The Annals of Statistics
1990, Vol. 18, No. 3, 1491–1495

ON PROBABILITIES OF EXCESSIVE DEVIATIONS FOR KOLMOGOROV–SMIRNOV, CRAMÉR–VON MISES AND CHI-SQUARE STATISTICS[1]

BY TADEUSZ INGLOT AND TERESA LEDWINA

Technical University of Wrocław

Let α_n be the classical empirical process and T: $D[0,1] \to R$. Assume T satisfies the Lipschitz condition. Using the Komlós–Major–Tusnády inequality, bounds for $P(T(\alpha_n) \ge x_n\sqrt{n})$ are obtained for every n and $x_n > 0$. Hence expansions for large deviations, as well as some moderate and Cramér-type large-deviations results for $T(\alpha_n)$, are derived.

1. Introduction. Let α_n be the classical empirical process and B be a Brownian bridge on $[0,1]$. Let $T_n = T(\alpha_n)$, where T: $D[0,1] \to R$ is such that $P(T(B) \ge y) > 0$ for all $y > 0$.

ASSUMPTION 1. There exists a constant c, $0 < c < \infty$, such that

$$|T(x) - T(y)| \le c \sup_{0 \le t \le 1} |x(t) - y(t)| \quad \text{for all } x, y \in D[0,1].$$

In Section 2 we give explicit bounds for $P(T_n \ge x_n\sqrt{n})$ valid for every n and $x_n > 0$. The main idea in getting bounds for $P(T_n \ge x_n\sqrt{n})$ is to replace this exact probability by the easier-to-calculate probability $P(T(B) \ge x_n\sqrt{n})$, i.e., by the tail of the asymptotic distribution of T_n. To show that these two excessive probabilities are close enough, the Kómlos–Major–Tusnády (KMT) (1975) inequality is applied. Note that if $x_n\sqrt{n} \to \infty$, then $P(T_n \ge x_n\sqrt{n})$ are called probabilities of excessive deviations and in particular cases $x_n = x$, $x_n = o(n^{-1/3})$ and $x_n = O(n^{-1/2}(\log n)^{1/2})$ are known as large, Cramér-type and moderate deviations, respectively.

ASSUMPTION 2. There exists a constant a, $0 < a < \infty$, such that

$$\log P(T(B) \ge y) = -(a/2)y^2(1 + o(1)) \quad \text{as } y \to \infty.$$

We show that from our bounds many earlier results on probabilities of excessive deviations obtained separately by special methods for each case can be easily derived. New applications for some quadratic statistics are also

Received March 1988; revised August 1989.
[1]Research supported by Grant CPBP 01.02.
AMS 1980 *subject classifications*. Primary 60F10; secondary 62G20, 62E20, 62E15.
Key words and phrases. Excessive deviations, large deviations, moderate deviations, Cramér-type deviations, strong approximation, Cramér–von Mises test, Kolmogorov–Smirnov test, chi-square test, Neyman's test, quadratic statistics.

indicated (cf. Section 3). In Section 2 it is also shown that

$$\lim_{x \to 0} x^{-2} \lim_{n \to \infty} n^{-1} \log P\big(T_n \geq x\sqrt{n}\,\big) = -a/2, \tag{1}$$

provided the limiting large deviations exist for x in a neighbourhood of 0 and

$$\lim_{n \to \infty} \big(nx_n^2\big)^{-1} \log P\big(T_n \geq x_n\sqrt{n}\,\big) = -a/2, \tag{2}$$

provided $x_n \to 0$ and $x_n\sqrt{n} \to \infty$ as $n \to \infty$. Obviously, Assumption 2 can be rewritten in the form

$$\lim_{y \to \infty} y^{-2} \lim_{n \to \infty} \log P(T_n \geq y) = -a/2. \tag{3}$$

This shows that under Assumptions 1 and 2, the order in which the above limits are taken does not matter. This is a crucial point in showing local coincidence of different notions of efficiency. Precise statement of remaining conditions needed to ensure the coincidence can be found in Bahadur (1971), Kallenberg (1983) and Wieand (1976). For some other statistics (e.g., linear rank statistics, L estimates and some U statistics), the equality of the limits (1)–(3) has been observed before, but until now there is no general theory treating the three limits simultaneously. Note that the limit (3) can be calculated using Kallianpur and Oodaira (1978), e.g., while the first general result of the type (1) has been obtained by Kallenberg and Ledwina (1987).

2. Results. For the sake of completeness we recall here the KMT (1975) inequality which is the main technical tool used in this paper. Let $U_1, U_2, \dots$ be i.i.d. random variables uniformly distributed on (0, 1) and let $F_n(t)$ be the empirical distribution function for $U_1, \dots, U_n$. Define

$$\alpha_n(t) = \sqrt{n}\,(F_n(t) - t), \qquad t \in [0, 1].$$

KMT (1975) have shown that there exist a probability space, sequence of processes $\{\alpha_n^*(t)\}$ and Brownian bridges $\{B_n(t)\}$ defined on it such that $\{\alpha_n\} =_D \{\alpha_n^*\}$ and for all n and x,

$$P\Big(\sup_{0 \leq t \leq 1} |\alpha_n^*(t) - B_n(t)| > n^{-1/2}(C \log n + x)\Big) \leq Le^{-lx}, \tag{4}$$

where C, L and l are positive absolute constants. For example, the constants can be chosen as $C = 12$, $L = 2$, and $l = \frac{1}{6}$ [cf. Bretagnolle and Massart (1989)].

Define the function g via $\log P(T(B) \geq y) = -(a/2)y^2(1 + g(y))$, $y \in R$.

THEOREM 1. *Suppose T: $D[0,1] \to R$ satisfies Assumption 1. Then for $K = ac/2l$, arbitrary $d > 1$, arbitrary $1 < p \leq 2$ and all n and $x_n > 0$, it holds*

$$\begin{aligned} P\big(T(\alpha_n) \geq x_n\sqrt{n}\,\big) &\leq \{1 + \exp(-(a/2)nx_n^pR_n)\} \\ &\quad \times \exp\Big\{-(a/2)nx_n^2\big(1 - Kx_n^{p-1}\big)^2(1 + r_n)\Big\}, \end{aligned} \tag{5}$$

and

$$(6)\quad \begin{aligned} P\big(T(\alpha_n) \ge x_n\sqrt{n}\big) &\ge \{1 - \exp(-(a/2)nx_n^pL_n)\} \\ &\quad \times \exp\Big\{-(a/2)nx_n^2\big(1 + dKx_n^{p-1}\big)^2(1 + l_n)\Big\}, \end{aligned}$$

where

$$R_n = 1 - x_n^{2-p}\big(1 - Kx_n^{p-1}\big)^2(1 + r_n) - 2(lC \log n + \log L)/anx_n^p$$

while

$$L_n = d - x_n^{2-p}\big(1 + dKx_n^{p-1}\big)^2(1 + l_n) - 2(lC \log n + \log L)/anx_n^p,$$

where $r_n = g(x_n(1 - Kx_n^{p-1})\sqrt{n})$, *while* $l_n = g(x_n(1 + dKx_n^{p-1})\sqrt{n})$.

PROOF. Put $H_n = \sup_{0 \le t \le 1}|\alpha_n^*(t) - B_n(t)|$. Since $\{\alpha_n^*\} =_D \{\alpha_n\}$ and $T(B_n) =_D T(B)$, it follows by (4) and Assumption 1 that

$$\begin{aligned} P\big(T(\alpha_n) \ge x_n\sqrt{n}\big) &\le P\big(T(B_n) \ge x_n\big(1 - Kx_n^{p-1}\big)\sqrt{n}\big) + P\big(H_n \ge c^{-1}Kx_n^p\sqrt{n}\big) \\ &\le \exp\Big\{-(a/2)x_n^2\big(1 - Kx_n^{p-1}\big)^2 n(1 + r_n)\Big\} \\ &\quad + L\exp\big\{-l\big(c^{-1}Kx_n^pn - C\log n\big)\big\}. \end{aligned}$$

Hence, by the definition of K and R_n, (5) follows. Analogously,

$$\begin{aligned} P\big(T(\alpha_n) \ge x_n\sqrt{n}\big) &\ge P\big(T(B_n) \ge x_n\big(1 + dKx_n^{p-1}\big)\sqrt{n}\big) \\ &\quad - P\big(H_n \ge c^{-1}dKx_n^p\sqrt{n}\big) \end{aligned}$$

yields (6). □

REMARK 1. Using some bounds for g, one can get from Theorem 1 explicit bounds for $P(T(\alpha_n) \ge x_n\sqrt{n})$ (cf. Section 3).

REMARK 2. Suppose the inner limits in (1) exist for small $x > 0$ and Assumption 2 holds. Applying Theorem 1 for arbitrary $d > 1$, $p = 2$ and fixed $x_n = x$ satisfying $0 < x < (\sqrt{d} - 1)/dK$, one gets

$$\begin{aligned} -(a/2)x^2(1 + dKx)^2 \le \lim_{n\to\infty} n^{-1}\log P\big(T(\alpha_n) \ge x\sqrt{n}\big) \\ \le -(a/2)x^2(1 - Kx)^2, \end{aligned}$$

and consequently

$$\lim_{x\to 0}\lim_{n\to\infty}\big(nx^2\big)^{-1}\log P\big(T(\alpha_n) \ge x\sqrt{n}\big) = -a/2.$$

REMARK 3. Suppose Assumption 2 holds. If $x_n \to 0$ and $x_n\sqrt{n} \to \infty$ as $n \to \infty$, then applying Theorem 1 for arbitrary $d > 1$ and $1 < p < 2$ one gets

$$\lim_{n\to\infty} (nx_n^2)^{-1} \log P\big(T(\alpha_n) \geq x_n\sqrt{n}\big) = -a/2.$$

REMARK 4. Let T be a measurable and finite seminorm on $D[0,1]$. Then by Kallianpur and Oodaira (1978), e.g., Assumption 2 is satisfied. In particular, Assumption 2 holds in all examples that follow.

3. Examples. By Remark 4, Theorem 1 is applicable for

$$T_{\mathrm{KS}}(x) = \sup_{0\leq t\leq 1} |x(t)|w(t),$$

where w is a nonnegative and bounded weight,

$$T_{\mathrm{CvM}}(x) = \left\{\int_0^1 |x(t)|^r w(t)\,dt\right\}^{1/r},$$

where w is a nonnegative, integrable weight and $r \geq 1$ and for

$$T_\chi(x) = \left\{\sum_{i=1}^{k} (x(a_i) - x(a_{i-1}))^2/(a_i - a_{i-1})\right\}^{1/2},$$

where $0 = a_0 < \cdots < a_k = 1$. Note that the limit (1) for T_{KS} and T_{CvM} for some weight functions have been calculated earlier; among others, by Bahadur (1971), Mogul'skii (1977) and Nikitin (1979, 1980). Moderate deviations for T_{CvM} and some excessive deviations for T_{KS} have been derived by Rubin and Sethuraman (1965) and Borovkov and Sycheva (1968). Recently, some excessive deviations for T_χ have been extensively studied in Kallenberg (1985). In particular, he derived some bounds for $P(T_\chi \geq x_n\sqrt{n})$. Let us very briefly discuss an application of our result in this case. Since $T_\chi^2(B)$ has the chi-square distribution with $k-1$ degrees of freedom, then for $y > k+1$, it holds that

$$2A(k)y^{(k-3)/2}e^{-y/2} \leq P\big(T_\chi^2(B) \geq y\big) \leq A(k)y^{(k-1)/2}e^{-y/2},$$

where $A(k) = \{2^{(k-1)/2}\Gamma((k-1)/2)\}^{-1}$. Hence, for $p_i = a_i - a_{i-1}$ satisfying, e.g., $p_i \geq \beta/k$, $i = 1,\dots,k$, for some $\beta > 0$, by Remark 1 one gets some explicit lower and upper bounds for $P(T_\chi^2(\alpha_n) \geq nx_n^2)$, valid for $nx_n^2 > k+1$, all $k = 2,3,\dots$ and all n [cf. Theorem 2.3 of Kallenberg (1985)].

Finally, consider a class of quadratic goodness-of-fit statistics introduced by Neuhaus (1988) and defined as

$$\sum_{k=1}^{\infty} \lambda_k(\psi_k', \alpha_n)^2, \tag{7}$$

where $\psi_k(t) = \sqrt{2}\cos \pi kt$, $t \in (0,1)$, are eigenfunctions and λ_k are eigenvalues of a kernel on $L_2(0,1)$ while $(\cdot,\cdot)$ denotes the scalar product in $L_2(0,1)$. Neuhaus recommended for practical use the statistic (7) with $\lambda_k = [(0.1\pi k)^{-1}\sin(0.1\pi k)]^4$, $k \geq 1$. Generally, suppose that λ_k's in (7) satisfy

$c_0^2 = \max\{\pi^2 k^2 \lambda_k,\ k \geq 1\} < \infty$. Since $\psi_k'(t) = -\pi k\sqrt{2}\sin \pi kt$, by the Bessel inequality, the functional

$$T(x) = \left\{ \sum_{k=1}^{\infty} \lambda_k (\psi_k', x)^2 \right\}^{1/2},$$

satisfies Assumption 1 with $c = c_0$. On the other hand, the smooth test of fit statistic introduced by Neyman (1937) can be represented also in the form (7) with ψ_k replaced by transformed Legendre polynomials on $(0, 1)$ and $\lambda_k = 1$ for $k = 1, \dots, k_0$ and $\lambda_k = 0$, otherwise. So, this statistic satisfies Assumption 1 also.

Acknowledgments. The second author is very grateful to W. C. M. Kallenberg and Ya. Yu. Nikitin, who kindly sent her some reprints of their papers. We have both benefitted from suggestions of a referee concerning the presentation of our results.

REFERENCES

BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics*. SIAM, Philadelphia.

BOROVKOV, A. A. and SYCHEVA, N. M. (1968). On asymptotically optimal nonparametric criteria. *Theory Probab. Appl.* **13** 359–393.

BRETAGNOLLE, J. and MASSART, P. (1989). Hungarian constructions from the nonasymptotic viewpoint. *Ann. Probab.* **17** 239–256.

KALLENBERG, W. C. M. (1983). Intermediate efficiency, theory and examples. *Ann. Statist.* **11** 170–182.

KALLENBERG, W. C. M. (1985). On moderate and large deviations in multinomial distributions. *Ann. Statist.* **13** 1554–1580.

KALLENBERG, W. C. M. and LEDWINA, T. (1987). On local and nonlocal measures of efficiency. *Ann. Statist.* **15** 1401–1420.

KALLIANPUR, G. and OODAIRA, H. (1978). Freidlin–Wentzel type estimates for abstract Wiener spaces. *Sankhyā Ser. A* **40** 116–137.

KÓMLOS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent R. V.'s and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131.

MOGUL'SKII, A. A. (1977). Remarks on large deviations for the ω^2 statistic. *Theory Probab. Appl.* **12** 166–171.

NEUHAUS, G. (1988). A class of quadratic goodness of fit tests. Addendum to Local asymptotics for linear rank statistics with estimated score functions. *Ann. Statist.* **16** 1342–1343.

NEYMAN, J. (1937). "Smooth tests" for goodness of fit. *Scand. Actuar. J.* **20** 149–199.

NIKITIN, YA. YU. (1979). Large deviations and asymptotic efficiency of integral-type statistics. I. *Zap. Naucn. Sem. Leningrad. Otdel. Mat. Inst. Steklov* **85** 175–187 (in Russian).

NIKITIN, YA. YU. (1980). Large deviations and asymptotic efficiency of integral-type statistics. II. *Zap. Naucn. Sem. Leningrad. Otdel. Mat. Inst. Steklov* **97** 151–175 (in Russian).

RUBIN, H. and SETHURAMAN, J. (1965). Probability of moderate deviations. *Sankhyā Ser. A* **27** 325–346.

WIEAND, H. S. (1976). A condition under which the Pitman and Bahadur approaches to efficiency coincide. *Ann. Statist.* **4** 1003–1011.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCŁAW
50–370 WROCŁAW
WYBRZEZE WYSPIAŃSKIEGO 27
POLAND