \beta(K_n, \theta_n) \geq -\frac{1}{2}(x_{\alpha_n} - u_n I^{1/2}(\theta_0))^2(1 + o(1)) \qquad (2)$$

or in the more universal form

$$\limsup_{n\to\infty}(nu_n^2 I(\theta_0))^{-1/2}(\sqrt{2|\ln \alpha(K_n)|} + (\sqrt{2|\ln \beta(K_n, \theta_n)|} \leq 1 \qquad (3)$$

If lower bound in (3) is attained, we say that a sequence of tests is efficient in the sense of Moderate Deviations (MD -efficient)

Nonparametric Hypothesis Testing

Inglot, Kallenberg and Ledwina (1993 - till now) – Kallenberg efficiency

Ingster, Suslina (2001-2003) -new type of efficiency

Signal Detection in Gaussian white noise

We observe a signal S in Gaussian white noise. The problem is to test the hypothesis $H_0 : S \equiv 0$ versus

$$H_\epsilon : ||S|| > \rho_\epsilon, \quad S \in U_\beta$$

where

$$U_\beta = \left\{ S : \int_0^1 S^2(t) + (S^{(\beta)}(t))^2 dt < P_0 \right\}.$$

In terms of ortonormal system of functions the same problem is the following.

We observe

$$y_j = s_j + \epsilon\xi_j, \quad 1 \le j < \infty,$$

where ξ_j are i.i.d.r.v's, $\xi_j\tilde{N}(0,1)$.

One needs to test the hypothesis

$$H_0 : s_j = 0, \quad 1 \le j < \infty, \qquad s = \{s_j\}_{j=1}^{\infty}$$

versus

$$s \in U_\epsilon = U_\epsilon(\beta) = \left\{ s : ||s|| > \rho_\epsilon, \ \sum_{j=1}^{\infty} (1 + (2\pi j)^{2\beta}) s_j^2 \le P_0 \right\}. \qquad (4)$$

For any test K_ϵ denote $\alpha(K_\epsilon)$ its type I error probability and $\beta(K_\epsilon, S)$ its type II error probability for the alternative S.

Denote

$$\beta(K_\epsilon, U_\epsilon) = \sup_{S \in U_\epsilon} \beta(K_\epsilon, S)$$

We say that the family of tests L_ϵ is asymptotically minimax in the sense of sharp asymptotic of moderate deviation probabilities (SMD - asymptotically minimax), if for any family of tests $M_\epsilon, \alpha(M_\epsilon) \le \alpha(L_\epsilon)$ there holds

$$\limsup_{\epsilon \to 0} \frac{\beta(L_\epsilon, U_\epsilon)}{\beta(M_\epsilon, U_\epsilon)} \le 1 \tag{5}$$

We say that the family of tests L_ϵ is asymptotically minimax in the sense of logarithmic asymptotic of moderate deviation probabilities (MD - asymptotically minimax) , if for any family of tests $M_\epsilon, \alpha(M_\epsilon) \le \alpha(L_\epsilon)$ there holds

$$\limsup_{\epsilon \to 0} \frac{\ln \beta(L_\epsilon, U_\epsilon)}{\ln \beta(M_\epsilon, U_\epsilon)} \ge 1. \tag{6}$$

The results are different for three zones of moderate deviations.

We show that standard asymptotically minimax tests for Sobolev spaces (see Ermakov (1989)) are asymptotically minimax in the sense of SMD-efficiency in zone

$$\epsilon^{\frac{4\beta}{4\beta+1}} << \rho_\epsilon << \epsilon^{\frac{3\beta}{3\beta+1}} \tag{7}$$

The same result hold in the sense of MD -efficiency for the zone

$$\epsilon^{\frac{4\beta}{4\beta+1}} << \rho_\epsilon << \epsilon^{\frac{2\beta}{2\beta+1}}. \tag{8}$$

The standard chi-squared tests turn out asymptotically minimax in the sense of MD - efficiency for the zone

$$\epsilon^{\frac{2\beta}{2\beta+1}} << \rho_\epsilon << 1 \tag{9}$$

First two zones. Define the sequence $\kappa_j^2 = (\lambda - \mu j^{2\beta})_+$, with λ, μ satisfying the equations

$$\sum_{j=1}^{\infty} \kappa_j^2 = \rho_\epsilon^2, \quad \sum_{j=1}^{\infty} j^{2\beta} \kappa_j^2 = P_0$$

Denote

$$A_\epsilon = \epsilon^{-4} \sum_{j=1}^{\infty} \kappa_j^4.$$

Define the family of test statistics

$$T_\epsilon = \epsilon^{-2} \sum_{j=1}^{\infty} \kappa_j^2 (y_j^2 - \epsilon^2)$$

and corresponding family of tests

$$L_\epsilon = \chi((2A_\epsilon)^{-1/2} T_\epsilon > \epsilon^2 x_{\alpha_\epsilon})$$

where x_{α_ϵ} satisfies the equation $\alpha_\epsilon = 1 - \Phi(x_{\alpha_\epsilon})$.

Theorem 1 *Assume*

$$\epsilon^{\frac{4\beta}{4\beta+1}} << \rho_\epsilon << \epsilon^{\frac{3\beta}{3\beta+1}} \tag{10}$$

and let $0 < \alpha_\epsilon < c < 1$. *Then for any family of tests* K_ϵ *such that* $\alpha(K_\epsilon) = \alpha_\epsilon(1 + o(1)), \beta(K_\epsilon, U_\epsilon) < c < 1$ *there holds*

$$\beta(K_\epsilon, U_\epsilon) \geq \Phi(x_{\alpha_\epsilon} - (A_\epsilon/2)^{1/2})(1 + o(1)). \tag{11}$$

The lower bound (11) is attained on the SMD asymptotically minimax family of tests L_ϵ.

Theorem 2 *Assume*

$$\epsilon^{\frac{4\beta}{4\beta+1}} << \rho_\epsilon << \epsilon^{\frac{2\beta}{2\beta+1}}. \tag{12}$$

and let $0 < \alpha_\epsilon < c < 1$. *Then for any family of tests* K_ϵ *such that* $\ln \alpha(K_\epsilon) = \ln \alpha_\epsilon(1 + o(1)), \beta(K_\epsilon, U_\epsilon) < c < 1$ *there holds*

$$\ln \beta(K_\epsilon, U_\epsilon) \geq -\frac{1}{2}(x_{\alpha_\epsilon} - (A_\epsilon/2)^{1/2})^2(1 + o(1)) \tag{13}$$

The lower bound (13) is attained on the MD asymptotically minimax family of tests L_ϵ.

Thus for any sequence of tests K_ϵ *with* $\alpha(K_\epsilon) < c < 1, \beta(K_\epsilon, U_\epsilon) < c < 1$ *there holds*

$$\limsup_{n\to\infty}(A_\epsilon/2)^{-1/2}(\sqrt{2|\ln \alpha(K_\epsilon)|} + (\sqrt{2|\ln \beta(K_n, U_\epsilon)|} \leq 1 \tag{14}$$

Define the chi-squared test statistics

$$T_{m\epsilon} = \sum_{j=1}^{m} (y_j^2 - \epsilon^2).$$

Let x_{α_ϵ} satisfies $\alpha_\epsilon = 1 - \Phi(x_{\alpha_\epsilon})$. Define the chi-squared tests

$$M_\epsilon = \chi(m^{-1/2} T_{m\epsilon} > x_{\alpha_\epsilon}),$$

Suppose that $m = m_\epsilon$ satisfies $\epsilon^{-2}\rho_\epsilon^2 >> m > c\epsilon^{-\frac{2}{2\beta+1}}$.

Theorem 3 *Assume*

$$\epsilon^{\frac{2\beta}{2\beta+1}} << \rho_\epsilon << 1. \qquad (15)$$

and let $0 < \alpha_\epsilon < c < 1$. *Then for any family of tests* K_ϵ *such that* $\ln \alpha(K_\epsilon) = \ln \alpha_\epsilon (1 + o(1)), \beta(K_\epsilon, U_\epsilon) < c < 1$ *there holds*

$$\ln \beta(K_\epsilon, U_\epsilon) \geq -\frac{1}{2}(x_{\alpha_\epsilon} - \epsilon^{-1}\rho_\epsilon)^2 (1 + o(1)) \qquad (16)$$

The lower bound (16) is attained on the MD asymptotically minimax family of tests M_ϵ.

Note that the stronger property holds for the family of tests M_ϵ. We have

$$\ln \beta(M_\epsilon, s) = -\frac{1}{2}(x_{\alpha_\epsilon} - \epsilon^{-1}||s||)^2(1 + o(1)). \tag{17}$$

uniformly in $s \in W_\epsilon(\beta) = \left\{ s : \sum_{j=1}^{\infty} j^{2\beta} s_j^2 \le P_0, ||s||^2 \epsilon^{-2} > c_\epsilon m \right\}$ where $c_\epsilon \to \infty$ as $\epsilon \to 0$. Therefore The family of tests M_ϵ is uniformly MD asymptotically efficient in the set $W_\epsilon(\beta)$. This implies that the family of tests M_ϵ is uniformly MD asymptotically efficient in the set $W_\epsilon(\beta)$ for all $0 < \beta_0 < \beta < \beta_1 < \infty$, where $m > C\epsilon^{-\frac{1}{2\beta_1+1}}$.

NONPARAMETRIC CONFIDENCE ESTIMATION IN MODERATE DEVIATION ZONE

Confidence sets are defined with confidence level $1-\alpha$ where α is usually small. Thus we have the standard settup of large and moderate deviation probabilities. We treat this problem if a priory information is given

$$V_\beta = \left\{ s : \sum_{j=1}^{\infty} (1 + j^{2\beta}) s_j^2 < P_0, s = \{s_j\}_{j=1}^{\infty} \right\}.$$

We say that $\hat{\Gamma}_\epsilon$ is a confidence set with confidence level $1-\alpha$ if

$$\inf_{s \in V_\beta} P_s(s \in \hat{\Gamma}_\epsilon) \geq 1 - \alpha.$$

For any set $\Theta \subset l_2$ denote

$$\text{diam}(\Theta) = \sup\{||\theta_1 - \theta_2|| : \theta_1, \theta_2 \in \Theta\}.$$

If

$$\epsilon^{\frac{4\beta}{4\beta+1}} << \rho_\epsilon << \epsilon^{\frac{2\beta}{2\beta+1}} \tag{18}$$

then

$$\limsup_{\epsilon \to 0} 8 A_\epsilon^{-1} \sup_{s \in V_\beta} \ln P_s(\mathrm{diam}(\hat{\Gamma}_\epsilon) > \rho_\epsilon) \geq -1 \tag{19}$$

If

$$\epsilon^{\frac{2\beta}{2\beta+1}} << \rho_\epsilon << 1 \tag{20}$$

then

$$\limsup_{\epsilon \to 0} 4 \epsilon^2 \rho_\epsilon^{-2} \sup_{s \in V_\beta} \ln P_s(\mathrm{diam}(\hat{\Gamma}_\epsilon) > \rho_\epsilon) \geq -1. \tag{21}$$

In terms of α_ϵ these statements have the following form.

If $|\ln \alpha_\epsilon| << \epsilon^{-\frac{1}{2\beta+1}}$, then

$$\limsup_{\epsilon \to 0} (\ln \alpha_\epsilon)^{-1} \sup_{s \in V_\beta} \ln P_s \left(\mathrm{diam}(\hat{\Gamma}_\epsilon) > \left(\frac{P_0 |\ln \alpha_\epsilon| \epsilon^4}{C_\beta} \right)^{\frac{\beta}{4\beta+1}} \right) \leq 1, \quad (22)$$

where

$$C_\beta = \frac{(2\beta+1)^2(4\beta+2)}{32\beta^2(4\beta+1)^{\frac{2\beta+1}{2\beta}}}.$$

If $|\ln \alpha_\epsilon| >> \epsilon^{-\frac{1}{2\beta+1}}$, then

$$\limsup_{\epsilon \to 0} (\ln \alpha_\epsilon)^{-1} \sup_{s \in V_\beta} \ln P_s \left(\mathrm{diam}(\hat{\Gamma}_\epsilon) > 2\epsilon |\ln \alpha_\epsilon|^{1/2} \right) \leq 1. \quad (23)$$

If $|\ln \alpha_\epsilon| >> \epsilon^{-\frac{1}{2\beta+1}}$, the lower bound (23) is attained on confidence sets

$$\hat{\Gamma}_\epsilon = \{s : \sum_{j=1}^{m} (y_j - s_j)^2 \le \epsilon^2 |\ln \alpha_\epsilon|, s \in V_\beta\} \qquad (24)$$

where $|\ln \alpha_\epsilon| >> m > c\epsilon^{-1/(1+2\beta)}$.

Such an arbitrariness in the choice of β allows do not consider adaptive setup for this model.

KERNEL BASED TESTS

In nonparametric estimation

$$\text{"consistency"} \Leftrightarrow \text{"compactness assumptions"}$$

In nonparametric hypothesis testing we need the assumptions of completely different type for consistency (distinguishability of two sets of hypothesis).

Consider the problem of testing hypothesis

$$H_0 : S \in \Theta_0$$

versus

$$H_1 : S \in \Theta_1$$

where Θ_0, Θ_1 are closed bounded sets in $L_2(0, 1)$.

Theorem 4 *The sets* Θ_0 *and* Θ_1 *are distinguishable iff there exists finite dimensional space* Π, *dim*$(\Pi) < \infty$, *such that*

$$Pr_{\Pi}\Theta_0 \cap Pr_{\Pi}\Theta_1 = \emptyset \tag{25}$$

Thus a priori information that S belongs to some compact set (for example Sobolev ball) is not necessary in nonparametric hypothesis testing. We could not find compactness assumptions in traditional nonparametric hypothesis testing.

Compactness assumptions in testing nonparametric hypothesis is a monkey counterpart of assumptions in nonparametric estimation. We do not need such an assumption in testing nonparametric hypothesis. All above mentioned results hold for the wider sets of alternatives

$$\{s : T_\epsilon(s) > \lambda_\epsilon \rho_\epsilon - \mu_\epsilon P_0, s \in l_2\} \tag{26}$$

Thus **the distance method** *is the most natural approach for the testing nonparametric hypothesis*.

Distance Method

$d_\epsilon(S - S_1)$ -seminorms in $L_2(0,1)$

Problem:

$$H_0 : d_\epsilon(S) = 0$$

versus

$$H_\epsilon : d_\epsilon(S) > \rho_\epsilon.$$

One needs to show that

$d_\epsilon(Y_\epsilon)$ *is asymptotically minimax*

Such a results have been proved for *Kolmogorov*, ω^2, *chi-squared and kernel-based tests*.

In this talk versions of Theorems 1-3 will be given for the kernel-based tests.

Let we observe random process $Y_\epsilon(t), t \in (0, 1)$ satisfying

$$dY_\epsilon(t) = S(t)dt + \epsilon q(t)dw(t), \tag{27}$$

where $dw(t)$ is Gaussian white noise and $q(t)$ is a weight function.

One needs to test the hypothesis $S(t) = 0, t \in (0, 1)$.

To test the hypothesis define the kernel estimator

$$\widehat{S}_\epsilon(t) = \frac{1}{h_\epsilon} \int K\left(\frac{t-s}{h_\epsilon}\right) Y_\epsilon(s)\, ds \tag{28}$$

with some kernel K and kernel width h_ϵ.

Define test statistics

$$T_\epsilon(Y_\epsilon) = \int_0^1 \widehat{S}_\epsilon^2(t) r(t)\, dt. \tag{29}$$

Here $r(t) > 0, t \in (0,1)$ is a weight function.

The sets of alternatives are induced the test statistics

$$U_\epsilon = \left\{ S : T_\epsilon(S) = \int_0^1 \left(\int_0^1 K_{h_\epsilon}(t-s) S(s) ds \right)^2 r(t)\, dt > \rho_\epsilon, S \in L_2(0,1) \right\}. \tag{30}$$

Here $K_{h_\epsilon}(s) = \frac{1}{h_\epsilon} K\left(\frac{s}{h_\epsilon}\right)$.

In Ermakov (2003) we proved that the test statistics $T_\epsilon(Y_\epsilon)$ are asymptotically minimax in the CLT -zone ($\rho_\epsilon \asymp \epsilon^2 h_\epsilon^{1/2}$). The Pitman efficiency was considered.

Assumptions

A1. support $K \subseteq [-1, 1]$, $K(t) = K(-t)$ for all $t \in (0, 1)$, $\int_{-1}^{1} K(t)dt = 1$ and $K(t)$ is bounded.

A2. The functions $r(t), q(t)$ are positive in $[0, 1]$and satisfies the Lipshitz conditions

$$|q(t) - q(s)| < C|t - s|^{\gamma}, \quad t, s \in [0, 1], \quad \gamma > \frac{1}{2}, \tag{31}$$

$$|r(t) - r(s)| < C|t - s|^{\gamma}, \quad t, s \in [0, 1], \quad \gamma > \frac{1}{2}. \tag{32}$$

Denote

$$d_\epsilon = d_\epsilon(h_\epsilon) = E_0 T_\epsilon(Y_\epsilon) = \frac{\epsilon^2}{h_\epsilon}\int_0^1 r(t)dt \int_{-1}^1 K^2(u) q^2(t-uh_\epsilon)du =$$

$$= \frac{\epsilon^2}{h_\epsilon}\int_0^1 q^2(t)r(t)dt \int_{-1}^1 K^2(u)du(1+o(h_\epsilon^{1/2})) = \frac{\epsilon^2}{h_\epsilon} d(1+o(h_\epsilon^{1/2}))$$

$$\sigma^2 = 2\int_{-2}^2 K_2^2(v)dv \int_0^1 q^4(t)r^2(t)dt.$$

Here we suppose t $q(t) = 0$, if $t \notin [0,1]$ and make use of the notation $K_2(t) = \int_{-1}^1 K(t-s)K(s)\, ds$

For any family of tests M_ϵ denote

$$\beta(M_\epsilon) = \beta(M_\epsilon, U_\epsilon) = \sup\{\beta(M_\epsilon, S) : S \in U_\epsilon\}.$$

Define the family of tests

$$L_\epsilon = \chi\left\{\epsilon^{-2} h_\epsilon^{1/2} \sigma^{-1} (T_\epsilon(Y_\epsilon) - d_\epsilon) > x_{\alpha_\epsilon}\right\}, \tag{33}$$

where x_{α_ϵ} satisfies the equation $\alpha_\epsilon = \Phi(-x_{\alpha_\epsilon})$.

The versions of Theorems 1 -3 were obtained for the following three zones of large deviations

$$\epsilon^2 h_\epsilon^{-1/2} <<< \rho_\epsilon <<< \epsilon^2 h_\epsilon^{-2/3} \tag{34}$$

$$\epsilon^2 h_\epsilon^{-1/2} <<< \rho_\epsilon <<< \epsilon^2 h_\epsilon^{-1} \tag{35}$$

$$\epsilon^2 h_\epsilon^{-1} <<< \rho_\epsilon <<< 1 \tag{36}$$

Theorem 5 . *Let*

$$\epsilon^2 h_\epsilon^{-1/2} <<< \rho_\epsilon <<< \epsilon^2 h_\epsilon^{-2/3} \tag{37}$$

and let $h_\epsilon \to 0$ *as* $\epsilon \to 0$. *Let* $\alpha_\epsilon \to 0$ *as* $\epsilon \to 0$. *Then the family of tests* L_ϵ *is SMD - asymptotically minimax and* $\alpha(L_\epsilon) = \alpha_\epsilon(1 + o(1))$,

$$\beta(L_\epsilon, U_\epsilon) = \Phi\left(x_{\alpha_\epsilon} - \frac{h_\epsilon^{1/2}\rho_\epsilon}{\epsilon^2\sigma}\right)(1 + o(1)). \tag{38}$$

Moreover for any family of signals $S_\epsilon \in U_\epsilon(\rho_\epsilon, h_\epsilon)$ *such that* $T_\epsilon(S_\epsilon) = \rho_\epsilon$, *there holds*

$$\beta(L_\epsilon, S_\epsilon) = \Phi\left(x_{\alpha_\epsilon} - \frac{h_\epsilon^{1/2}\rho_\epsilon}{\epsilon^2\sigma}\right)(1 + o(1)). \tag{39}$$

Theorem 6 . *Let*

$$\epsilon^2 h_\epsilon^{-1/2} <<< \rho_\epsilon <<< \epsilon^2 h_\epsilon^{-1} \tag{40}$$

and $h_\epsilon \to 0$ *as* $\epsilon \to 0$*. Let* $\alpha_\epsilon \to 0$ *as* $\epsilon \to 0$*. Then the family of tests* L_ϵ *is MD-asymptotically minimax and* $\ln \alpha(L_\epsilon) = \ln \alpha_\epsilon(1 + o(1))$

$$\ln \beta(L_\epsilon, U_\epsilon) = -\frac{1}{2}\left(\sqrt{2|\ln \alpha(L_\epsilon)|} - h_\epsilon^{1/2} \rho_\epsilon \epsilon^{-2} \sigma^{-1}\right)^2 (1 + o(1)). \tag{41}$$

Moreover for any family of signals $S_\epsilon \in U_\epsilon(\rho_\epsilon, h_\epsilon)$ *such that* $T(h_\epsilon, S_\epsilon) = \rho_\epsilon$*, in (41) the equality is attained*

$$\ln \beta(L_\epsilon, S_\epsilon) = -\frac{1}{2}\left(\sqrt{2|\ln \alpha(L_\epsilon)|} - h_\epsilon^{1/2} \rho_\epsilon \epsilon^{-2} \sigma^{-1}\right)^2 (1 + o(1)) \tag{42}$$

Define the family of tests

$$L_{1\epsilon}(Y_\epsilon) = \chi(\sigma_1 \epsilon^{-1} T_\epsilon(Y_\epsilon) > x_{\alpha_\epsilon}),$$

where $\sigma_1^2 = \max_{x\in(0,1)} r(x)q^2(x)$.

Theorem 7 . *Let*

$$\rho_\epsilon \epsilon^2 h_\epsilon^{-1} <<< \rho_\epsilon <<< 1 \qquad (43)$$

and $h_\epsilon \to 0$ *as* $\epsilon \to 0$. *Let* $\alpha_\epsilon \to 0$ *as* $\epsilon \to 0$. *Then the family of tests* $L_{1\epsilon}$ *is MD-asymptotically minimax and* $\ln \alpha(L_{1\epsilon}) = \ln \alpha_\epsilon(1 + o(1))$,

$$\ln \beta(L_{1\epsilon}, U_\epsilon) = -\frac{1}{2}\left(\sqrt{2|\ln \alpha(L_\epsilon)|} - \epsilon^{-1}\sigma_1^{-1}\rho_\epsilon^{1/2}\right)^2 (1 + o(1)). \qquad (44)$$