ON THE NORMAL APPROXIMATION FOR SUMS OF INDEPENDENT VARIABLES

BY

L. LE CAM*

TECHNICAL REPORT NO. 42
JANUARY 1985

*RESEARCH PARTIALLY SUPPORTED BY
NATIONAL SCIENCE FOUNDATION GRANT MCS80-02698

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

1. Introduction. In his famous monograph [1], Paul Lévy states a result that gives necessary and sufficient conditions for a sum $S = \sum_j X_j$ of independent variables to have an approximately normal distribution. The only condition imposed is the independence. There are no negligibility requirements. Lévy's statement and the accompanying arguments have sometimes been criticized as non-rigorous or too vague. Actually the statement makes perfect sense intuitively and the argument can be made rigorous. The present paper is an attempt to a rigorous presentation, following almost exactly the steps indicated by Lévy. A rigorous presentation, for the case where variances exist and converge to the variance of the limiting distribution, was given by Zolotarev in [5]. The general Normal case is covered by a paper of Mačys [3]. Zolotarev treats a more general problem in [6]. However, the techniques of proof are different from those used by Lévy. Here we consider the general case. In some lemmas we have used results that are more precise than those available to Lévy in 1937. However, these are inessential modifications of the main arguments. After some preliminaries, the theorem is stated in Section 2 below. Section 3 gives the proof of a number of auxiliary lemmas. Section 4 concludes the proof of the theorem.

2. Notation and results. Let X_j; $j = 1,2,\ldots$ be a finite or infinite sequence of independent random variables and let $S = \sum_j X_j$ be their sum, assumed to exist if the sequence is infinite. The problem raised by Lévy is to find out conditions that imply that the distribution $\mathscr{L}(S)$ of S is close to a Gaussian $\mathscr{N}(\mu,\sigma^2)$ distribution. For this to make sense one has to introduce distances between distributions. We shall use two of them: the Lévy distance $\lambda(P,Q)$ between the probability measures P and Q and the Kolmogorov vertical distance $\rho(P,Q)$. The distance $\lambda(P,Q)$ is the infimum of the numbers ε such that

$$-\varepsilon + P\{(-\infty,x-\varepsilon]\} \leq Q\{(-\infty,x]\} \leq P\{(-\infty,x+\varepsilon]\} + \varepsilon$$

for all values of x. The distance ρ is given by $\rho(P,Q) = \sup_x |P\{(-\infty,x]\} - Q\{(-\infty,x]\}|$. The two are related by the inequalities

$$\lambda \leq \rho \leq \lambda + C(\lambda)$$

where $C(\lambda) = \min\{C_P(\lambda), C_Q(\lambda)\}$ with for instance $C_P(\lambda) = \sup_x P\{[x,x+\lambda]\}$. We shall be concerned here with conditions that are necessary and sufficient for approximability of $\mathscr{L}(S)$ by $\mathscr{N}(\mu,\sigma^2)$ in the sense of the Kolmogorov distance. This distance is invariant by one to one monotone increasing transformations. Hence the size of σ^2 is unimportant. One can standardize and look for approximations by $\mathscr{N}(0,1)$.

We shall often write PQ for the convolution of two measures P and Q and write G_σ for the Gaussian measure $\mathscr{N}(0,\sigma^2)$.

One of the main results we need is a theorem conjectured by Lévy and proved by Cramér as follows.

THEOREM 1. *Let the convolution* PQ *be* $\mathcal{N}(0,1)$ *then there are numbers* μ_i, σ_i, $i = 1, 2$ *such that* $P = \mathcal{N}(\mu_1, \sigma_1^2)$ *and* $Q = \mathcal{N}(\mu_2, \sigma_2^2)$ *with* $\mu_1 + \mu_2 = 0$ *and* $\sigma_1^2 + \sigma_2^2 = 1$.

An easy consequence of Theorem 1 is a result proved by Lévy in [2].

THEOREM 2. *There is a function* g *on* $[0,1]$ *to* $[0,1]$ *with the following properties:*

i) $g(\delta) \geq \delta$ *and* $g(\delta)$ *decreases to zero as* δ *decreases to zero.*

ii) *If* $G = \mathcal{N}(0,1)$ *and* $\lambda(PQ,G) \leq \delta$ *then there is a* $G_{\mu,\sigma} = \mathcal{N}(\mu,\sigma^2)$ *such that* $\lambda(P,G_{\mu,\sigma}) \leq g(\delta)$.

Now consider our sequence $\{X_j\}$ and let $\{X_j'\}$ be an independent replica of it. Let $S' = \sum_j X_j'$ and let $T = S - S'$. According to Theorem 2 and the bounds between Lévy and Kolmogorov distances recalled above, the sum S will be approximately Gaussian if and only if T is approximately Gaussian. Thus *it will be sufficient to study the case where the independent variables* X_j *have distributions that are symmetric around zero.*

We shall make that assumption in the remainder of the present paper.

For our next statement we shall need a variant of Theorem 2 applicable to the symmetric case as follows.

THEOREM 2'. *There is a function* f *on* $[0,1]$ *to* $[0,1]$ *with the following properties:*

i) $f(\delta) \geq \delta$ *and* $f(\delta)$ *decreases to zero as* δ *decreases to zero.*

ii) *If* $G = \mathcal{N}(0,1)$ *and* P *and* Q *are probability measures symmetric around zero such that* $\lambda(PQ,G) \leq \delta$ *then there is a* σ *such that for* $G_\sigma = \mathcal{N}(0,\sigma^2)$ *one has* $\lambda(P,G_\sigma) \leq f(\delta)$.

We are aware of results of Sapogov [4] that give bounds on the function g. However we shall not use these bounds in order to show that Lévy's argument can be carried out without actual knowledge of the bounds, although the statements would be more precise if the bounds were used.

THEOREM 3. *Let the variables* X_j *be independent symmetrically distributed around zero. Let* $D_j^2 = E(1 \wedge X_j^2)$. *Then there are functions* $\varepsilon(\delta)$, $\theta(\delta)$, $\omega(\delta)$ *all tending to zero as* $\delta \to 0$ *with the following properties. Let* J *be the subset of the integers where* $D_j^2 < \varepsilon(\delta)$. *Then, if* $p[\mathscr{L}(S), G_1] < \delta$ *one has*

i) $\sum_j \{P|X_j| \geq \theta(\delta);\ j \in J\} \leq \omega(\delta)$.

ii) *For each* $j \in J^c$ *there is a Gaussian* G_{σ_j} *such that*

$$\sum_j \{\rho(P_j, G_{\sigma_i}):\ j \in J^c\} \leq (64\pi)[f(\delta)]^{1/g} .$$

This Theorem admits a converse. However to get a converse in terms of the Kolmogorov distance one must assume that $\sum_j D_j^2$ or something similar is not too close to zero. Here is such a converse.

THEOREM 4. *Let the* X_j *be independent and symmetrically distributed around zero. Assume that for some subset* J *of the integers one has*

i) $\sum_j \{P\{|X_j| > \varepsilon\}:\ j \in J\} \leq \varepsilon$

ii) *For each* $j \in J^c$ *there is a* σ_j *such that*

$\sum_j \{\rho(P_j, G_{\sigma_j}):\ j \in J^c\} < \varepsilon$.

Let

$$\tau^2 = \sum_j \{\sigma_j^2:\ j \in J^c\} + \sum_j \{E\{\varepsilon \wedge [X_j]\}^2:\ j \in J\} .$$

Then

$$\rho[\mathscr{L}(S), G_\tau] \leq K \frac{\varepsilon}{\tau} + 2\varepsilon$$

for a certain universal constant K.

The combination of these two Theorems is what Lévy had stated in his own way: In order that $\mathscr{L}(S)$ be close to $\mathscr{N}(0,1)$ it is necessary and sufficient that i) any term that is not negligible be close to Gaussian and ii) the maximum of the negligible terms be itself negligible. Lévy seems to have been thinking of "nonnegligible" as something like our $D_j^2 \geq \varepsilon$ for a *fixed* ε. Hence the number of nonneglibible terms has to stay finite. With the condition used here in Theorem 3 that $D_j^2 \geq \varepsilon(\delta)$, the number of such terms may tend to infinity as $\delta \to 0$. Hence the stronger formulation in terms of $\sum_j \{\rho(P_1; G_{\sigma_j}): j \in J^c\}$.

3. Auxiliary lemmas. As stated we assume that the variables X_j are independent and symmetrically distributed around zero. We shall use a splitting technique described by Lévy in [1].

Let $(\xi_j, \eta_j, U_j, V_j)$ be independent random variables such that $\mathcal{L}(\xi_j) = \mathcal{L}(\eta_j)$ and $P(\xi_j = 1) = 1 - P(\xi_j = 0) = \alpha_j$. Assume that X_j has the same distribution as $Y_j = (1-\xi_j)U_j + \xi_j V_j$. The technique consists in replacing the sum $S = \sum Y_j$ by a sum $T = \sum_j (1-\eta_j)U_j + \sum_j \xi_j V_j$, thus removing the difference $S - T = \sum_j (\eta_j - \xi_j)U_j$.

A splitting of X_j in this form can be obtained in several manners. One possibility is to take numbers θ_j with $P\{|X_j| \geq \theta_j\} \leq \alpha_j$ and $\mathcal{L}(U_j) = \mathcal{L}\{X_j \,|\, |X_j| < \theta_j\}$, $\mathcal{L}(V_j) = \mathcal{L}[X_j \,|\, |X_j| \geq \theta_j]$. A more refined procedure would be to take independent variables W_j, uniformly distributed on $[-1,+1]$ and write that X_j has the same distribution as a certain nondecreasing function $\varphi_j(W_j)$ such that $\varphi(-x) = -\varphi(x)$ for $x \geq 0$. One could then take $\xi_j = I[|W_j| \geq 1-\alpha_j]$ and $\mathcal{L}(U_j) = \mathcal{L}\{X_j \,|\, |W_j| < 1-\alpha_j\}$ and so forth.

LEMMA 1. *Let* $(\xi_j, \eta_j, U_j, V_j)$ *yield a splitting of* X_j *as described. Assume that* $\sup_j \alpha_j \leq \alpha$ *and that the* U_j *have symmetric distribution around zero. Then*

$$\rho[\mathcal{L}(S), \mathcal{L}(T)] \leq 13\alpha^{1/3}$$

PROOF. Take a number $\tau \geq 0$ and let $U_j = U_j' + U_j''$ where $U_j' = U_j$ if $|U_j| \leq \tau$ and $U_j' = 0$ otherwise. Let $\beta_j^2 = E(U_j')^2$. The variance of $\sum_j (\eta_j - \xi_j)U_j'$ is equal to $2\sum \alpha_j(1-\alpha_j)\beta_j^2 \leq 2\alpha(1-\alpha)\sum_j \beta_j^2$. Using Chebyshev's inequality and taking account that

$$\Pr[\sum_j(\eta_j-\xi_j)U_j'' \neq 0] \leq 2\alpha(1-\alpha)\sum_j P[|U_j|>\tau]$$

one can write

$$\Pr\{|\sum_j(\eta_j-\xi_j)U_j|>\tau\} \leq 2\alpha(1-\alpha)D^2(\tau)$$

where $D^2(\tau) = \sum_j E\{1\wedge\left|\frac{U_j}{\tau}\right|^2\}$. Thus

$$\Pr\{|S-T|>\tau\} \leq 2\alpha(1-\alpha)D^2(\tau) .$$

An application of standard inequalities for Kolmogorov distances and moduli of continuity yields

$$\rho[\mathscr{L}(S),\mathscr{L}(T)] \leq [\Gamma_S(\tau)\wedge\Gamma_T(\tau)] + 2\alpha(1-\alpha)D^2(\tau)$$

where for instance $\Gamma_S(\tau) = \sup_x P[x<S\leq x+\tau]$. According to Esseen's modification of Kolmogorov's concentration inequalities, one has

$$\Gamma_T(\tau) \leq \frac{2\sqrt{2\pi}}{(\sqrt{1-\alpha})D(\tau)} .$$

Now consider two cases. It may be that $2\alpha D^2(0) \leq 13\alpha^{1/3}$. Then $\Pr\{\sum(\eta_i-\xi_i)U_j \neq 0\} \leq 2\alpha D^2(0)$ and the desired result follows. If on the contrary $2\alpha D^2(0) > 13\alpha^{1/3}$, note that $D^2(\tau)$ decreases continuously and tends to zero at infinity. Thus there is a smallest value τ such that $(1-\alpha)^3D^3(\tau) = \frac{1}{\alpha}\sqrt{\frac{\pi}{2}}$. This value minimizes the expression

$$2\left\{\frac{\sqrt{2\pi}}{D(\tau)\sqrt{1-\alpha}} + \alpha(1-\alpha)D^2(\tau)\right\}$$

the value of the minimum is

$$4[2^{1/3}+2^{-1/3}]\pi^{1/3}\alpha^{1/3} \leq 13\alpha^{1/3} .$$

This completes the proof of the lemma.

Now let us return to the Normal approximation with $\rho[\mathcal{L}(S), \mathcal{N}(0,1)] \leq \delta$. According to the above we also have

$$\rho[\mathcal{L}(T), \mathcal{N}(0,1)] = \gamma \leq \delta + 13\alpha^{1/3} .$$

Thus we shall work with T instead of S. Note that T is a sum of independent terms $\sum_j (1-\eta_j)U_j + \sum_j \xi_j V_j$. Therefore Lévy's theorem can be applied to the sum $\sum \xi_j V_j$.

LEMMA 2. *Let* $\gamma = \rho[\mathcal{L}(T), \mathcal{N}(0,1)]$. *Assume* $32f(\gamma) \leq 1$ *and assume that in the above splitting there is a subset* J *where the* V_j *are either identically zero or such that*

$$|V_j| \geq \theta = (1.6)\{1 + \sqrt{2|\log f(\gamma)|}\}f(\gamma) .$$

Then $\sum_{j \in J} P\{|\xi_j V_j| \geq \theta\} \leq 4f(\gamma)$.

PROOF. Let H be a subset of J where $|V_j| \geq \theta$ and where $\eta = \sum_{j \in H} \alpha_j \leq \frac{1}{4}$. Then, if $W = \sum\{\xi_j V_j : j \in H\}$ one has $\Pr[W = 0] \geq 1 - \sum \alpha_j \geq 1 - \eta$ and $\Pr\{|W| \geq \Theta\} \geq \sum_j \alpha_j \prod_{k \neq j} (1-\alpha_h) \geq \eta(1-\eta)$.

According to the Cramér-Lévy theorem, there is a Gaussian $\mathcal{N}(0,\sigma^2) = G_\sigma$ such that $\lambda(\mathcal{L}(W), G_\sigma) \leq f(\gamma)$. Since $\Pr[W = 0] \geq 1 - \eta$, such a Gaussian measure must be such that

$$G_\sigma[-f(\gamma), f(\gamma)] \geq 1 - \eta - f(\gamma) .$$

However, here, $\eta + f(\gamma) < \frac{1}{2}$. Therefore $(1.6)f(\gamma) \geq \sigma$. Since $\Pr[|W| \geq \theta] \geq \eta(1-\eta)$ one must also have

$$G_\sigma\{[-\theta + f(\gamma), \theta - f(\gamma)]^c\} \geq \eta(1-\eta) - f(\gamma) .$$

Using the usual upper bound on tail probabilities, this gives

$$f(\gamma) + \sqrt{\frac{2}{\pi}}\, \frac{\sigma}{\theta - f(\gamma)} \exp\{-\frac{1}{2\sigma^2}\, [\theta - f(\gamma)]^2\} \geq \eta(1-\eta) .$$

Here θ is chosen so that

$$\theta - f(\gamma) \geq (1.6) f(\gamma)\sqrt{2|\log f(\gamma)|} \geq \sigma\sqrt{2|\log f(\gamma)|} .$$

Thus $\eta(1-\eta) \leq 2f(\gamma)$ and $\eta \leq 4f(\gamma)$, giving the required bound for the subset H. Now note that $13\alpha^{1/3} \leq \gamma \leq f(\gamma)$. Thus $\alpha \leq (\frac{f(\gamma)}{13})^3$. Thus any other element of J could be added to H without violating the condition $\eta \leq \frac{1}{4}$. It must therefore be true that $\sum_{j\in J} \alpha_j \leq 4f(\gamma)$. This concludes the proof of the lemma.

LEMMA 3. *Let* P *be a probability measure such that* $\lambda(P, G_\sigma) \leq \varepsilon$ *for* $G_\sigma = \mathcal{N}(0, \sigma^2)$. *Then* $\rho(P, G_\sigma) \leq \varepsilon(1 + \frac{1}{2\sigma})$.

PROOF. For G_σ an interval of length ε has a probability at most $\frac{1}{\sigma\sqrt{2\pi}}\, \varepsilon$.

LEMMA 4. *Let* P_1 *and* P_2 *be two probability measures. Let* $D_i^2(\tau) = \int \{1 \wedge (\frac{x}{\tau})^2\} dP_i$. *Then* $|D_1^2(\tau) - D_2^2(\tau)| \leq 2\rho(P_1, P_2)$.

PROOF. Let P_i' be the distribution of $\left(\frac{X_i}{\tau}\right)^2$. Then $\rho(P_1', P_2') \leq 2\rho(P_1, P_2)$. Furthermore, if P_i'' is the distribution of $Y_i = 1 \wedge \left(\frac{X_i}{\tau}\right)^2$, then $\rho(P_1'', P_2'') \leq \rho(P_1', P_2')$. Let F_i be the cumulative distribution of Y_i. Then $EY_k = \int_0^1 [1 - F_i(y)]\, dy$. The result follows.

We shall need another inequality using the Lévy distance instead of the Kolmogorov distance.

LEMMA 5. *Let* $\lambda(P_j, G_{\sigma_j}) \leq \lambda$. *Then* $2\sigma_j^2 \geq D_j^2(1) - 2\lambda(1 + \frac{1}{2}\lambda)$ *with*

$D_j^2(z) = E\,1 \wedge \left(\frac{X_j}{\tau}\right)^2$ *as usual.*

PROOF. Let Z be a random variable with $\mathscr{L}(Z) = G_\sigma$. Then for $y > \lambda \geq 0$ one can write $\Pr[Z > y - \lambda] \geq P[X > y] - \lambda$ and a similar inequality for the negative tail of the distribution. Combining them, one obtains

$$\Pr[|Z| + \lambda > y] \geq P[|X| > y] - 2\lambda \ ,$$

or equivalently, for $\lambda^2 < v \leq 1$

$$\Pr\{[|Z| + \lambda]^2 > v\} \geq P\{1 \wedge |X|^2 > v\} - 2\lambda \ .$$

Integrating on $[0,1]$ gives

$$\int_0^1 \Pr[[|Z| + \lambda]^2 > v]\,dv \geq \int_0^1 \Pr\{1 \wedge |X|^2 > v\}\,dv - 2\lambda - \lambda^2 \ .$$

Integration by parts then yields

$$E[|Z| + \lambda]^2 \geq E\{1 \wedge |X|^2\} - 2\lambda - \lambda^2 \ .$$

However $[|Z| + \lambda]^2 \leq 2[|Z|^2 + \lambda^2]$. Therefore

$$2\,E\,Z^2 \geq E(1 \wedge X^2) - 2\lambda - 3\lambda^2 \ .$$

LEMMA 6. *For* $S = \sum_j X_j$ *assume* $\rho[\mathscr{L}(S), G] \leq \delta$. *Fix an* $\varepsilon > 0$ *and a* $t > 0$. *Let*

$$C(t) = \sqrt{\frac{2}{\pi}} \int_0^{\frac{1}{2}} e^{-x^2/2}\,dx \ .$$

Assume $C(t) > 2\delta$. *Then the set of integers* j *such that* $D_j^2(t) \geq \varepsilon$ *has cardinality at most* $\frac{1}{\varepsilon} \frac{8\pi}{[C(t) - 2\delta]^2}$.

PROOF. The concentration $\sup_x P[x \le S \le x+t]$ is at most

$$\frac{2\sqrt{2\pi}}{\sqrt{\sum_j D_j^2(t)}} .$$

The result follows.

LEMMA 7. *Assume* $\delta < .09$. *Let* ε *be such that* $\varepsilon \ge 4f(\delta)[1+2f(\delta)]$. *Let* $J^c(\varepsilon)$ *be the set of indices* j *such that* $D_j^2(1) \ge \varepsilon$. *Then*

$$\Delta = \sum_j \{\rho(P_j, G_{\sigma_j}) \colon j \in J^c(\varepsilon)\} \le \frac{32\pi}{\varepsilon} f(\delta)[1+\frac{1}{\sqrt{\varepsilon}}]$$

Also, for $0 < t < 1$ *one has*

$$\alpha = \sup_j \{P\{|X_j| \ge t\} \colon j \in J(\varepsilon)\} \le \frac{\varepsilon}{t^2} .$$

PROOF. The last statement follows from Chebyshev's inequality. The first one is obtained as a combination of Lemma 5, Lemma 6 and Lemma 3.

4. Proof of Theorems 3 and 4. To prove Theorem 3, let $\varepsilon(\delta) = [f(\delta)]^{16/27}$. This gives a certain set $J = J(\varepsilon)$. According to Lemma 7 and replacing $1 + \frac{1}{\sqrt{\varepsilon}}$ by $\frac{2}{\sqrt{\varepsilon}}$, the sum $\sum_j \{\rho(P_j, G_{\sigma_j}): j \in J^c\}$ does not exceed $(64)\pi[f(\delta)]^{1/9}$.

Similarly, one can take $t^2 = (13)^3[f(\delta)]^{7/27}$. Then

$$\alpha = \sup_j \{P|X_j| \geq (13)^{3/2}[f(\delta)]^{7/54} : j \in J\} \leq \frac{1}{(13)^3}[f(\delta)]^{1/3} .$$

Thus $13\alpha^{1/3} \leq [f(\delta)]^{1/9}$. Let $\gamma = \delta + [f(\delta)]^{1/9}$ and let $\theta(\delta)$ be the maximum of $(13)^{3/2}[f(\delta)]^{7/54}$ and $(1.6)f(\gamma)\{1+\sqrt{2|\log f(\gamma)|}\}$. Now Lemma 2 says that

$$\sum_j \{P[|X_j| \geq \theta(\delta)] : j \in J\} \leq \omega(\delta) = 4f(\gamma) .$$

This concludes the proof of Theorem 3.

To prove Theorem 4, let $\sigma^2 = \sum_j [\sigma_j^2 : j \in J^c]$ and $\beta^2 = \sum_j \{E[\varepsilon^2 \wedge X_j^2] : j \in J\}$. Let $V = \sum_j \{X_j, j \in J^c\}$, $W = \sum_j \{X_j : j \in J\}$. Then $\rho[\mathscr{L}(V), G_\sigma] \leq \sum_j \{\rho(P_j, G_{\sigma_j}) : j \in J^c\} \leq \varepsilon$. Also, if $Y_j = [\varepsilon \wedge X_j]\operatorname{sign} X_j$ and $Z = \sum_j [Y_j : j \in J]$ one will have $\rho[\mathscr{L}(W), \mathscr{L}(Z)] \leq \varepsilon$. Thus it is enough to bound the distance between G_τ and the convolution of G_σ with $\mathscr{L}(Z)$. For that one can use the procedure commonly employed to obtain the Berry-Esseen bounds. This will yield the result as stated, since the distance between G_τ and G_σ convoluted with $\mathscr{L}(Z)$ will take the form $K\varepsilon \sum_j \{E(\varepsilon \wedge X_j)^2 : j \in J\} \frac{1}{\tau^3}$.

Note that the bound will be usable only if τ is large compared to ε. Note also that it would more pleasant aesthetically to use $(\tau')^2 = \sigma^2 + \sum_j \{D_j^2 : j \in J\}$ instead of the τ^2 of the theorem. This can be done if for instance $(\tau')^2 \geq 2\sqrt{\varepsilon}$. Alternately, one could bound the Lévy distance $\lambda[\mathscr{L}(S), G_{\tau'}]$, instead of the Kolmogorov distance.

REFERENCES

[1] LÉVY, P. (1937). *Theorie de l'Addition des Variables Aléatoires.* Gauthier Villans, Paris.

[2] LÉVY, P. (1935). Propriétés asymptotiques des sommes de variables indépendantes ou enchainées. *J. de Math. Pures et Appliquées* **14** #4 347-402.

[3] MAČYS, J. (1968). Sur la convergence des répartitions de sommes de variables aléatoires indépendantes vers les lois de la classe I_0 de Linnik. *Comptes Rendus. Acad. Sciences Paris* **267** 316-317.

[4] SAPOGOV, N. A. (1951). The stability problem for a theorem of Cramér. *Izv. Akad. Nauk SSSR, Ser. Mat.* **15** 205-218.

[5] ZOLOTAREV, V. M. (1967). A generalization of the Lindeberg-Feller theorem. *Theorem of Probability and its Applications* **12** #4 608-618.

[6] ZOLOTAREV, V. M. (1970). Théoremès limites généraux pour les sommes de variables indépendantes. *Comptes Rendus. Acad. Sciences Paris* **270** 899-902.

TECHNICAL REPORTS

Statistics Department

University of California, Berkeley

1. BREIMAN, L. and FREEDMAN, D. (Nov. 1981, revised Feb. 1982). How many variables should be entered in a regression equation? Jour. Amer. Statist. Assoc., March 1983, 78, No. 381, 131-136.

2. BRILLINGER, D. R. (Jan. 1982). Some contrasting examples of the time and frequency domain approaches to time series analysis. Time Series Methods in Hydrosciences, (A. H. El-Shaarawi and S. R. Esterby, eds.) Elsevier Scientific Publishing Co., Amsterdam, 1982, pp. 1-15.

3. DOKSUM, K. A. (Jan. 1982). On the performance of estimates in proportional hazard and log-linear models. Survival Analysis, (John Crowley and Richard A. Johnson, eds.) IMS Lecture Notes - Monograph Series, (Shanti S. Gupta, series ed.) 1982, 74-84.

4. BICKEL, P. J. and BREIMAN, L. (Feb. 1982). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. Ann. Prob., Feb. 1982, 11. No. 1, 185-214.

5. BRILLINGER, D. R. and TUKEY, J. W. (March 1982). Spectrum estimation and system identification relying on a Fourier transform. The Collected Works of J. W. Tukey, vol. 2, Wadsworth, 1985, 1001-1141.

6. BERAN, R. (May 1982). Jackknife approximation to bootstrap estimates. Ann. Statist., March 1984, 12 No. 1, 101-118.

7. BICKEL, P. J. and FREEDMAN, D. A. (June 1982). Bootstrapping regression models with many parameters. Lehmann Festschrift, (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.) Wadsworth Press, Belmont, 1983, 28-48.

8. BICKEL, P. J. and COLLINS, J. (March 1982). Minimizing Fisher information over mixtures of distributions. Sankhyā, 1983, 45, Series A, Pt. 1, 1-19.

9. BREIMAN, L. and FRIEDMAN, J. (July 1982). Estimating optimal transformations for multiple regression and correlation.

10. FREEDMAN, D. A. and PETERS, S. (July 1982, revised Aug. 1983). Bootstrapping a regression equation: some empirical results. JASA, 1984, 79, 97-106.

11. EATON, M. L. and FREEDMAN, D. A. (Sept. 1982). A remark on adjusting for covariates in multiple regression.

12. BICKEL, P. J. (April 1982). Minimax estimation of the mean of a mean of a normal distribution subject to doing well at a point. Recent Advances in Statistics, Academic Press, 1983.

14. FREEDMAN, D. A., ROTHENBERG, T. and SUTCH, R. (Oct. 1982). A review of a residential energy end use model.

15. BRILLINGER, D. and PREISLER, H. (Nov. 1982). Maximum likelihood estimation in a latent variable problem. Studies in Econometrics, Time Series, and Multivariate Statistics, (eds. S. Karlin, T. Amemiya, L. A. Goodman). Academic Press, New York, 1983, pp. 31-65.

16. BICKEL, P. J. (Nov. 1982). Robust regression based on infinitesimal neighborhoods. Ann. Statist., Dec. 1984, 12, 1349-1368.

17. DRAPER, D. C. (Feb. 1983). Rank-based robust analysis of linear models. I. Exposition and review.

18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.

19. FREEDMAN, D. A. and FIENBERG, S. (Feb. 1983, revised April 1983). Statistics and the scientific method, Comments on and reactions to Freedman, A rejoinder to Fienberg's comments. Springer New York 1985 Cohort Analysis in Social Research, (W. M. Mason and S. E. Fienberg, eds.).

20. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Jan. 1984). Using the bootstrap to evaluate forecasting equations. J. of Forecasting. 1985, Vol. 4, 251-262.

21. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Aug. 1983). Bootstrapping an econometric model: some empirical results. JBES, 1985, 2, 150-158.

22. FREEDMAN, D. A. (March 1983). Structural-equation models: a case study.

23. DAGGETT, R. S. and FREEDMAN, D. (April 1983, revised Sept. 1983). Econometrics and the law: a case study in the proof of antitrust damages. Proc. of the Berkeley Conference, in honor of Jerzy Neyman and Jack Kiefer. Vol I pp. 123-172. (L. Le Cam, R. Olshen eds.) Wadsworth, 1985.

24. DOKSUM, K. and YANDELL, B. (April 1983). Tests for exponentiality. Handbook of Statistics, (P. R. Krishnaiah and P. K. Sen, eds.) 4, 1984.

25. FREEDMAN, D. A. (May 1983). Comments on a paper by Markus.

26. FREEDMAN, D. (Oct. 1983, revised March 1984). On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist., 1984, 12, 827-842.

27. DOKSUM, K. A. (Dec. 1983). An extension of partial likelihood methods for proportional hazard models to general transformation models. Ann. Statist., 1987, 15, 325-345.

28. BICKEL, P. J., GOETZE, F. and VAN ZWET, W. R. (Jan. 1984). A simple analysis of third order efficiency of estimate Proc. of the Neyman-Kiefer Conference, (L. Le Cam, ed.) Wadsworth, 1985.

29. BICKEL, P. J. and FREEDMAN, D. A. Asymptotic normality and the bootstrap in stratified sampling. Ann. Statist. 12 470-482.

30. FREEDMAN, D. A. (Jan. 1984). The mean vs. the median: a case study in 4-R Act litigation. JBES. 1985 Vol 3 pp. 1-13.

31. STONE, C. J. (Feb. 1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., Dec. 1984, 12, 1285-1297.

32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.

33. STONE, C. J. (Oct. 1984). Additive regression and other nonparametric models. Ann. Statist., 1985, 13, 689-705.

34. STONE, C. J. (June 1984). An asymptotically optimal histogram selection rule. Proc. of the Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer (L. Le Cam and R. A. Olshen, eds.), II, 513-520.

35. FREEDMAN, D. A. and NAVIDI, W. C. (Sept. 1984, revised Jan. 1985). Regression models for adjusting the 1980 Census. Statistical Science. Feb 1986, Vol. 1, No. 1, 3-39.

36. FREEDMAN, D. A. (Sept. 1984, revised Nov. 1984). De Finetti's theorem in continuous time.

37. DIACONIS, P. and FREEDMAN, D. (Oct. 1984). An elementary proof of Stirling's formula. Amer. Math Monthly. Feb 1986, Vol. 93, No. 2, 123-125.

38. LE CAM, L. (Nov. 1984). Sur l'approximation de familles de mesures par des familles Gaussiennes. Ann. Inst. Henri Poincaré, 1985, 21, 225-287.

39. DIACONIS, P. and FREEDMAN, D. A. (Nov. 1984). A note on weak star uniformities.

40. BREIMAN, L. and IHAKA, R. (Dec. 1984). Nonlinear discriminant analysis via SCALING and ACE.

41. STONE, C. J. (Jan. 1985). The dimensionality reduction principle for generalized additive models.

42. LE CAM, L. (Jan. 1985). On the normal approximation for sums of independent variables.

43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?

44. BRILLINGER, D. R. (1985). The natural variability of vital rates and associated statistics. Biometrics, to appear.

45. BRILLINGER, D. R. (1985). Fourier inference: some methods for the analysis of array and nonGaussian series data. Water Resources Bulletin, 1985, 21, 743-756.

46. BREIMAN, L. and STONE, C. J. (1985). Broad spectrum estimates and confidence intervals for tail quantiles.

47. DABROWSKA, D. M. and DOKSUM, K. A. (1985, revised March 1987). Partial likelihood in transformation models with censored data.

48. HAYCOCK, K. A. and BRILLINGER, D. R. (November 1985). LIBDRB: A subroutine library for elementary time series analysis.

49. BRILLINGER, D. R. (October 1985). Fitting cosines: some procedures and some physical examples. Joshi Festschrift, 1986. D. Reidel.

50. BRILLINGER, D. R. (November 1985). What do seismology and neurophysiology have in common? - Statistics! Comptes Rendus Math. Rep. Acad. Sci. Canada. January, 1986.

51. COX, D. D. and O'SULLIVAN, F. (October 1985). Analysis of penalized likelihood-type estimators with application to generalized smoothing in Sobolev Spaces.

52. O'SULLIVAN, F. (November 1985). A practical perspective on ill-posed inverse problems: A review with some new developments. To appear in Journal of Statistical Science.

53. LE CAM, L. and YANG, G. L. (November 1985, revised March 1987). On the preservation of local asymptotic normality under information loss.

54. BLACKWELL, D. (November 1985). Approximate normality of large products.

55. FREEDMAN, D. A. (June 1987). As others see us: A case study in path analysis. Journal of Educational Statistics. 12, 101-128.

56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.

57. LE CAM, L. (February 1986). On the Bernstein - von Mises theorem.

58. O'SULLIVAN, F. (January 1986). Estimation of Densities and Hazards by the Method of Penalized likelihood.

59. ALDOUS, D. and DIACONIS, P. (February 1986). Strong Uniform Times and Finite Random Walks.

60. ALDOUS, D. (March 1986). On the Markov Chain simulation Method for Uniform Combinatorial Distributions and Simulated Annealing.

61. CHENG, C-S. (April 1986). An Optimization Problem with Applications to Optimal Design Theory.

62. CHENG, C-S., MAJUMDAR, D., STUFKEN, J. & TURE, T. E. (May 1986, revised Jan 1987). Optimal step type design for comparing test treatments with a control.

63. CHENG, C-S. (May 1986, revised Jan. 1987). An Application of the Kiefer-Wolfowitz Equivalence Theorem.

64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.

65. ALDOUS, D. (JUNE 1986). Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes.

66. PITMAN, J. (JULY 1986, revised November 1986). Stationary Excursions.

67. DABROWSKA, D. and DOKSUM, K. (July 1986, revised November 1986). Estimates and confidence intervals for median and mean life in the proportional hazard model with censored data.

68. LE CAM, L. and YANG, G.L. (July 1986). Distinguished Statistics, Loss of information and a theorem of Robert B. Davies (Fourth edition).

69. STONE, C.J. (July 1986). Asymptotic properties of logspline density estimation.

71. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.

72. LEHMANN, E.L. (July 1986). Statistics - an overview.

73. STONE, C.J. (August 1986). A nonparametric framework for statistical modelling.

74. BIANE, PH. and YOR, M. (August 1986). A relation between Lévy's stochastic area formula, Legendre polynomial, and some continued fractions of Gauss.

75. LEHMANN, E.L. (August 1986, revised July 1987). Comparing Location Experiments.

76. O'SULLIVAN, F. (September 1986). Relative risk estimation.

77. O'SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.

78. PITMAN, J. & YOR, M. (September 1987). Further asymptotic laws of planar Brownian motion.

79. FREEDMAN, D.A. & ZEISEL, H. (November 1986). From mouse to man: The quantitative assessment of cancer risks. To appear in Statistical Science.

80. BRILLINGER, D.R. (October 1986). Maximum likelihood analysis of spike trains of interacting nerve cells.

81. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.

82. DOKSUM, K.J. and LO, A.Y. (November 1986). Consistent and robust Bayes Procedures for Location based on Partial Information.

83. DABROWSKA, D.M., DOKSUM, K.A. and MIURA, R. (November 1986). Rank estimates in a class of semiparametric two-sample models.

84. BRILLINGER, D. (December 1986). Some statistical methods for random process data from seismology and neurophysiology.

85. DIACONIS, P. and FREEDMAN, D. (December 1986). A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 1987, **23**, 397-423.

86. DABROWSKA, D.M. (January 1987). Uniform consistency of nearest neighbour and kernel conditional Kaplan - Meier estimates.

87. FREEDMAN, D.A., NAVIDI, W. and PETERS, S.C. (February 1987). On the impact of variable selection in fitting regression equations.

88. ALDOUS, D. (February 1987, revised April 1987). Hashing with linear probing, under non-uniform probabilities.

89. DABROWSKA, D.M. and DOKSUM, K.A. (March 1987, revised January 1988). Estimating and testing in a two sample generalized odds rate model.

90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.

91. DIACONIS, P and FREEDMAN, D.A. (April 1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. To appear in the Journal of Applied Probability.

92. DABROWSKA, D.M. (April 1987, revised September 1987). Kaplan-Meier estimate on the plane.

92a. ALDOUS, D. (April 1987). The Harmonic mean formula for probabilities of Unions: Applications to sparse random graphs.

93. DABROWSKA, D.M. (June 1987, revised Feb 1988). Nonparametric quantile regression with censored data.

94. DONOHO, D.L. & STARK, P.B. (June 1987). Uncertainty principles and signal recovery.

95. CANCELLED

96. BRILLINGER, D.R. (June 1987). Some examples of the statistical analysis of seismological data. To appear in *Proceedings, Centennial Anniversary Symposium, Seismographic Stations, University of California, Berkeley.*

97. FREEDMAN, D.A. and NAVIDI, W. (June 1987). On the multi-stage model for carcinogenesis. To appear in Environmental Health Perspectives.

98. O'SULLIVAN, F. and WONG, T. (June 1987). Determining a function diffusion coefficient in the heat equation.

99. O'SULLIVAN, F. (June 1987). Constrained non-linear regularization with application to some system identification problems.

100. LE CAM, L. (July 1987, revised Nov 1987). On the standard asymptotic confidence ellipsoids of Wald.

101. DONOHO, D.L. and LIU, R.C. (July 1987). Pathologies of some minimum distance estimators. Annals of Statistics, June, 1988.

102. BRILLINGER, D.R., DOWNING, K.H. and GLAESER, R.M. (July 1987). Some statistical aspects of low-dose electron imaging of crystals.

103. LE CAM, L. (August 1987). Harald Cramér and sums of independent random variables.

104. DONOHO, A.W., DONOHO, D.L. and GASKO, M. (August 1987). Macspin: Dynamic graphics on a desktop computer. IEEE Computer Graphics and applications, June, 1988.

105. DONOHO, D.L. and LIU, R.C. (August 1987). On minimax estimation of linear functionals.

106. DABROWSKA, D.M. (August 1987). Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap.

107. CHENG, C-S. (August 1987). Some orthogonal main-effect plans for asymmetrical factorials.

108. CHENG, C-S. and JACROUX, M. (August 1987). On the construction of trend-free run orders of two-level factorial designs.

109. KLASS, M.J. (August 1987). Maximizing $E \max_{1 \le k \le n} S_k^+ / ES_n^+$: A prophet inequality for sums of I.I.D. mean zero variates.

110. DONOHO, D.L. and LIU, R.C. (August 1987). The "automatic" robustness of minimum distance functionals. Annals of Statistics, June, 1988.

111. BICKEL, P.J. and GHOSH, J.K. (August 1987, revised June 1988). A decomposition for the likelihood ratio statistic and the Bartlett correction — a Bayesian argument.

112. BURDZY, K., PITMAN, J.W. and YOR, M. (September 1987). Some asymptotic laws for crossings and excursions.

113. ADHIKARI, A. and PITMAN, J. (September 1987). The shortest planar arc of width 1.

114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.

115. BICKEL, P.J. and RITOV, Y. (Sept. 1987, revised Aug 1988). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (Sept.1987, revised Aug. 1988). Achieving information bounds in non and semiparametric models.

117. RITOV, Y. (October 1987). On the convergence of a maximal correlation algorithm with alternating projections.

118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.

119. HESSE, C.H. (October 1987). An asymptotic expansion for the mean of the passage-time distribution of integrated Brownian Motion.

120. DONOHO, D. and LIU, R. (October 1987, revised March 1988). Geometrizing rates of convergence, II.

121. BRILLINGER, D.R. (October 1987). Estimating the chances of large earthquakes by radiocarbon dating and statistical modelling. To appear in *Statistics a Guide to the Unknown.*

122. ALDOUS, D., FLANNERY, B. and PALACIOS, J.L. (November 1987). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains.

123. DONOHO, D.L., MACGIBBON, B. and LIU, R.C. (Nov.1987, revised July 1988). Minimax risk for hyperrectangles.

124. ALDOUS, D. (November 1987). Stopping times and tightness II.

125. HESSE, C.H. (November 1987). The present state of a stochastic model for sedimentation.

126. DALANG, R.C. (December 1987, revised June 1988). Optimal stopping of two-parameter processes on nonstandard probability spaces.

127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.

129. SMITH, D.L. (December 1987). Exponential bounds in Vapnik-Červonenkis classes of index 1.

130. STONE, C.J. (Nov.1987, revised Sept. 1988). Uniform error bounds involving logspline models.

131. Same as No. 140

132. HESSE, C.H. (December 1987). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes

133. DONOHO, D.L. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean, I.

134. DUBINS, L.E. and SCHWARZ, G. (December 1987). A sharp inequality for martingales and stopping-times.

135. FREEDMAN, D.A. and NAVIDI, W. (December 1987). On the risk of lung cancer for ex-smokers.

136. LE CAM, L. (January 1988). On some stochastic models of the effects of radiation on cell survival.

137. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the uniform consistency of Bayes estimates for multinomial probabilities.

137a. DONOHO, D.L. and LIU, R.C. (1987). Geometrizing rates of convergence, I.

138. DONOHO, D.L. and LIU, R.C. (January 1988). Geometrizing rates of convergence, III.

139. BERAN, R. (January 1988). Refining simultaneous confidence sets.

140. HESSE, C.H. (December 1987). Numerical and statistical aspects of neural networks.

141. BRILLINGER, D.R. (January 1988). Two reports on trend analysis: a) An Elementary Trend Analysis of Rio Negro Levels at Manaus, 1903-1985 b) Consistent Detection of a Monotonic Trend Superposed on a Stationary Time Series

142. DONOHO, D.L. (Jan. 1985, revised Jan. 1988). One-sided inference about functionals of a density.

143. DALANG, R.C. (February 1988). Randomization in the two-armed bandit problem.

144. DABROWSKA, D.M., DOKSUM, K.A. and SONG, J.K. (February 1988). Graphical comparisons of cumulative hazards for two populations.

145. ALDOUS, D.J. (February 1988). Lower bounds for covering times for reversible Markov Chains and random walks on graphs.

146. BICKEL, P.J. and RITOV, Y. (Feb.1988, revised August 1988). Estimating integrated squared density derivatives.

147. STARK, P.B. (March 1988). Strict bounds and applications.

148. DONOHO, D.L. and STARK, P.B. (March 1988). Rearrangements and smoothing.

149. NOLAN, D. (March 1988). Asymptotics for a multivariate location estimator.

150. SEILLIER, F. (March 1988). Sequential probability forecasts and the probability integral transform.

151. NOLAN, D. (March 1988). Limit theorems for a random convex set.

152. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On a theorem of Kuchler and Lauritzen.

153. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the problem of types.

154. DOKSUM, K.A. (May 1988). On the correspondence between models in binary regression analysis and survival analysis.

155. LEHMANN, E.L. (May 1988). Jerzy Neyman, 1894-1981.

156. ALDOUS, D.J. (May 1988). Stein's method in a two-dimensional coverage problem.

157. FAN, J. (June 1988). On the optimal rates of convergence for nonparametric deconvolution problem.

158. DABROWSKA, D. (June 1988). Signed-rank tests for censored matched pairs.

159. BERAN, R.J. and MILLAR, P.W. (June 1988). Multivariate symmetry models.

160. BERAN, R.J. and MILLAR, P.W. (June 1988). Tests of fit for logistic models.

161. BREIMAN, L. and PETERS, S. (June 1988). Comparing automatic bivariate smoothers (A public service enterprise).

162. FAN, J. (June 1988). Optimal global rates of convergence for nonparametric deconvolution problem.

163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform.

164. BICKEL, P.J. and KRIEGER, A.M. (July 1988). Confidence bands for a distribution function using the bootstrap.

165. HESSE, C.H. (July 1988). New methods in the analysis of economic time series I.

166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.

167. BREIMAN, L., STONE, C.J. and KOOPERBERG, C. (August 1988). Confidence bounds for extreme quantiles.

168. LE CAM, L. (August 1988). Maximum likelihood an introduction.

169. BREIMAN, L. (August 1988). Submodel selection and evaluation in regression-The conditional case and little bootstrap.

170. LE CAM, L. (September 1988). On the Prokhorov distance between the empirical process and the associated Gaussian bridge.

171. STONE, C.J. (September 1988). Large-sample inference for logspline models.

172. ADLER, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.

173. MILLAR, P.W. (October 1988). Optimal estimation in the non-parametric multiplicative intensity model.

174. YOR, M. (October 1988). Interwinings of Bessel processes.

175. ROJO, J. (October 1988). On the concept of tail-heaviness.

176. ABRAHAMS, D.M. and RIZZARDI, F. (September 1988). BLSS - The Berkeley interactive statistical system: An overview.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:

Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.