Weierstraß–Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

On large and moderate large deviations of empirical bootstrap measure

Michael S. Ermakov

submitted: 10th October 1996

Mechanical Engineering
Problems Institute
Russian Academy of Sciences
Bolshoy pr., V.O., 61
199178 St. Petersburg
Russia

Preprint No. 271
Berlin 1996

1991 Mathematics Subject Classification. 60F10, 62C12, 62E25.
Key words and phrases. Large deviations, bootstrap method, empirical measure.

Edited by
Weierstraß–Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
D — 10117 Berlin
Germany

Fax: + 49 30 2044975
e–mail (X.400): c=de;a=d400-gw;p=WIAS-BERLIN;s=preprint
e–mail (Internet): preprint@wias-berlin.de

Summary. We find the asymptotics for the large and moderate large deviation probabilities of common distribution of the empirical measure and the empirical bootstrap measure (empirical measure obtaining by the bootstrap method). For the most widespread statistical functionals depending on empirical measure we compare their asymptotics of moderate large deviation probabilities with similar asymptotics given by the bootstrap procedure.

1. Introduction. Let S be a Hausdorff space, $\Im$ the σ-field of Borel sets in S and Λ the space of all probability measures (pms) on $(S,\Im)$. Let $X_1,\dots,X_n$ be i.i.d.r.v.'s taking values in S according to a pm $P_0 \in \Lambda$ and let $\hat{P}_n$ be the empirical measure of $X_1,\dots,X_n$. The distributions of statistics depending on the sample $X_1,\dots,X_n$ are often analyzed on the base of the bootstrap procedure (see Hall (1992), Mammen (1992) and Efron and Tibshirany(1993)). For a given statistics $V(X_1,\dots,X_n)$, we simulate independent samples $X_1^*,\dots,X_n^*$ with the distribution $\hat{P}_n$ and consider the empirical distribution of $V(X_1^*,\dots,X_n^*)$ as an estimator of the distribution of $V(X_1,\dots,X_n)$. What is of especial interest, are the estimates of large and moderate large deviation probabilities of $V(X_1,\dots,X_n)$. From this viewpoint it is natural to compare probabilities of large and moderate large deviations for $V(X_1,\dots,X_n)$ and $V(X_1^*,\dots,X_n^*)$. In the paper we carry out such a comparison in a slightly different setting. The statistics $V(X_1,\dots,X_n)$ usually can be represented as a functional $T(\hat{P}_n)$ of the empirical measure $\hat{P}_n$: $V(X_1,\dots,X_n) = T(\hat{P}_n)$. Similarly, $V(X_1^*,\dots,X_n^*) = T(P_n^*)$, where P_n^* is the empirical measure of $X_1^*,\dots,X_n^*$. Thus, the problem is reduced to the study of large and moderate deviations of statistical functionals $T(\hat{P}_n)-T(P)$ and $T(P_n^*)-T(\hat{P}_n)$.

The problems related to large and moderate large deviation probabilities of empirical measures have been studied in many papers (see Sanov (1957), Hoadley (1967), Stone (1974), Sievers (1978), Groeneboom, Oosterhoff, Ruymgaart (1979) (GOR), Borovkov and Mogulskii (1980), and Ermakov (1992),(1995)). These papers contain complete results proved under rather general assumptions. Our goal here is to develop similar techniques for large and moderate large deviations of $P_n^* \times \hat{P}_n$ and to use these techniques to compare the probabilities of the deviations $T(\hat{P}_n)-T(P)$ and $T(P_n^*)-T(\hat{P}_n)$. Thus, we intend to consider the following two settings: the asymptotics of the probabilities $P_0(P_n^* \times \hat{P}_n \in \bar{\Omega})$ with $\bar{\Omega} \in \Lambda^2 = \Lambda \times \Lambda$

and the asymptotics of the probabilities $P_0(P_n^* \times \hat{P}_n \in \bar{\Omega}_n)$ where $\bar{\Omega}_n \subset \Lambda^2$ and P_0 is a limiting point of $\bar{\Omega}_n$.

The problem of large deviations for empirical bootstrap measure has been studied earlier in Chaganty (1993). These results were obtained on the base of another approach and were given in terms of the topology of weak convergence instead of the τ-topology considered in the paper. It seems that the approach developed in GOR (1979) can be easily extended to the large deviations problem for the empirical bootstrap measure. Similar situation takes place also for the moderate large deviations, with the only difference that here more essential modification of analytical technique of GOR (1979) and Ermakov (1992), (1995) is required.

Note that the results on large deviations for $P_n^* \times \hat{P}_n$ are far from being "computable", except some special cases (see Chaganty (1993)). At the same time the moderate large deviation theorems allow easily to compare the probabilities of deviations of $T(\hat{P}_n) - T(P)$ and $T(P_n^*) - T(\hat{P}_n)$ for the majority of widespread statistics. Using the same technique as in GOR (1979) and Ermakov (1992), (1995), we also consider the problem of moderate large deviation in the parametric bootstrap setting. The large deviation problem for parametric bootstrap has been studied earlier in Chaganty (1993) in terms of the topology of weak convergence.

In this paper we use the following notations. We denote by C, c arbitrary positive constants, by $\chi(A)$ the indicator of an event A, and by $[t]$ the integral part of a real number t.

2. Large deviation probabilities. Introduce the τ-topology of weak convergence in the space Λ. We say that a sequence $Q_n \in \Lambda$ converges to $Q \in \Lambda$ in τ-topology iff

$$\lim_{n\to\infty} \int_S f\,dQ_n = \int_S f\,dQ$$

for each bounded $\Im$-measurable function $f : S \to R^1$. In what follows all topological notions, except otherwise is explicitly stated, relate to τ-topology (for details, see GOR (1979)). For any set $\Omega \subset \Lambda$ the closure and the interior of Ω are denoted by $\mathrm{cl}\,(\Omega)$ and $\mathrm{int}\,(\Omega)$ respectively. The τ-topology in Λ^2 is defined as the corresponding product topology.

For any $P, Q \in \Lambda$ we introduce the Kullback-Leibler information number $K(Q, P)$ as

$$K(Q,P) = \int_S q \log q\,dP, \quad q = \frac{dQ}{dP}$$

if Q is absolutely continuous w.r.t. P, and $K(Q,P) = \infty$ otherwise. It is known (see, e.g., GOR (1979)) that the asymptotics of large deviation probabilities of empirical measures $\hat{P}_n$ can be expressed in terms of the Kullback-Leibler information numbers. The analog of Kullback-Leibler information numbers for the large deviation probabilities of $P_n^* \times \hat{P}_n$ is the functional $K_b(\bar{Q}, P)$ which is defined as $K_b(\bar{Q}, P) = K(Q_2, Q_1) + K(Q_1, P)$ for any $\bar{Q} = Q_1 \times Q_2 \in \Lambda^2$ and $P \in \Lambda$. We also set $K_b(\bar{\Omega}, P) = \inf\{K_b(\bar{Q}, P) : \bar{Q} \in \bar{\Omega}\}$ for any $\bar{\Omega} \subset \Lambda^2$ and $P \in \Lambda$.

Theorem 2.1. *Let $P_0 \in \Lambda$ and let $\bar{\Omega} \subset \Lambda^2$. Let $K_b(\text{int}\,(\bar{\Omega}), P_0) = K_b(\text{cl}\,(\bar{\Omega}), P_0)$. Then*

$$\lim_{n\to\infty} n^{-1} \log P_0(P_n^* \times \hat{P}_n \in \bar{\Omega}) = -K_b(\bar{\Omega}, P_0). \tag{2.1}$$

Remark. Theorem 2.1 admits the following interpretation. The righthand side of (2.1) equals the infimum of the product of asymptotics of two probabilities. The first is the asymptotics of probability that the empirical measure $\hat{P}_n$ belongs to "a small vicinity of pm Q_1". The second is the asymptotics of probability that the empirical bootstrap measure P^* belongs to "a small vicinity of Q_2" under the condition that $\hat{P}_n$ belongs to "a small vicinity of Q_1".

The proof of Theorem 2.1 in its main elements is based on the same arguments as the proof of theorem on large deviations of empirical measures (see GOR (1979)). In particular, the analogies of Lemmas 2.3-2.5 in GOR (1979) are valid for our setting as well. The differences in the proof of the analogy of the main Lemma 3.1 in GOR (1979) are clearly seen from similar arguments for the moderate large deviation setting (see the proof of Theorem 3.2 below).

Chaganty (1993) has developed a special technique to prove the convergence of $K_b(\bar{Q}_n, P)$ to $K_b(\bar{Q}, P)$ if a sequence of pms $\bar{Q}_n$ converges to $\bar{Q}$ in weak topology. As a consequence, Chaganty obtained a version of Theorem 2.1 related to weak topology. Passing to such a limit cannot be used straightforwardly in our setting, and here we apply the partition technique. For any $P \in \Lambda$, $\bar{Q} = Q_1 \times Q_2 \in \Lambda^2$ and a partition $\Pi = \{S_j\}_{j=1}^m$ of S consisting of Borel sets S_j, let

$$K_b(\bar{Q}, P|\Pi) = \sum_{j=1}^m \left(Q_2(S_j) \log \frac{Q_2(S_j)}{Q(S_j)} + Q_1(S_j) \log \frac{Q_1(S_j)}{P(S_j)} \right).$$

For any $\bar{\Omega} \subset \Lambda^2$ we set $K_b(\bar{\Omega}, P|\Pi) = \inf\{K_b(\bar{Q}, P|\Pi) : \bar{Q} \in \Omega\}$ With this notation, all analytical estimates are based on the relation (compare with Lemma 2.4 in GOR (1979))

$$K_b(\bar{\Omega}, P) = \sup\{K_b(\bar{\Omega}, P) : \Pi \text{ is a partition of } S\} \tag{2.2}$$

if $K_b(\bar{\Omega}, P) = K_b(\text{cl}(\bar{\Omega}, P)$.

The proof of (2.2) makes use of the same arguments as the proof of similar relation (2.5) in GOR (1979) and is based on the fact that the set $\Gamma = \{\bar{Q} : K_b(\bar{Q}, P) < C < \infty, \bar{Q} \in \Lambda^2\}$ is uniformly absolutely continuous w.r.t. P and is both compact and sequentially compact in the τ-topology. These statements are proved similarly to that one of Lemma 2.3 in GOR (1979).

3. Moderate large deviation probabilities. Let us introduce the linear space Λ_0 generated by all differences $P - Q$ with $P, Q \in \Lambda$ and define the τ-topologies on Λ_0 and Λ_0^2 similarly to that one on Λ and Λ^2, respectively.

For any $P, Q \in \Lambda$, introduce the Hellinger distance

$$\rho(P,Q) = \left(\int_S \left(\left(\frac{dP}{dR}\right)^{1/2} - \left(\frac{dQ}{dR}\right)^{1/2} \right)^2 dR \right)^{1/2}, \quad R = \frac{1}{2}(P+Q). \tag{3.1}$$

Define also, for any $G \in \Lambda_0$ and $P \in \Lambda$, the functional $\rho_0(G, P)$ such that

$$\rho_0^2(G : P) = \int_S \left(\frac{dG}{dP}\right)^2 dP \tag{3.2}$$

if G is absolutely continuous w.r.t. P and $\rho_0(G : P) = \infty$ otherwise. The functionals ρ and ρ_0 are the standard tools for the study of moderate large deviation probabilities of empirical measures $\hat{P}_n$ (see Borovkov and Mogulskii (1980) and Ermakov (1992), (1995)). It is well known that for any $G \in \Lambda_0$ satisfying the inclusion $P + uG \in \Lambda$ and such that $\rho_0(G : P) < \infty$ it holds

$$K(P+uG, P) = 2\rho^2(P+uG, P) + o(u^2) = \frac{1}{2}u^2\rho^2(G : P) + o(u^2)$$

as $u \to 0$.

For any $\bar{Q} = Q_1 \times Q_2 \in \Lambda^2$ and $\bar{G} = G_1 \times G_2 \in \Lambda_0^2$ denote $\bar{Q} + \bar{G} = (Q_1 + G_1) \times (Q_2 + G_2)$. We set

$$\rho_b^2(\bar{Q}, \bar{P}) = \rho^2(Q_2, Q_1) + \rho^2(Q_1, P),$$

$$\rho_{0b}^2(\bar{G}, P) = \rho_0^2(G_2 - G_1 : P) + \rho_0^2(G_1, P).$$

The functionals ρ_b and ρ_{0b} play the same role in the study of moderate large deviations of $P_n^* \times \hat{P}_n$ as the functional K_b in the study of large deviations. It is easily seen that for any $\bar{G} = G_1 \times G_2 \in \Lambda_0$ such that $P + uG_1, P + uG_2 \in \Lambda$ and $\rho_{0b}(\bar{G} : P) < \infty$ it holds

$$K_b((\bar{P} + u\bar{G}), P) = 2\rho_b((\bar{P} + u\bar{G}), P) + o(u^2) = \frac{1}{2}u^2\rho_{0b}(\bar{G} : P) + o(u^2)$$

as $u \to 0$. Here $\bar{P} = P \times P$

For any $\Omega \subset \Lambda$, $\bar{\Omega} \subset \Lambda^2$ and $\Omega_0 \subset \Lambda_0$, $\bar{\Omega}_0 \subset \Lambda_0^2$ define

$$\rho(\Omega, P) = \inf\{\rho(Q, P) : Q \in \Omega\},$$

$$\rho_0(\Omega_0 : P) = \inf\{\rho_0(G : P) : G \in \Omega_0\},$$

$$\rho_b(\bar{\Omega}, P) = \inf\{\rho_b(\bar{Q}, P) : Q \in \Omega\}$$

and $\rho_{0b}(\bar{\Omega}_0, P) = \inf\{\rho_{0b}(\bar{Q}, P) : \bar{Q} \in \bar{\Omega}_0\}$.

Theorem 3.1. *Let $P_0 \in \Lambda$ and $\bar{\Omega}_0 \subset \Lambda_0^2$. Let*

$$\rho_{0b}(\text{int}\,(\bar{\Omega}_0) : P_0) = \rho_{0b}(\text{cl}\,(\bar{\Omega}_0) : P_0). \tag{3.3}$$

Then, for any sequence $b_n \to 0$, $nb_n^2 \to \infty$ as $n \to \infty$,

$$\lim_{n\to\infty} (nb_n^2)^{-1} \log P_0(P_n^* \times \hat{P}_n \in P_0 \times P_0 + b_n\bar{\Omega}_0) = -\frac{1}{2}\rho_0^2(\bar{\Omega}_0 : P_0). \tag{3.4}$$

Remark. A similar version of a theorem on moderate large deviation probabilities of empirical measures is proved in Borovkov and Mogulskii (1980).

The asymptotics (3.4) can be applied for the study of moderate large deviations of homogeneous functionals. At the same time, the majority of statistical fuctionals are only approximately homogeneous. To obtain the results for this latter case, we have developed the following approach (see Ermakov (1992),(1995)).

Let a sequence of pms P_n converge to pm P_0 and let $\bar{P}_0 = P_0 \times P_0$ be a limiting point of sets $\bar{\Omega}_n \subset \Lambda^2$. Denote $b_n = \rho(\bar{\Omega}_n, \bar{P}_n)$ and suppose that $nb_n^2 \to \infty$ as $n \to \infty$.

Assume that

A1. There exists a sequence $H_n \in \Lambda_0$ such that H_n are absolutely continuous w.r.t. P_0, $\rho(P_n, P_0 + b_nH_n) = o(b_n)$ as $n \to \infty$ and for any sequence $C_n \to \infty$ as $n \to \infty$

$$\lim_{n\to\infty} \int_S \left(\frac{dH_n}{dP_0}\right)^2 \chi\left(\left|\frac{dH_n}{dP_0}\right| > C_n\right) dP_0 = 0. \tag{3.5}$$

A1 implies that pms P_n can be approximated by a sequence of pms $P_0 + b_nH_n$ such that dH_n/dP_0 are uniformly integrable in $L_2(P_0)$.

A2. There exist an open set $\bar{\Omega}_0 \subset \Lambda_0^2$ and a function ω, $\omega(x)/x \to 0$ as $x \to 0$ such that

i. For any sequence $\bar{G}_n \in \bar{\Omega}_0$ there exists a sequence $\bar{Q}_n \in \bar{\Omega}_n$ such that $\rho_b(\bar{Q}_n, \bar{P}_0 + b_n\bar{G}_n) < \omega(\rho_b(\bar{P}_0, \bar{P}_0 + b_n\bar{G}_n)$.

ii. For any sequence $\bar{Q}_n \in \bar{\Omega}_n$ there exists a sequence $\bar{G}_n \in \bar{\Omega}_0$ such that $\rho_b(\bar{Q}_n, \bar{P}_0 + b_n\bar{G}_n) < \omega(\rho_b(\bar{P}_0, \bar{Q}_n))$.

A2 implies that the sequence of sets $\bar{\Omega}_n$ can be approximated in the Hellinger metric by the sets $\bar{P}_0 + b_n\bar{\Omega}_0$.

Theorem 3.2. *Assume* A1,A2. *Let* $b_n \to 0$, $nb_n^2 \to \infty$ *as* $n \to \infty$ and let $\rho_{0b}(\text{cl}\,\bar{\Omega}_0 : P_0) = \rho_{0b}(\text{int}\,\bar{\Omega}_0 : P_0)$. *Then*

$$\lim_{n\to\infty} \frac{\log P_n(P_n^* \times \hat{P}_n \in \bar{\Omega}_n)}{2n\rho_b^2(\bar{\Omega}_n, \bar{P}_n)} = -1. \tag{3.6}$$

The analogy of Theorem 3.2 is also valid for the moderate large deviations of $P_k^* \times \hat{P}_n$, where P_k^* is the empirical measure of independent sample $X_1^*, \dots, X_k^*$ with the pm $\hat{P}_n$.

Suppose that $k/n \to \nu > 0$ as $n \to \infty$.

For any $\bar{Q} = Q_1 \times Q_2 \in \Lambda$ and $\bar{G} = G_1 \times G_2 \in \Lambda_0^2$ denote

$$\rho_\nu^2(\bar{Q}, P) = \nu\,\rho^2(Q_2, Q_1) + \rho^2(Q_1, P),$$

$$\rho_{0\nu}^2(\bar{G} : P) = \nu\,\rho_0^2(G_2 - G_1 : P) + \rho_0^2(G_1 : P).$$

For any $\bar{\Omega} \subset \Lambda^2$ and $\bar{\Omega}_0 \subset \Lambda_0^2$ we set $\rho_\nu(\bar{\Omega}, P) = \inf\{\rho_\nu(\bar{Q}, P) : \bar{Q} \in \bar{\Omega}\}$ and $\rho_{0\nu}(\bar{\Omega}_0 : P) = \inf\{\rho_{0\nu}(\bar{G} : P) : \bar{G} \in \bar{\Omega}_0\}$

Theorem 3.3. *Assume* A1,A2. *Let* $b_n \to 0$, $nb_n^2 \to \infty$ *as* $n \to \infty$ and let $\rho_{0\nu}\text{cl}\,(\bar{\Omega}_0) : P_0) = \rho_{0\nu}(\text{int}\,(\bar{\Omega}_0) : P_0)$. *Then*

$$\lim_{n\to\infty} \frac{\log P_n(P_k^* \times \hat{P}_n \in \bar{\Omega}_n)}{2n\rho_\nu^2(\bar{\Omega}_n, \bar{P}_n)} = -1. \tag{3.7}$$

Let $H_n = b_n H$, $H \in \Lambda_0$. *Then, in the other terms,* (3.7) *can be written as*

$$\lim_{n\to\infty} (nb_n^2)^{-1} \log P_n(P_k^* \times \hat{P}_n \in \bar{\Omega}_n) = -\frac{1}{2}\rho_{0\nu}^2(\bar{\Omega}_0 - \bar{H} : P_0)$$

where $\bar{H} = (\nu H) \times H$.

Clearly, similar version of Theorem 3.3 can be established also for the large deviation probabilities of empirical bootstrap measure. The proof of Theorem 3.3 is akin to that one of Theorem 3.2 and is omitted. From now on, we assume $k = n$.

Example 3.1. Differentiable statistical functionals. A linear approximation of statistics is the standard tool to prove its asymptotic normality (see Serfling (1980), Denker (1985)). Here we apply the same technique for the study of moderate large deviation. We assume that the functional $T : \Lambda \to R^1$ admits a linear approximation of the following type.

B. There exist a bounded function r and a norm N in Λ_0 continuous in τ-topology such that
$$\left|T(Q) - T(P_0) - \int_S r\, d(Q - P_0)\right| < \omega(N(Q - P_0)).$$
where $\omega(t)$ is an increasing function such that $\omega(t)/t \to 0$ as $t \to 0$.

Assumptions of such a type without that severe limitations on the function r and the norm N (continuiuty in τ-topology) were used for the proof of asymptotic normality of statistics $T(\hat{P}_n)$ (see Serfling (1980)) and in implicit form also for the study of moderate large deviations (see Jureckova, Kallenberg and Veraverbeke (1988), Inglot, Kallenberg and Ledwina (1990),(1992), and Ermakov (1994)). At the same time the corresponding technique to weaken these limitations is developed in the theory of large deviations of empirical measures in Groeneboom and Shorack (1981). That allows to suppose that a similar results can be obtained for this setting as well.

If B holds, then, as it follows easily from Theorems 3.1 and 3.2, for any sequence P_n converging to P and satisfying A1 we have
$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_n(T(P_n^*) - T(\hat{P}_n) > b_n) =$$
$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_n\left(\int_S r\, d(P_n^* - \hat{P}_n\right) > b_n) =$$
$$-\frac{1}{2} \inf\left\{\int_S ((g_2 - g_1)^2 + g_1^2)\, dP_0 : \int_S (g_2 - g_1) r\, dP_0 > \int_S r^2\, dP_0\right\} =$$
$$-\frac{1}{2}\int_S r^2\, dP_0, \tag{3.8}$$
$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_n(T(\hat{P}_n) - T(P_n) > b_n) =$$
$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_n\left(\int_S r\, d(\hat{P}_n - P_n) > b_n\right) =$$
$$-\frac{1}{2} \inf\left\{\int_S g^2\, dP_0 : \int_S gr\, dP_0 > \int_S r^2 dP_0\right\} = -\frac{1}{2}\int_S r^2 dP_0. \tag{3.9}$$
Thus, the asymptotics of moderate large deviation probabilities of $T(P_n^*) - T(\hat{P}_n)$ and $T(\hat{P}_n) - T(P)$ coincide. At the same time by Theorem 3.2
$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_n(T(P_n^*) - T(P_n) > b_n) =$$
$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_n(\int_S r d(P_n^* - P_n) > b_n) = -\frac{1}{4}\int_S r^2 dP_0. \tag{3.10}$$
The proof of (3.8) is very easy and (3.9),(3.10) are obtained by a similar technique. Introduce the inverse function $\bar{\omega}(t)$ of $\omega(s)$ such that $\bar{\omega}(t) = s$ implies $\omega(s) = t$. By Theorem 3.2 and B we have
$$|P_n(T(P^*) - T(\hat{P}_n) > b_n) - P_n(\int_S r\, d(P_n^* - \hat{P}_n) > b_n)| <$$

$$P_n(N(P_n^* - \hat{P}_n) > \bar{\omega}(b_n)) < \exp\{-Cn\bar{\omega}^2(b_n)\}.$$

The asymptotics of $P_n(\int_S r\, d(P_n^* - \hat{P}_n) > b_n)$ follows directly from Theorem 3.2.

Example 3.2. Variance. Let $T(P) = Var_P[X] = E_P[X^2] - (E_P[X])^2$. Then the problem of identifying the asymptotics of the right-hand side of (3.8) can be reduced to minimization of the functional

$$\int_S ((g_2 - g_1)^2 + g_1^2)\, dP_0 \tag{3.11}$$

under the constraint that

$$\int_S x^2(g_2(x) - g_1(x))\, P_0(dx) - \Big(\int_S x^2(1 + g_2(x))\, P_0(dx)\Big)^2 + \Big(\int_S x(1 + g_2(x))\, P_0(dx)\Big)^2 =$$

$$\int_S x^2(g_2(x) - g_1(x))\, P_0(dx) - 2E_{P_0}[X] \int_S x(g_2(x) - g_1(x))\, P_0(dx) +$$

$$\int_S x(g_2(x) - g_1(x))\, P_0(dx) \int_S x(g_1(x) + g_2(x))\, P_0(dx) =$$

$$\int_S x^2(g_2(x) - g_1(x))\, P_0(dx) - 2E_{P_0}[X] \int_S x(g_2(x) - g_1(x))\, P_0(dx) + o(b_n) > b_n \tag{3.12}$$

Thus, the functional $T(P)$ admits the linear approximation and

$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_0(T(P_n^*) - T(\hat{P}_n) > b_n) =$$

$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_0(T(\hat{P}_n) - T(P_0) > b_n) = -\frac{1}{2} E[X^2 - 2X\, E[X]]^2. \tag{3.13}$$

This example can be considered as a particular case of Example 3.1, since $Var_P[X]$ has the influence function.

Example 3.3. Homogeneous functionals. It is easily seen that the analog of (3.10) holds also in the case of an arbitrary norm $N : \Lambda_0 \to R^1$ continuous in τ-topology

$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_0(N(P_n^* - P_0) > b_n) = -\frac{1}{4} \rho_0^2(\Omega_0 : P_0). \tag{3.14}$$

Here $\Omega_0 = \{G : N(G) > 1, G \in \Lambda_0\}$.

For the moderate large deviations of empirical measure $\hat{P}_n$ (see Ermakov (1995)) we have

$$\lim_{n\to\infty} (nb_n^2)^{-1/2} \log P_0(N(\hat{P}_n - P_0) > b_n) = -\frac{1}{2} \rho_0^2(\Omega_0 : P_0). \tag{3.15}$$

In particular, the relations (3.14) and (3.15) are valid for the functional N defined as the functional of the Kolmogorov and the omega-square test statistics

$$N(P - P_0) = \max\{|F(x) - x| : x \in S\}$$

or

$$N(P - P_0) = \Big(\int_0^1 (F(x) - x)^2 dx\Big)^{1/2}.$$

Here $S = (0,1)$, P_0 is the pm of the uniform distribution and F is the distribution function of pm P.

Let us prove that

$$\lim_{n\to\infty}(nb_n^2)^{-1/2}\log P_0(N(P_n^* - P_0) - N(\hat{P}_n - P_0) > b_n) = -\frac{1}{2}\rho_0^2(\Omega_0 : P_0). \quad (3.16)$$

For any $a, d > 0$ denote $\Omega(a) = \{P : N(P - P_0) \le a,\ P \in \Lambda\}$ and $\Phi(d) = \{P : N(P - P_0) \ge d,\ P \in \Lambda\}$, $\Phi(d) - \Omega(a) = \{G : G = Q - P,\ Q \in \Phi(d),\ P \in \Omega(a)\}$. Relation (3.16) is given by Theorem 3.1 combined with the following computation:

$$\inf\{\int_S((g_2 - g_1)^2 + g_1^2)\,dP_0 : N(G_2) - N(G_1) \ge b_n,$$

$$g_1 = dG_1/dP_0, g_2 = dG_2/dP_0, G_1, G_2 \in \Lambda_0\} =$$

$$= \inf\{\rho_0^2(\Phi(d) - \Omega(a) : P_0) + \rho_0^2(\Omega(a) - P_0 : P_0) : d - a > b_n\} \quad (3.17)$$

By Theorem 3.2 in Ermakov (1995), there exists a function $h \in L_2(P_0)$, $\int_S h dP_0 = 0$ such that the family of pms W_λ, $\lambda > 0$, $dW_\lambda/dP_0 = c_\lambda + \lambda h_\lambda(x)$, $h_\lambda(x) = h(x)\chi(h(x) > -c_\lambda/\lambda)$, $c_\lambda \to 1$ as $\lambda \to 0$, satisfies as follows $W_\lambda \in \Phi(\lambda + o(\lambda))$ and

$$\rho^2(\Phi(\lambda), P_0) = \rho^2(W_\lambda, P_0)(1 + o(1)) = \frac{1}{4}\lambda^2 \int_S h^2(s)\,dP_0(1 + o(1)) \quad (3.18)$$

as $\lambda \to 0$. Note that (3.17) is valid also in the case of $\Omega(0) = \{P_0\}$. Therefore the right-hand side of (3.18) equals

$$\inf\{(d - a)^2 + a^2 : d - a > b_n\} \int_S h^2(s)\,ds\,(1 + o(1)) =$$

$$b_n^2\rho_0^2(\Omega_0 : P_0)(1 + o(1)), \quad (3.19)$$

and (3.16) follows.

The moderate large deviation probabilities of $T(P_n^*) - T(\hat{P}_n)$ with $T(P) = N^\gamma(P - P_0)$, $\gamma > 1$, can be expressed similarly to (3.17),(3.19). Here we get

$$\lim_{n\to\infty}(nb_n^2)^{-1/2}\log P_0(N^\gamma(P_n^* - P_0) - N^\gamma(\hat{P}_n - P_0) > b_n) =$$

$$\inf\{(d - a)^2 + a^2 : d^\gamma - a^\gamma > b_n\} > \int_S h^2(s)ds(1 + o(1))$$

This relation shows that for $\gamma > 1$ the asymptotics of $\log P(T(P_n^*) - T(\hat{P}_n) > b_n)$ does not coincide with the corresponding asymptotics of $\log P(T(\hat{P}_n) - T(P) > b_n)$.

4. Moderate large deviations in the parametric bootstrap. Let $X_1, \dots, X_n$ be i.i.d.r.v.'s with pm $P_{\theta_0} \in \Theta \subset R^d$ and let $T(\hat{P}_n)$ be an estimator of θ_0. In the parametric bootstrap procedure, given $T(\hat{P}_n) = \hat{\theta}_n$, the bootstrap sample is a sequence of i.i.d. observations $X_1^*, \dots, X_n^*$ from $P_{\hat{\theta}_n}$. We shall study these procedures under the following assumptions;

C1. There exist charges $H_1, \dots, H_d \in \Lambda_0$ and a function ω, $\omega(x)/x \to 0$ as $x \to 0$, such that $\rho\left(P_\theta, P_{\theta_0} + \sum_{i=1}^d (\theta_i - \theta_{0i}) H_i\right) < \omega(\rho(P_\theta, P_{\theta_0}))$. Here $\theta = (\theta_1, \dots, \theta_d)$ and $\theta_0 = (\theta_{01}, \dots, \theta_{0d})$.

C2. The charges H_i, $1 \le i \le d$, are absolutely continuous w.r.t. P_{θ_0} and have bounded densities $h_i = dH_i/dP_{\theta_0}$, $1 \le i \le d$.

Suppose that $T(P) = (T_1(P), \dots, T_d(P))$.

C3. There exist bounded functions r_i, $1 \le i \le d$, and a norm $N : \Lambda_0 \to R^1$ which is continuous in the τ-topology such that

$$\left| T_i(P) - T_i(P_{\theta_0}) - \int_S r_i d(P - P_{\theta_0}) \right| \le \omega(N(P - P_{\theta_0})) \tag{4.1}$$

for all $1 \le i \le d$.

For any $G \in \Lambda_0$ denote

$$I(G, P_{\theta_0}) = \sum_{i,j=1}^d \int_S r_i \, dG \int_S h_i h_j dP_{\theta_0} \int_S r_j \, dG.$$

For $\bar{G} = G_1 \times G_2 \in \Lambda_0^2$ we set

$$\rho_{0p}^2(\bar{G}, P) = \rho_0^2 \left(G_2 - \sum_{i=1}^d \int_S r_i \, dG_1 \, H_i : P_{\theta_0} \right) + I(G_1, P_{\theta_0}).$$

For any $\bar{\Omega} \subset \Lambda_0^2$, let $\rho_{0p}(\bar{\Omega}_0 : P) = \inf\{\rho_{0p}(\bar{G} : P) : \bar{G} \in \bar{\Omega}_0\}$.

Theorem 4.1. *Assume* A2 *and* B1-B3. *Let* $b_n \to 0$, $nb_n^2 \to \infty$ *as* $n \to \infty$. *Then*

$$\lim_{n \to \infty} (nb_n^2)^{-1} \log P(P_n^* \times P_{\hat{\theta}_n} \in \bar{\Omega}_n) = -\frac{1}{2} \rho_{0p}^2(\bar{\Omega}_0 : P_{\theta_0}) \tag{4.2}$$

Example 4.1. Let $\theta \in R^1$ and let $\hat{\theta}_n = T_1(\hat{P}_n)$ be the estimator of θ satisfying C3 with the functional T_1, satisfying (4.1). Let $T_2 : \Lambda \to R^1$ satisfy (4.1) with $i = 2$. Then the moderate large deviation probabilities of $T_2(P_n^*) - T_2(P_{\hat{\theta}_n})$ and $T_2(P_n^*) - T_2(P_{\theta_0})$ have the following asymptotics. Then

$$\lim_{n \to \infty} (nb_n^2)^{-1} \log P_{\theta_0}(T_2(P_n^*) - T_2(P_{\hat{\theta}_n}) > b_n) =$$

$$-\frac{1}{2} \inf \left\{ \left(\int_S r_1 g_1 \, dP_{\theta_0} \right)^2 \int_S h^2 dP_{\theta_0} + \int_S \left(g_2 - h \int_S r_1 g_1 \, dP_{\theta_0} \right)^2 dP_{\theta_0} : \right.$$

$$\left. \int_S r_2 \left(g_2 - h \int_S r_1 g_1 \, dP_{\theta_0} \right) dP_{\theta_0} > \int_S r_2^2 \, dP_{\theta_0} \right\} = -\frac{1}{2} \int_S r_2^2 dP_{\theta_0} \tag{4.3}$$

and

$$\lim_{n\to\infty}(nb_n^2)^{-1}\log P_{\theta_0}(T_2(P_n^*)-T_2(P_{\theta_0})>b_n)=-\frac{1}{2}\inf\left\{\left(\int_S r_1g_1\,dP_{\theta_0}\right)^2\int_S h^2\,dP_{\theta_0}+\right.$$

$$\left.\int_S\left(g_2-h\int_S r_1g_1\,dP_{\theta_0}\right)^2 dP_{\theta_0}:\int_S r_2g_2dP_{\theta_0}>\int_S r_2^2dP_{\theta_0}\right\}=$$

$$-\frac{1}{2}\frac{\left(\int_S r_2^2dP_{\theta_0}\right)^2\int_S h^2dP_{\theta_0}}{\int_S r_2^2\,dP_{\theta_0}\int_S h^2\,dP_{\theta_0}+\left(\int_S r_2h\,dP_{\theta_0}\right)^2}. \tag{4.4}$$

5. Proofs of Theorems 3.2 and 4.1. The reasoning in its main elements is akin to Ermakov (1995). The arguments are based on the approximation of the problem in question by the associated problem for the multinomial distributions. The approximation is realized on the base of partitions $\Pi=\{S_j\}_1^m$ of the Hausdorff space S consisting of the finite number of Borel sets S_j. We analyze the probabilities of a number of observations $(X_{i_1}^*,X_{i_2})$ in each set $S_{j_1}\times S_{j_2}$ and, as a result, obtain the assertions of Theorem 3.2 and 4.1.

Proof of Theorem 3.2. For any partition $\Pi=\{S_j\}_1^m$ and any $Q,P\in\Lambda$, $G\in\Lambda_0$, $\bar{Q}=Q_1\times Q_2\in\Lambda^2$, $\bar{G}=G_1\times G_2\in\Lambda_0^2$ denote

$$\rho^2(Q,P|\Pi)=\sum_{j=1}^{k}(Q^{1/2}(S_j)-P^{1/2}(S_j))^2,$$

$$\rho_b^2(\bar{Q},P|\Pi)=\rho^2(Q_2,Q_1|\Pi)+\rho^2(Q_1,P|\Pi),$$

$$\rho_0^2(G:P|\Pi)=\sum_{j=1}^{m}\frac{G^2(S_j)}{P(S_j)},$$

$$\rho_{0b}^2(\bar{G}:P|\Pi)=\rho_0^2(G_2-G_1:P|\Pi)+\rho_0^2(G_1:P|\Pi).$$

It is easy to see that

$$\rho(Q,P)=\sup_\Pi\rho(Q,P|\Pi),\qquad \rho_b(\bar{Q},P)=\sup_\Pi\rho_b(\bar{Q},P|\Pi),$$

$$\rho_0(G:P)=\sup_\Pi\rho(G:P|\Pi),\qquad \rho_{0b}(\bar{G}:\Pi)=\sup_\Pi\rho_{0b}(\bar{G}:P|\Pi).$$

Here the supremum is taken over the all partitions Π of S.

For any partition $\Pi=\{S_j\}_1^m$ and any sets $\bar{\Omega}\subset\Lambda^2$ and denote $\rho_b(\bar{\Omega},P|\Pi)=\inf\{\rho_b(\bar{Q},P|\Pi):\bar{Q}\in\bar{\Omega}\}$ and $\rho_{0b}(\bar{\Omega}_0,P|Pi)=\inf\{\rho_{0b}(\bar{G}:P|\Pi):\bar{G}\in\bar{\Omega}\}$ respectively.

Suppose that the sequence H_n converges to $H\in\Lambda_0$. Since the sequence of densities dH_n/dP is compact set in $L_2(P)$ the general setting is akin to this one.

Similarly to the main line of the proof of Sanov theorem the analytical technique will be based on the following relation (compare with (2.2))

$$\rho_{0b}(\bar{\Omega}_0 : P) = \sup\{\rho_{0b}(\bar{\Omega}_0, P|\Pi) : \Pi \text{ is a partition of } S\} \tag{5.1}$$

if $\rho_{0b}(\bar{\Omega}_0 : P) = \rho_0(\mathrm{cl}\,(\bar{\Omega}_0) : P)$.

By A1,A2, (5.1) implies

$$\rho_b(\bar{\Omega}_n, P_n)(1+o(1)) = \sup\{\rho_b(\Omega_n, P_n|\Pi) : \Pi \text{ is a partition of } S\} \tag{5.2}$$

as $n \to \infty$.

The proof of (5.1) is akin to that one of Lemma 2.4 in GOR (1979).

By A1,A2, the convergence H_n to H implies that there exists $\bar{G} = G_1 \times G_2 \in \mathrm{cl}\,(\bar{\Omega}_0)$ such that

$$\rho_b^2(\bar{\Omega}_n, P_n) = \rho_b^2(\bar{\Omega}_n, P_0 + b_n H_n)(1+o(1)) = \rho_b^2(P_0 + b_n\bar{\Omega}_0, P_0 + b_n H)(1+o(1)) =$$

$$\rho_b^2(P_0 + b_n\bar{G}_n, P_0 + b_n H)(1+o(1)) = \frac{1}{2} b_n^2 \rho_{0b}^2(\bar{G} - \bar{H} : P_0)(1+o(1)).$$

Here $\bar{H} = H \times H$ and $\bar{G} - \bar{H} = (G_1 - H) \times (G_2 - H)$:

Let $\Pi = \{S_j\}_1^m$ be a partition of S such that $p_j = P(S_j) > 0$, $1 \le j \le m$. Denote $p_{nj} = P_{nj}(S_j)$, $h_j = H(S_j)$, $g_{1nj} = G_{1n}(S_j)$, $g_{2nj} = G_{2n}(S_j)$, $g_{1j} = G_1(S_j)$, $g_{2j} = G_{2j}(S_j)$ for all $1 \le j \le k$. Put

$$\gamma_n^2 = \rho_b^2(\bar{G} : P_n|\Pi), \quad \gamma^2 = \frac{1}{4}\rho_{0b}(\bar{G} - \bar{H} : P_0|\Pi).$$

It is clear that

$$\rho_b^2(\Omega_n, P_n|\Pi) = b_n^2\gamma_n^2(1+o(1)) = \frac{1}{4} b_n^2\gamma^2(1+o(1)) \tag{5.3}$$

as $n \to \infty$.

By the Stirling formula, we get

$$P_n(P_n^* \times \hat{P}_n \in \bar{\Omega}_n) \le P_n(\rho_b(P_n^* \times \hat{P}_n, P_n|\Pi) > \rho_b(\bar{\Omega}_n, \bar{P}_n|\Pi)) =$$

$$\sum^* \frac{(n!)^2}{(nz_{n1})!\dots(nz_{nm})!(nx_{n1})!\dots(nx_{nm})!} \prod_{j=1}^m p_{nj}^{nz_{nj}} \prod_{j=1}^m z_{nj}^{nx_{nj}} \le$$

$$C\sum^* \exp\left\{(1-m)\log n - \frac{1}{2}\sum_{j=1}^m \log z_{nj} - \frac{1}{2}\sum_{j=1}^m \log x_{nj} - \right.$$

$$\left. n\sum_{j=1}^m z_{nj}\log\frac{z_{nj}}{p_{nj}} - n\sum_{j=1}^m x_{nj}\log\frac{x_{nj}}{z_{nj}}\right\} = B(Y_n) \tag{5.4}$$

Here $\sum^*$ denotes the summation over the set Y_n of all $z_n = (z_{n1}, \dots, z_{nm})$ and $x_n = (x_{n1}, \dots, x_{nm})$ such that $nz_{n1}, \dots, nz_{nm}$ and $nx_{n1}, \dots, nx_{nm}$ are whole nonnegative numbers, $z_{n1}+\ldots+z_{nm}=1$, $x_{n1}+\ldots x_{nm}=1$ and $J_{1n}(z_n)+J_{2n}(x_n,z_n)>b_n^2\gamma_n^2$. Here

$$J_{1n}(z_n)=\sum_{j=1}^{m}(z_{nj}^{1/2}-p_{nj}^{1/2})^2,\quad J_{2n}(x_n,z_n)=\sum_{j=1}^{m}(x_{nj}^{1/2}-z_{nj}^{1/2})^2.$$

Fix $\epsilon>0$. For all whole numbers i_1, $i_1>-1/\epsilon-1$, and all integers i_2 introduce the sets $U_{ni_1}=\{z_n:(1+\epsilon i_1)b_n^2\gamma_n^2<J_{1n}(z_n)\le(1+\epsilon(i_1+1))b_n^2\gamma_n^2,\ z_n=(z_{n1},\ldots,z_{nm})$, $nz_{n1},\ldots,nz_{nm}$ are whole nonnegative numbers, $z_{n1}+\ldots+z_{nm}=1\}$ and $W_{ni_2}(z_n)=\{x_n:\epsilon i_2b_n^2\gamma_n^2<J_{2n}(x_n,z_n)\le\epsilon(i_2+1)b_n^2\gamma_n^2,\ x_n=(x_{n1},\ldots,x_{nm})$, $nx_{n1},\ldots,nx_{nm}$ are whole nonnegative numbers, $x_{n1}+\ldots+x_{nm}=1\}$. Denote $Y_{ni_1,i_2}=\{(z_n,x_n): z_n\in U_{ni_1},x_n\in W_{ni_2}(z_n)\}$.

Put $d_\epsilon=-[1/\epsilon]-1$.

The numbers of elements U_{ni_1} and $W_{ni_2}(z_n)$ for $i_1>d_\epsilon+1$, $i_2>1$ and $z_n\in U_{ni_1}$ equal

$$C\epsilon(nb_n\gamma_n)^{2\lambda}(1+\epsilon i_1)^{\lambda-1}\prod_{j=1}^{m}p_{nj}^{m}(1+o(1)) \tag{5.5}$$

and

$$C\epsilon(nb_n\gamma_n)^{2\lambda}(\epsilon i_2)^{\lambda-1}\prod_{j=1}^{m}z_{nj}^{m}(1+o(1)), \tag{5.6}$$

respectively. Here $\lambda=(m-1)/2$.

Expanding $\log(z_{nj}/p_{nj})$ and $\log(x_{nj}/z_{nj})$ in Taylor series by powers of $(z_{nj}^{1/2}-p_{nj}^{1/2})p_{nj}^{-1/2}$ and $(x_{nj}^{1/2}-z_{nj}^{1/2})z_{nj}^{-1/2}$ respectively we get

$$I(Y_{ni_1i_2})<C\epsilon^2(1+\epsilon i_1)^{\lambda-1}(\epsilon i_2)^{\lambda-1}(nb_n^2\gamma_n^2)^{2\lambda}\exp\{-2n(1+\epsilon(i_1+i_2))b_n^2\gamma_n^2\}. \tag{5.7}$$

The detailed estimates in (5.7) are akin to (6.5) in Ermakov (1995).

Define the sequence $\epsilon=\epsilon_n$ such that $\epsilon_n\to0$, $n\epsilon_nb_n^2\gamma_n^2\to\infty$ as $n\to\infty$. Then we have

$$\sum_{i_1>d_n}\sum_{i_1+i_2\ge0}B(Y_{ni_1i_2})\le$$
$$C(nb_n^2\gamma_n)^{2\lambda}\int_0^\infty t^{\lambda-1}\exp\{-2nb_n^2\gamma_n^2t\}\int_{1-t}^\infty s^{\lambda-1}\exp\{-2nb_n^2\gamma_n^2s\}dsdt<$$
$$C(nb_n^2\gamma_n^2)^{2\lambda}\int_1^\infty u^{2\lambda-1}\exp\{-2nb_n^2\gamma_n^2u\}du=$$
$$\exp\{-2nb_n^2\gamma_n^2(1+o(1))\}=\exp\{-2nb_n^2\gamma^2(1+o(1))\} \tag{5.8}$$

Thus it remains to estimate

$$\sum_{i_2>1/\epsilon_n}B(Y_{nd_{\epsilon_n}i_2})\le P(\rho_b(P_n^*\times\hat P_n,P|\Pi)>\rho_b(\bar\Omega_n,P|\Pi)(1+o(1))\,|\,\rho(\hat P_n,P|\Pi)>\epsilon_n^{-1}b_n^2\gamma_n^2)<$$
$$\sum_{i>-d_n}C\epsilon_n(\epsilon_ni)^{\lambda-1}(nb_n^2\gamma_n^2)^{\lambda}\exp\{-2ni\epsilon_nb_n^2\gamma_n^2\}=\exp\{-2nb_n^2\gamma^2(1+o(1))\}, \tag{5.9}$$

$$\sum_{i_1>0} I(Y_{ni_10}) \le CP(\rho(\hat{P}_n, P|\Pi) > \rho(\bar{\Omega}_n, \bar{P}|\Pi)(1+o(1))) <$$

$$\exp\{-2nb_n^2\gamma^2(1+o(1))\}. \tag{5.10}$$

Here (5.9) follows from (6.6) in Ermakov (1995) and the estimates in (5.10) are similar to that in (5.9). Now (5.8)-(5.10) together implies (3.6).

Proof of Theorem 4.1. The difference in the arguments has no principal character in comparison with the proof of Theorem 3.2. Thus we introduce only new notations and point out auxilliary relations allowing to pass to the arguments of the proof of Theorem 3.2.

For any partition Π and set $\bar{\Omega} \subset \Lambda^2$ denote

$$\rho_p(\bar{\Omega}, P_{\theta_0}) = \inf\{\rho_b(\bar{Q}, P_{\theta_0}) : \bar{Q} = Q_2 \times P_\theta \in \bar{\Omega}\},$$

$$\rho_p(\bar{\Omega}, P_{\theta_0}|\Pi) = \inf\{\rho_b(\bar{Q}, P_{\theta_0}|\Pi) : \bar{Q} = Q_2 \times P_\theta \in \bar{\Omega}\}.$$

By B1–B3, for any $\theta_n \to \theta_0$, $\theta_n = (\theta_{n1}, \dots, \theta_{nd})$, as $n \to \infty$ and Q_n converging to Q_{θ_0} in τ-topology we have

$$\rho^2(P_\theta, P_{\theta_0}) = \frac{1}{4}\sum_{i,j=1}^{d} (\theta_{ni} - \theta_{0i}) \int_S h_i h_j dP_{\theta_0} (\theta_{nj} - \theta_{0j})(1+o(1)), \tag{5.11}$$

$$\rho\left(P_{T(Q_n)}, P_{\theta_0} + \sum_{i=1}^{d} \int_S r_i d(Q_n - P_{\theta_0})\, H_i\right) \le \omega(\rho(P_{T(Q_n)}, P_{\theta_0})), \tag{5.12}$$

$$\rho(P_{T(Q_n)}, P_{\theta_0}) = \rho\left(P_{\theta_0} + \sum_{i=1}^{d} \int_S r_i d(Q_n - P_{\theta_0})\, H_i, P_{\theta_0}\right) + O(\omega(\rho(P_{T(Q_n)}, P_{\theta_0}))), \tag{5.13}$$

$$\rho^2(P_{T(Q_n)}, P_{\theta_0}) = \frac{1}{4}\sum_{i,j=1}^{d} \int_S r_i d(Q_n - P_{\theta_0}) \int_S h_i h_j dP_{\theta_0} \int_S r_j d(Q_n - P_{\theta_0}) +$$

$$O\left(\sum_{i=1}^{d} \left|\int_S r_i d(Q_n - P_{\theta_0})\right| \omega(N(Q_n - P_{\theta_0}))\right) \tag{5.14}$$

Hence we get

$$\rho_p(\Omega_n, P_{\theta_0}) = \frac{1}{4} b_n^2 \rho_{0p}(\bar{\Omega}_0 : P)(1+o(1)) \tag{5.15}$$

as $n \to \infty$.

The further arguments are akin to the proof of Theorem 3.2. We write

$$P_{\theta_0}(P_n^* \times P_{\hat{\theta}_n} \in \bar{\Omega}_n) \le P_{\theta_0}(\rho_b(P_n^* \times P_{\hat{\theta}_n}, P_{\theta_0}|\Pi) > \rho_p(\bar{\Omega}_n, P_{\theta_0}|\Pi)) \tag{5.16}$$

and make use versions of relations (5.11) - (5.16) written for a given partition Π.

References

1. Borovkov,A.A. and Mogulskii,A.A. (1980) On probabilities of large deviations in topological spaces. II. Sibirsk. Math.Zh. 21(5) 12-26
2. Chaganty,N.R. (1993) Large deviations for joint distributions and statistical applications. Technical Report TR 93-2 Old Dominion University Norfolk. To appear Sankhia 1996.
3. Efron,B. and Tibshirany,R.J. (1993) An Introduction to the Bootstrap. Chapman and Hall, London, New York.
4. Ermakov,M.S. (1992) Large deviations for empirical probability measures and statistical tests. Zapiski Nauchn. Semin. POMI RAN 207 37-59 English transl.
5. Ermakov,M.S. (1994) On lower bounds of moderate large deviation probabilities of tests and estimators. Zapiski Nauchn.Semin.POMI RAN 216 52-61
6. .Ermakov,M.S. (1995) Asymptotic minimaxity of tests of Kolmogorov and omega-square type. Theory Probab. Appl. 40 54-67.
7. Groeneboom,P., Oosterhoff,J. and Ruymgaart,F.H. (1979) Large deviation theorems for empirical probability measures. Ann. Probab. 7 553-586
8. Groeneboom,P. and Shorack,G.R.(1981) Large deviations of goodness of fit statistics and linear combinations of order statistics. Ann. Probab. 9 971-987.
9. Hall,P. (1992) The Bootstrap and Edgeworth Expansion. Springer Verlag, New York.
10. Hoadley,A.B. (1967) On the probability of large deviations of functions of several empirical c.d.f.'s. Ann. Math. Statist. 38 360-381
11. Inglot,T.,Kallenberg,W.C.M. and Ledwina,T. (1992) Strong moderate deviation theorems. Ann. Probab. 20 987-1003
12. Inglot.T. and Ledwina,T. (1990) On probabilities of excessive deviations for Kolmogorov-Smirnov, Cramer-von Mises and chi-square statistics. Ann. Statist. 18 1491-1495
13. Jureckova,J., Kallenberg,W.C.M. and Veraverbeke,N. (1988) Moderate and Cramer type large deviation theorems for M-estimators. Statist. Probab. Lett. 6 191-199
14. Mammen,E. (1992) When does bootstrap work? Springer Verlag, New York.
15. Sanov,I. (1957) On the probability of large deviations of random variables. Math. Sbornik, 42 (In Russian) English translation: Selected Translations in Mathematical Statistics and Probability, 1961, 1 213-244.

16. Sievers,G.(1978) Estimates of location: A large deviation comparison. Ann. Statist., 6 610-618

17. Stone,M. (1974) Large deviations of empirical probability measures. Ann. Statist., 2 362-366.

Recent publications of the
Weierstraß–Institut für Angewandte Analysis und Stochastik

Preprints 1996

242. Susumu Okada, Siegfried Prößdorf: On the solution of the generalized airfoil equation.

243. Alexey K. Lopatin: Oscillations and dynamical systems: Normalization procedures and averaging.

244. Grigori N. Milstein: Stability index for invariant manifolds of stochastic systems.

245. Luis Barreira, Yakov Pesin, Jörg Schmeling: Dimension of hyperbolic measures – A proof of the Eckmann–Ruelle conjecture.

246. Leonid M. Fridman, Rainer J. Rumpel: On the asymptotic analysis of singularly perturbed systems with sliding mode.

247. Björn Sandstede: Instability of localised buckling modes in a one–dimensional strut model.

248. Björn Sandstede, Christopher K.R.T. Jones, James C. Alexander: Existence and stability of N–pulses on optical fibers with phase–sensitive amplifiers.

249. Vladimir Maz'ya, Gunther Schmidt: Approximate wavelets and the approximation of pseudodifferential operators.

250. Gottfried Bruckner, Sybille Handrock–Meyer, Hartmut Langmach: On the identification of soil transmissivity from measurements of the groundwater level.

251. Michael Schwarz: Phase transitions of shape memory alloys in soft and hard loading devices.

252. Gottfried Bruckner, Masahiro Yamamoto: On the determination of point sources by boundary observations: uniqueness, stability and reconstruction.

253. Anton Bovier, Véronique Gayrard: Hopfield models as generalized random mean field models.

254. Matthias Löwe: On the storage capacity of the Hopfield model.

255. Grigori N. Milstein: Random walk for elliptic equations and boundary layer.

256. Lutz Recke, Daniela Peterhof: Abstract forced symmetry breaking.

257. Lutz Recke, Daniela Peterhof: Forced frequency locking in S^1–equivariant differential equations.

258. Udo Krause: Idealkristalle als Abelsche Varietäten.

259. Nikolaus Bubner, Jürgen Sprekels: Optimal control of martensitic phase transitions in a deformation–driven experiment on shape memory alloys.

260. Christof Külske: Metastates in disordered mean field models: random field and Hopfield models.

261. Donald A. Dawson, Klaus Fleischmann: Longtime behavior of a branching process controlled by branching catalysts.

262. Tino Michael, Jürgen Borchardt: Convergence criteria for waveform iteration methods applied to partitioned DAE systems in chemical process simulation.

263. Michael H. Neumann, Jens–Peter Kreiss: Bootstrap confidence bands for the autoregression function.

264. Silvia Caprino, Mario Pulvirenti, Wolfgang Wagner: Stationary particle systems approximating stationary solutions to the Boltzmann equation.

265. Wolfgang Dahmen, Angela Kunoth, Karsten Urban: Biorthogonal spline–wavelets on the interval – Stability and moment conditions.

266. Daniela Peterhof, Björn Sandstede, Arnd Scheel: Exponential dichotomies for solitary–wave solutions of semilinear elliptic equations on infinite cylinders.

267. Andreas Rathsfeld: A wavelet algorithm for the solution of a singular integral equation over a smooth two–dimensional manifold.

268. Jörg Schmeling, Serge E. Troubetzkoy: Dimension and invertibility of hyperbolic endomorphisms with singularities.

269. Erwin Bolthausen, Dmitry Ioffe: Harmonic crystal on the wall: a microscopic approach.

270. Nikolai N. Nefedov, Klaus R. Schneider: Delayed exchange of stabilities in singularly perturbed systems.