A note on a maximal Bernstein inequality

PÉTER KEVEI[1] and DAVID M. MASON[2]

[1]*CIMAT, Callejón Jalisco S/N, Mineral de Valenciana, Guanajuato 36240, Mexico*
E-mail: kevei@ cimat.mx
[2]*Statistics Program, University of Delaware, 213 Townsend Hall Newark, DE 19716, USA*
E-mail: davidm@udel.edu

Dedicated to the memory of Sándor Csörgő

We show somewhat unexpectedly that whenever a general Bernstein-type maximal inequality holds for partial sums of a sequence of random variables, a maximal form of the inequality is also valid.

Keywords: Bernstein inequality; dependent sums; maximal inequality; mixing; partial sums

1. Introduction and statement of main result

Let $X_1, X_2, \ldots,$ be a sequence of independent random variables such that for all $i \geq 1$, $EX_i = 0$ and for some $\kappa > 0$ and $v > 0$ for integers $m \geq 2$, $E|X_i|^m \leq vm!\kappa^{m-2}/2$. The classic Bernstein inequality (cf. [13], page 855) says that, in this situation, for all $n \geq 1$ and $t \geq 0$,

$$\mathbf{P}\left\{\left|\sum_{i=1}^{n} X_i\right| > t\right\} \leq 2\exp\left\{-\frac{t^2}{2vn + 2\kappa t}\right\}.$$

Moreover (cf. [12], Theorem B.2), its maximal form also holds; that is, we have

$$\mathbf{P}\left\{\max_{1 \leq j \leq n}\left|\sum_{i=1}^{j} X_i\right| > t\right\} \leq 2\exp\left\{-\frac{t^2}{2vn + 2\kappa t}\right\}.$$

It turns out that, under a variety of assumptions, a sequence of not necessarily independent random variables $X_1, X_2, \ldots,$ will satisfy a generalized Bernstein-type inequality of the following form: For suitable constants $A > 0$, $a > 0$, $b \geq 0$ and $0 < \gamma < 2$ for all $m \geq 0$, $n \geq 1$ and $t \geq 0$,

$$\mathbf{P}\{|S(m+1, m+n)| > t\} \leq A\exp\left\{-\frac{at^2}{n + bt^\gamma}\right\}, \tag{1.1}$$

where, for any choice of $1 \leq k \leq l < \infty$, we denote the partial sum $S(k, l) = \sum_{i=k}^{l} X_i$. Here are some examples.

Example 1. Let $X_1, X_2, \ldots,$ be a stationary sequence satisfying $EX_1 = 0$ and $\mathrm{Var}\, X_1 = 1$. For each integer $n \geq 1$ set $S_n = X_1 + \cdots + X_n$ and $B_n^2 = \mathrm{Var}(S_n)$. Assume that for some $\sigma_0^2 > 0$ we have $B_n^2 \geq \sigma_0^2 n$ for all $n \geq 1$. Statulevičius and Jakimavičius [15] and Saulis and Statulevičius [14] prove that the partial sums satisfy (1.1) with constants depending on a Bernstein

condition on the moments of X_1 and the particular mixing condition that the sequence may fulfill. In fact, all values of $1 \leq \gamma < 2$ are attainable. Their Bernstein-type inequalities are derived via a result of [1] relating cumulant behavior to tail behavior, which says that for an arbitrary random variable ξ with expectation 0, whenever there exist $\gamma \geq 0$, $H > 0$ and $\Delta > 0$ such that its cumulants $\Gamma_k(\xi)$ satisfy $|\Gamma_k(\xi)| \leq (k!/2)^{1+\gamma} H / \Delta^{k-2}$ for $k = 2, 3, \ldots$, then for all $x \geq 0$

$$\mathbf{P}\{\pm\xi > x\} \leq \exp\left\{-\frac{x^2}{2(H + (x/\Delta^{1/(1+2\gamma)})^{(1+2\gamma)/(1+\gamma)})}\right\}. \tag{1.2}$$

In Example 1, $\xi = S_n/B_n$ and $\Delta = d\sqrt{n}$ for some $d > 0$.

Example 2. Doukhan and Neumann [4] have shown, using the result in (1.2), that if a sequence of mean zero random variables $X_1, X_2, \ldots$, satisfies a general covariance condition, then the partial sums satisfy (1.1). Refer to their Theorem 1 and Remark 2, and also see [8].

Example 3. Assume that $X_1, X_2, \ldots$, is a strong mixing sequence with mixing coefficients $\alpha(n)$, $n \geq 1$, satisfying for some $c > 0$, $\alpha(n) \leq \exp(-2cn)$. Also assume that $EX_i = 0$ and for some $M > 0$ for all $i \geq 1$, $|X_i| \leq M$. Theorem 2 of Merlevéde, Peligrad and Rio [9] implies that for some constant $C > 0$ for all $t \geq 0$ and $n \geq 1$,

$$\mathbf{P}\{|S_n| > t\} \leq \exp\left(-\frac{Ct^2}{nv^2 + M^2 + tM(\log n)^2}\right), \tag{1.3}$$

with $S_n = \sum_{i=1}^n X_i$ and where $v^2 = \sup_{i>0}(\text{Var}(X_i) + 2\sum_{j>i} |\text{cov}(X_i, X_j)|) > 0$.

To see how the last example satisfies (1.1), notice that for any $0 < \eta < 1$ there exists a $D_1 > 0$ such that for all $t \geq 0$ and $n \geq 1$,

$$nv^2 + M^2 + tM(\log n)^2 \leq n(v^2 + M^2) + D_1 t^{1+\eta}. \tag{1.4}$$

Thus (1.1) holds with $\gamma = 1 + \eta$ for suitable $A > 0$, $a > 0$ and $b \geq 0$.

For any choice of $1 \leq i \leq j < \infty$ define

$$M(i, j) = \max\{|S(i, i)|, \ldots, |S(i, j)|\}.$$

We shall show, somewhat unexpectedly, that if a sequence of random variables $X_1, X_2, \ldots$, satisfies a Bernstein-type inequality of the form (1.1), then, without any additional assumptions, a modified version of it also holds for $M(1 + m, n + m)$ for all $m \geq 0$ and $n \geq 1$.

Theorem 1. *Assume that, for constants $A > 0$, $a > 0$, $b \geq 0$ and $\gamma \in (0, 2)$, inequality (1.1) holds for all $m \geq 0, n \geq 1$ and $t \geq 0$. Then for every $0 < c < a$ there exists a $C > 0$ depending only on A, a, b and γ such that for all $n \geq 1$, $m \geq 0$ and $t \geq 0$,*

$$\mathbf{P}\{M(m + 1, m + n) > t\} \leq C \exp\left\{-\frac{ct^2}{n + bt^\gamma}\right\}. \tag{1.5}$$

Remark 1. Notice that though $c < a$, c can be chosen arbitrarily close to a.

Remark 2. Theorem 1 was motivated by Theorem 2.2 of Móricz, Serfling and Stout [11], who showed that whenever for a suitable positive function $g(i, j)$ of $(i, j) \in \{1, 2, \ldots\} \times \{1, 2, \ldots\}$, positive function $\phi(t)$ defined on $(0, \infty)$ and constant $K > 0$, for all $1 \le i \le j < \infty$ and $t > 0$,

$$\mathbf{P}\{|S(i, j)| > t\} \le K \exp\{-\phi(t)/g(i, j)\},$$

then there exist constants $0 < c < 1$ and $C > 0$ such that for all $n \ge 1$ and $t > 0$,

$$\mathbf{P}\{M(1, n) > t\} \le C \exp\{-c\phi(t)/g(1, n)\}.$$

Earlier, Móricz [10] proved that in the special case when $\phi(t) = t^2$ one can choose $c < 1$ arbitrarily close to 1 by making $C > 0$ large enough. This inequality is clearly not applicable to obtain a maximal form of the generalized Bernstein inequality.

Remark 3. We do not know whether there exist examples for which (1.1) holds for some $0 < \gamma < 1$ and $b > 0$. However, since the proof of our theorem remains valid in this case, we shall keep it in the statement.

Remark 4. The version of Theorem 1 obtained by replacing everywhere $|S(m + 1, m + n)|$ by $S(m + 1, m + n)$ and $M(m + 1, m + n)$ by $M^+(m + 1, m + n) = \max_{m+1 \le j \le n+m}(S(m + 1, j) \vee 0)$ remains true with little change in the proof.

Remark 5. Theorem 1 also remains valid for sums of Banach space valued random variables with absolute value $|\cdot|$ replaced by norm $\|\cdot\|$.

Remark 6. In statistics, maximal exponential inequalities are crucial tools to determine the exact rate of almost sure pointwise and uniform consistency of kernel estimators of the density function and the regression function. The literature in this area is huge. See, for instance, [2,3,5–7,16] and the references therein. These results only treat the case of i.i.d. observations. Dependent versions of our maximal Bernstein inequalities should be useful to determine exact rates of almost sure consistency of kernel estimators based on data that possess a certain dependence structure. In fact, some work in this direction has already been accomplished in Section 4.2 of [4]. To carry out such an application in the present paper is well beyond its scope.

Theorem 1 leads to the following bounded law of the iterated logarithm.

Corollary 1. *Under the assumptions of Theorem 1, with probability* 1,

$$\limsup_{n \to \infty} \frac{|S(1, n)|}{\sqrt{n \log \log n}} \le \frac{1}{\sqrt{a}}. \tag{1.6}$$

Remark 7. In general, one cannot replace "$\le$" by "$=$" in (1.6). To see this, let $Y, Z_1, Z_2, \ldots$ be a sequence of independent random variables such that Y takes on the value 0 or 1 with probability $1/2$ and $Z_1, Z_2, \ldots$ are independent standard normals. Now define $X_i = YZ_i$, $i = 1, 2, \ldots$. It

is easily checked that assumption (1.1) is satisfied with $A = 2$, $a = 1/2$, $b = 0$ and $\gamma = 1$. When $Y = 1$ the usual law of the iterated logarithm gives with probability 1,

$$\limsup_{n \to \infty} |S(1, n)| / \sqrt{n \log \log n} = \sqrt{2} = 1/\sqrt{a}, \qquad (1.7)$$

whereas, when $Y = 0$ the above limsup is obviously 0. This agrees with Corollary 1, which says that with probability 1 the limsup is $\leq \sqrt{2}$. However, we see that with probability $1/2$ it equals $\sqrt{2}$ and with probability $1/2$ it equals 0.

Theorem 1 is proved in Section 2 and the proof of Corollary 1 is given in Section 3.

2. Proof of theorem

The case $b = 0$ is a special case of Theorem 1 of [10]. Therefore we shall always assume that $b > 0$. Choose any $0 < c < a$. We prove our theorem by induction on n. Notice that by the assumption, for any integer $n_0 \geq 1$ we may choose $C > An_0$ to make the statement true for all $1 \leq n \leq n_0$. This remark will be important, because at some steps of the proof we assume that n is large enough. Also, since the constants A, a, b and γ in (1.1) are independent of m, we can assume $m = 0$ without loss of generality in our proof.

Assume the statement holds up to some $n \geq 2$. (The constant C will be determined in the course of the proof.)

Case 1: Fix a $t > 0$ for which

$$t^\gamma \leq \alpha n \qquad (2.1)$$

for some $0 < \alpha < 1$ to be specified later. (In any case, we assume that $\alpha n \geq 1$.) Using an idea of [11], we may write for arbitrary $1 \leq k \leq n$, $0 \leq q \leq 1$ and $p + q = 1$ the inequality

$$\mathbf{P}\{M(1, n + 1) > t\}$$
$$\leq \mathbf{P}\{M(1, k) > t\} + \mathbf{P}\{|S(1, k)| > pt\} + \mathbf{P}\{M(k + 1, n + 1) > qt\}.$$

Let

$$u = \frac{n + t^\gamma b(q^\gamma - q^2)}{1 + q^2}.$$

Notice that

$$\frac{t^2}{u + bt^\gamma} = \frac{q^2 t^2}{n - u + bq^\gamma t^\gamma}. \qquad (2.2)$$

Set

$$k = \lceil u \rceil. \qquad (2.3)$$

Using the induction hypothesis and (1.1) we obtain

$$\mathbf{P}\{M(1, n+1) > t\}$$

$$\leq C \exp\left\{-\frac{ct^2}{k+bt^\gamma}\right\} + A \exp\left\{-\frac{ap^2t^2}{k+bp^\gamma t^\gamma}\right\} + C \exp\left\{-\frac{cq^2t^2}{n-k+bq^\gamma t^\gamma}\right\}. \tag{2.4}$$

Notice that we chose k to make the first and third terms in the right-hand side of (2.4) almost equal, and since by (2.3)

$$\frac{t^2}{k+bt^\gamma} \leq \frac{q^2t^2}{n-k+bq^\gamma t^\gamma},$$

the first term is greater than or equal to the third.

First we handle the second term in (2.4), showing that for $0 \leq t \leq (\alpha n)^{1/\gamma}$,

$$\exp\left\{-\frac{ap^2t^2}{k+bp^\gamma t^\gamma}\right\} \leq \exp\left\{-\frac{ct^2}{n+1+bt^\gamma}\right\}.$$

For this we need to verify that for $0 \leq t \leq (\alpha n)^{1/\gamma}$,

$$\frac{ap^2}{k+bp^\gamma t^\gamma} > \frac{c}{n+1+bt^\gamma}, \tag{2.5}$$

which is equivalent to

$$ap^2(n+1+bt^\gamma) > c(k+bp^\gamma t^\gamma).$$

Using that

$$k = \lceil u \rceil \leq u+1 = 1 + \frac{1}{1+q^2}[n+b(q^\gamma-q^2)t^\gamma],$$

it is enough to show

$$n\left(ap^2 - \frac{c}{1+q^2}\right) + t^\gamma\left(ap^2b - cbp^\gamma - \frac{cb}{1+q^2}(q^\gamma-q^2)\right) + ap^2 - c > 0.$$

Note that if the coefficient of n is positive, then we can choose α in (2.1) small enough to make the above inequality hold, even if the coefficient of t^γ is negative. So in order to guarantee (2.5) (at least for large n) we only have to choose the parameter p so that $ap^2 - c > 0$ – which implies that

$$ap^2 - \frac{c}{1+q^2} > 0 \tag{2.6}$$

holds – and then select α small enough.

Next we treat the first and third terms in (2.4). By the remark above, it is enough to handle the first term. Let us examine the ratio of $C \exp\{-ct^2/(k+bt^\gamma)\}$ and $C \exp\{-ct^2/(n+1+bt^\gamma)\}$.

Notice again that since $u + 1 \geq k$,

$$
\begin{aligned}
n + 1 - k \geq n - u = n &- \frac{n + b(q^\gamma - q^2)t^\gamma}{1 + q^2} \\
&= \frac{q^2 n - b(q^\gamma - q^2)t^\gamma}{1 + q^2} \\
&\geq n \frac{q^2 - \alpha b(q^\gamma - q^2)}{1 + q^2} \\
&=: c_1 n.
\end{aligned}
$$

At this point we need that $0 < c_1 < 1$. Thus we choose α small enough so that

$$
q^2 - \alpha b(q^\gamma - q^2) > 0. \tag{2.7}
$$

Also, using $t \leq (\alpha n)^{1/\gamma}$, we get the bound

$$
(n + 1 + bt^\gamma)(k + bt^\gamma) \leq n^2(1 + \alpha b)^2 =: c_2 n^2,
$$

which holds if n is large enough. Therefore, we obtain for the ratio

$$
\exp\left\{-ct^2\left(\frac{1}{k + bt^\gamma} - \frac{1}{n + 1 + bt^\gamma}\right)\right\} \leq \exp\left\{-\frac{cc_1 t^2}{c_2 n}\right\} \leq e^{-1},
$$

whenever $cc_1 t^2/(c_2 n) \geq 1$, that is, $t \geq \sqrt{c_2 n/(cc_1)}$. Substituting back into (2.4), for $t \geq \sqrt{c_2 n/(cc_1)}$ and $t \leq (\alpha n)^{1/\gamma}$ we obtain

$$
\mathbf{P}\{M(1, n + 1) > t\}
$$

$$
\leq \left(\frac{2}{e}C + A\right)\exp\{-ct^2/(n + 1 + bt^\gamma)\} \leq C\exp\{-ct^2/(n + 1 + bt^\gamma)\},
$$

where the last inequality holds for $C > Ae/(e - 2)$.

Next assume that $t < \sqrt{c_2 n/(cc_1)}$. In this case, choosing C large enough, we can make the bound > 1, namely

$$
C\exp\left\{-\frac{ct^2}{n + 1 + bt^\gamma}\right\} \geq C\exp\left\{-\frac{cc_2 n}{cc_1 n}\right\} = Ce^{-c_2/c_1} \geq 1,
$$

if $C > e^{c_2/c_1}$.

Case 2: Now we must handle the case $t > (\alpha n)^{1/\gamma}$. Here we apply the inequality

$$
\mathbf{P}\{M(1, n + 1) > t\} \leq \mathbf{P}\{M(1, n) > t\} + \mathbf{P}\{|S(1, n + 1)| > t\}.
$$

Using assumption (1.1) and the induction hypothesis, we have

$$
\mathbf{P}\{M(1, n + 1) > t\} \leq C\exp\left\{-\frac{ct^2}{n + bt^\gamma}\right\} + A\exp\left\{-\frac{at^2}{n + 1 + bt^\gamma}\right\}.
$$

We will show that the right-hand side $\leq C \exp\{-ct^2/(n+1+bt^\gamma)\}$. For this it is enough to prove

$$\exp\left\{-ct^2\left(\frac{1}{n+bt^\gamma} - \frac{1}{n+1+bt^\gamma}\right)\right\} + \frac{A}{C}\exp\left\{-\frac{t^2(a-c)}{n+1+bt^\gamma}\right\} \leq 1. \qquad (2.8)$$

First assume that $\gamma \leq 1$. Using the bound following from $t > (\alpha n)^{1/\gamma}$, we get

$$\frac{t^2}{(n+bt^\gamma)(n+bt^\gamma+1)} \geq \frac{t^2}{(\alpha^{-1}+b)(2\alpha^{-1}+b)t^{2\gamma}} =: t^{2-2\gamma}c_3 \geq c_3.$$

We have that the right-hand side of (2.8) for $a \geq c$ is less than

$$e^{-cc_3} + \frac{A}{C} \leq 1$$

for C large enough.

For $1 < \gamma < 2$ we have to use a different argument. For t large enough (i.e., for n large enough, since $t > (\alpha n)^{1/\gamma}$) we have

$$\exp\left\{-\frac{ct^2}{(n+bt^\gamma)(n+bt^\gamma+1)}\right\} \leq \exp\{-cc_3t^{2-2\gamma}\} \leq 1 - \frac{cc_3t^{2-2\gamma}}{2}.$$

We also have for $C > A$,

$$\frac{A}{C}\exp\left\{-\frac{t^2(a-c)}{n+1+bt^\gamma}\right\} \leq \exp\left\{-t^{2-\gamma}\frac{a-c}{2\alpha^{-1}+b}\right\}.$$

It is clear that since $a > c$, for t large enough, that is, for n large enough,

$$\frac{cc_3t^{2-2\gamma}}{2} > \exp\left\{-t^{2-\gamma}\frac{a-c}{2\alpha^{-1}+b}\right\}.$$

The proof is complete.

3. Proof of corollary

Choose any $\lambda > 1$ and set $m_r = \lceil\lambda^r\rceil$ for $r = 1, 2, \dots$. Now, using inequality (1.5), we get

$$\mathbf{P}\{M(1, m_r) > \sqrt{c^{-1}m_{r+1}\log\log m_r}\}$$

$$\leq C\exp\left\{-\frac{m_{r+1}\log\log m_r}{m_r + b(c^{-1}m_{r+1}\log\log m_r)^{\gamma/2}}\right\}.$$

Since as $r \to \infty$

$$\frac{m_{r+1}\log\log m_r}{m_r + b(c^{-1}m_{r+1}\log\log m_r)^{\gamma/2}} = (1 + o(1))\lambda\log r,$$

it is readily checked that for r_0 large enough so that $\log \log m_{r_0} > 0$,

$$\sum_{r=r_0}^{\infty} \mathbf{P}\left\{ M(1, m_r) > \sqrt{c^{-1} m_{r+1} \log \log m_r} \right\} < \infty$$

and thus, since $m_{r+1}/m_r = \lambda + o(1)$, we get by the Borel–Cantelli lemma that with probability 1

$$\limsup_{r \to \infty} \frac{M(1, m_r)}{\sqrt{m_r \log \log m_r}} \leq \sqrt{\lambda c^{-1}}. \tag{3.1}$$

Next we see that for all $r \geq r_0$

$$\max_{m_r \leq n < m_{r+1}} \frac{|S(1, n)|}{\sqrt{n \log \log n}} \leq \frac{M(1, m_{r+1})}{\sqrt{m_r \log \log m_r}}.$$

Thus by (3.1), with probability 1,

$$\limsup_{r \to \infty} \max_{m_r \leq n < m_{r+1}} \frac{|S(1, n)|}{\sqrt{n \log \log n}}$$

$$\leq \limsup_{r \to \infty} \frac{M(1, m_{r+1})}{\sqrt{m_r \log \log m_r}}$$

$$= \limsup_{r \to \infty} \frac{M(1, m_{r+1})}{\sqrt{m_{r+1} \log \log m_{r+1}}} \frac{\sqrt{m_{r+1} \log \log m_{r+1}}}{\sqrt{m_r \log \log m_r}} \leq \lambda \sqrt{c^{-1}}.$$

Hence, since $\lambda > 1$ can be chosen arbitrarily close to 1 and $c < a$ arbitrarily close to a, we have proved (1.6).

Acknowledgements

References

[1] Bentkus, R. and Rudzkis, R. (1980). On exponential estimates of the distribution of random variables. *Lithuanian Math. J.* **20** 15–30 (in Russian). MR0575427

[2] Deheuvels, P. and Mason, D.M. (1992). Functional laws of the iterated logarithm for the increments of empirical and quantile processes. *Ann. Probab.* **20** 1248–1287. MR1175262

[3] Deheuvels, P. and Mason, D.M. (1994). Functional laws of the iterated logarithm for local empirical processes indexed by sets. *Ann. Probab.* **22** 1619–1661. MR1303659

[4] Doukhan, P. and Neumann, M.H. (2007). Probability and moment inequalities for sums of weakly dependent random variables, with applications. *Stochastic Process. Appl*. **117** 878–903. MR2330724

[5] Einmahl, U. and Mason, D.M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab*. **13** 1–37. MR1744994

[6] Einmahl, U. and Mason, D.M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist*. **33** 1380–1403. MR2195639

[7] Giné, E. and Guillou, A. (2002). Rates of strong uniform consistency for multivariate kernel density estimators. En l'honneur de J. Bretagnolle, D. Dacunha-Castelle, I. Ibragimov. *Ann. Inst. H. Poincaré Probab. Statist*. **38** 907–921. MR1955344

[8] Kallabis, R. and Neumann, M.H. (2006). An exponential inequality under weak dependence. *Bernoulli* **12** 333–350. MR2218558

[9] Merlevède, F., Peligrad, M. and Rio, E. (2009). Bernstein inequality and moderate deviations under strong mixing conditions. In *High Dimensional Probability V: The Luminy Volume* (C. Houdré, V. Koltchinskii, D.M. Mason and M. Peligrad, eds.) 273–292. Beachwood, OH: IMS.

[10] Móricz, F.A. (1979). Exponential estimates for the maximum of partial sums. I. Sequences of rv's. Special issue dedicated to George Alexits on the occasion of his 80th birthday. *Acta Math. Acad. Sci. Hungar*. **33** 159–167. MR0515130

[11] Móricz, F.A., Serfling, R.J. and Stout, W.F. (1982). Moment and probability bounds with quasisuper-additive structure for the maximum partial sum. *Ann. Probab*. **10** 1032–1040. MR0672303

[12] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants. (French) Mathématiques & Applications (Berlin)* **31**. Berlin: Springer. MR2117923

[13] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. New York: Wiley. MR0838963

[14] Saulis, L. and Statulevičius, V.A. (1991). *Limit Theorems for Large Deviations*. Dordrecht: Kluwer. MR1171883

[15] Statulevičius, V.A. and Jakimavičius, D.A. (1988). Estimates for semiinvariants and centered moments of stochastic processes with mixing. I. *Litovsk. Mat. Sb*. **28** 112–129; translation in *Lithuanian Math. J*. **28** 67–80. MR0949647

[16] Stute, W. (1982). A law of the logarithm for kernel density estimators. *Ann. Probab*. **10** 414–422. MR0647513