Math-Net.Ru

Общероссийский математический портал

О. Johnson, Convergence of the Poincaré constant,
Теория вероятн. и ее примен., 2003, том 48, вы-
пуск 3, 615–620

DOI: <https://doi.org/10.4213/tvp276>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.224.135.184

26 апреля 2023 г., 17:13:58

16. Logan B., Mallows C., Rice S., Shepp L. Limit distributions of self-normalized sums. — Ann. Probab., 1973, v. 1, p. 788–809.
17. Maller A. A theorem on products of random variables, with application to regression. — Austral. J. Statist., 1981, v. 23, p. 177–185.
18. Пеплов В. В. Суммы независимых случайных величин. М.: Наука, 1972, 416 с.
19. Pinelis I. Extremal probabilistic problems and Hotelling's T^2 test under a symmetric condition. — Ann. Statist., 1994, v. 22, № 1, p. 357–368.
20. Shao Q.-M. Self-normalized large deviations. — Ann. Probab., 1997, v. 25, № 1, p. 285–328.
21. Shao Q.-M. Recent development in self-normalized limit theorems. — Asymptotic Methods in Probability and Statistics. Ed. by B. Szyszkowicz et al. Amsterdam: Elsevier, 1998, p. 467–480.
22. Shao Q.-M. A Cramér type large deviation result for Student's t -statistic. — J. Theoret. Probab., 1999, v. 12, № 2, p. 385–398.
23. Шараметов Ш. Неравенство Берри-Эссеена для статистики Стьюдента. — Узб. матем. журн., 1995, № 2, с. 101–112.
24. Slavova V. V. On the Berry–Esseen bound for Student's statistic. — Lecture Notes in Math., 1985, v. 1155, p. 355–390.
25. Wang Q., Jing B. Y. An exponential nonuniform Berry–Esseen bound for self-normalized sums. — Ann. Probab., 1999, v. 27, № 4, p. 2068–2088.

Поступила в редакцию
9.VI.2003

© 2003 г.

JOHNSON O.*

CONVERGENCE OF THE POINCARÉ CONSTANT

Постоянная Пуанкаре R_Y для случайной величины Y связывает $L^2(Y)$ -норму функции g с ее производной g' . Поскольку разность $R_Y - D_Y$ неотрицательна и равна нулю, тогда и только тогда, когда Y — нормальная случайная величина, то эту разность можно рассматривать как расстояние до нормального распределения. В статье устанавливается наилучшая возможная скорость сходимости этого расстояния в центральной предельной теореме. Кроме того, мы показываем, что R_Y конечна для дискретных смесей нормально распределенных величин, что позволяет получить скорость сходимости в центральной предельной теореме в смысле относительной энтропии.

Ключевые слова и фразы: постоянная Пуанкаре, спектральная щель, центральная предельная теорема, информация по Фишеру.

1. Introduction and results. Poincaré (or spectral gap) inequalities provide a relation between L^2 -norms on functions and their derivatives.

Definition 1.1 ([3]). Given a random variable Y , define the Poincaré constant R_Y :

$$R_Y = \sup_{g \in H_1(Y)} \frac{Dg(Y)}{Eg'(Y)^2},$$

where $H_1(Y)$ is the space of absolutely continuous functions on the real line such that $Dg(Y) > 0$ and $Eg'(Y)^2 < \infty$.

R_Y will not in general be finite, however it will be finite for the normal and other strongly unimodal distributions (see, for example, [10], [8], [6], [5], [11], [3]).

We will exploit various relationships between the Poincaré constant and Fisher information.

* Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WB, United Kingdom; e-mail: otj1000@cam.ac.uk

Definition 1.2. For a random variable Y with smooth density p , define the score function $\rho_Y(y) = p'(y)/p(y)$ and Fisher information $I(Y) = \mathbf{E}\rho(Y)^2$.

Notice that for a given Y , if g is a local maximum of $\mathbf{D}g(Y)/\mathbf{E}g'(Y)^2$, then for all functions h and small t

$$\frac{\mathbf{D}(g+th)}{\mathbf{E}(g'+th')^2} \leq \frac{\mathbf{D}g}{\mathbf{E}g'^2} = R_g,$$

so multiplying out, $0 \leq t^2(R_g \mathbf{E}h'^2 - \mathbf{E}h^2) + 2t(R_g \mathbf{E}g'h' - \mathbf{E}gh)$, which can only hold in an interval around zero if

$$R_g \mathbf{E}g'h' = \mathbf{E}gh. \quad (1)$$

Integration by parts implies therefore that $g = -R_g(\rho_Y g' + g'')$, so local maxima correspond to eigenfunctions of the Laplacian $\mathbf{D}_Y g = \rho_Y g' + g''$, and the global maximum to the least strictly negative eigenvalue (hence the alternative name of spectral gap inequality).

Example 1.1. The Poincaré constant can be infinite. For example, consider the discrete random variable, where $\mathbf{P}\{X = 1\} = \mathbf{P}\{X = -1\} = \frac{1}{2}$. Then, we can choose g such that $g'(-1) = g'(1) = \varepsilon$, but $g(-1) = -1$, $g(1) = 1$, so that $\mathbf{D}g(X) = 1$, but $\mathbf{E}g'(X)^2 = \varepsilon^2$. This argument will work for any discrete random variable, indeed any random variable whose support is not an interval.

However, our first main result shows that discrete random variables perturbed by small normals have a finite Poincaré constant:

Theorem 1.1. Consider X , a random variable with variance σ^2 taking a finite number of values with probabilities $p_1, p_2, \dots, p_n$ respectively, and Z_τ an independent normal with mean zero and variance τ . Then $Y_\tau = X + Z_\tau$ satisfies a Poincaré inequality with constant

$$R_{Y_\tau} = \tau \left(1 + \frac{\sigma^2}{\tau \min_s p_s} \exp \left(\frac{\sigma^2}{\tau \min_s p_s} \right) \right).$$

Proof. See Section 2.

Note that part 8 of Theorem 1.1 in [12] also shows that R_{Y_τ} is finite. However, our bound has an explicit dependence on σ^2 , and so has independent interest.

In [9], we show that finiteness of the Poincaré constant gives an explicit rate of convergence of relative entropy distance in the Central Limit Theorem. This is a strong result, and implies convergence in L^1 .

Theorems 2, 3, and 4 in [3] provide the following results.

Lemma 1.1. For the constant R_X defined above

- 1) $R_{aX+b} = a^2 R_X$;
- 2) if X, Y are independent, then $R_{X+Y} \leq R_X + R_Y$;
- 3) $R_X \geq \mathbf{D}X$, with equality if and only if X is normal;
- 4) if R_X is finite, then $\mathbf{E} \exp(|X - \mathbf{E}X|/(12\sqrt{R_X})) \leq 2$, so X has moments of all orders;
- 5) if $R_{X_n}/\mathbf{D}X_n \rightarrow 1$, then $\mathbf{E}w(X_n) \rightarrow \mathbf{E}w(Z)$, where Z is normal, for any continuous w with $|w(t)| < \exp(ct)$, for sufficiently small c .

The first three properties are reminiscent of those of Fisher Information — a subadditive relation holds and the minimising case characterises the normal distribution. In analogy with the approach to the Central Limit Theorem developed by Brown [4] and Barron [1], we add an extra term into the subadditive relation $R_{X+Y} \leq R_X + R_Y$, which is sandwiched as convergence occurs. This gives us an answer to the question posed by Chen and Lou [7], of identifying the limit of the Poincaré constant in the Central Limit Theorem. This was also answered by Utev [12], though without the explicit rate of convergence that we provide.

Theorem 1.2. Consider $X_1, X_2, \dots$ i.i.d., with $R = R_{X_i}$ and $I = I(X)$ finite. Defining $U_n = (X_1 + \dots + X_n)/\sqrt{n\sigma^2}$, then there exists a constant C , depending only on I and R , such that $R_{U_n} - 1 \leq C/n$.

Proof. See Section 3.

We can argue that this is the best possible rate, up to the choice of the constant. Without loss of generality assume that X has variance 1. Considering $g(x) = x^2 - 1$, we

know that $R_X \geq (\mathbf{E} X^4 - 1)/4$. Since $\mathbf{E} U_{2^n}^4 - 1 = (\mathbf{E} U_n^4 - 1)/2 > 0$, if $R_{U_n} - 1 = f(n)/n$ then $f(2^k)/2^k = R_{U_{2^k}} - 1 \geq (\mathbf{E} X^4 - 1)/2^k$, so $f(2^k) \geq \mathbf{E} X^4 - 1$ and hence does not tend to zero.

Since we can perturb random variables by adding small normals to ensure that the Fisher information is finite, we use this to prove a strong form of the Central Limit Theorem.

Theorem 1.3. *If $X_1, X_2, \dots$ are i.i.d. random variables with mean 0, variance σ^2 , and finite $R_X = R$, then $U_n = (X_1 + \dots + X_n)/\sqrt{n\sigma^2}$ has the property that: $\mathbf{E} w(U_n) \rightarrow \mathbf{E} w(Z)$, where Z is standard normal, for any continuous w such that $w(t) \leq \exp(c|t|)$, where $c < c_0 = 1/(12\sqrt{R})$.*

P r o o f. Given a random variable U , define $U^\tau \sim U + Z_\tau$. Now, since $R_{U_n^\tau} \leq R_{U_n} + \tau \leq R + \tau$, by Lemma 1.1, part 4 taking τ_0 such that $c = 1/(12\sqrt{R + \tau_0})$: $\mathbf{E} \exp(c|U_n^\tau|) \leq 2$, for all n , if $\tau \leq \tau_0$, $c < c_0$.

Now, since U_n^τ have uniformly bounded Fisher information $I(U_n^\tau) \leq I(Z_\tau) = 1/\tau$ and $R_{U_n^\tau} \leq R + \tau$, Theorem 1.2 implies that $\mathbf{E} w(U_n^\tau) \rightarrow \mathbf{E} w(Z^\tau)$.

Hence, since $|\mathbf{E} w(U_n) - \mathbf{E} w(Z)| \leq |\mathbf{E} w(U_n^\tau) - \mathbf{E} w(U_n)| + |\mathbf{E} w(U_n^\tau) - \mathbf{E} w(Z^\tau)| + |\mathbf{E} w(Z) - \mathbf{E} w(Z^\tau)|$, we need only show that given ε , $|\mathbf{E} w(U^\tau) - \mathbf{E} w(U)| \leq \varepsilon$ for τ small enough. This follows by uniform integrability arguments (see Theorem 25.12 in [2]), since $\mathbf{E} |w(U^\tau)|^p \leq \mathbf{E} \exp(cp|U^\tau|) \leq 2$, for some small p , and since $w(U^\tau)$ converges weakly to $w(U)$.

2. Finiteness of R for mixtures of normals.

P r o o f o f T h e o r e m 1.1. Without loss of generality, consider X taking a finite number of values $a_1 > a_2 > \dots > a_n$ with probabilities $p_1, p_2, \dots, p_n$, respectively, where $\mathbf{E} X = \sum_i p_i a_i = 0$.

We introduce the «squared span» $M = \max(|a_1^2 - a_2^2|, |a_2^2 - a_3^2|, \dots, |a_{n-1}^2 - a_n^2|, (a_1 - a_2)^2, (a_2 - a_3)^2, \dots, (a_{n-1} - a_n)^2)$, and write p for $\min_s p_s$.

By Theorem 1 of [3], we need to check that for some R and all x , the density f_τ of Y satisfies

$$\int_x^\infty y f_\tau(y) dy \leq R f_\tau(x). \quad (2)$$

Since $f_\tau(y) = \sum_i p_i \phi_\tau(y - a_i)$, the left-hand side of equation (2) becomes (defining $u_j = \sum_{i=1}^j p_i a_i$ and $a_{n+1} = -\infty$)

$$\begin{aligned} \int_x^\infty y f_\tau(y) dy &= \sum_{i=1}^n p_i \int_x^\infty (y - a_i) \phi_\tau(y - a_i) dy + \sum_{i=1}^n p_i a_i \int_x^\infty \phi_\tau(y - a_i) dy \\ &= \sum_{i=1}^n p_i \tau \phi_\tau(x - a_i) + \sum_{i=1}^n p_i a_i \int_{x-a_i}^\infty \phi_\tau(y) dy \\ &= \tau f_\tau(x) + \sum_{i=1}^n p_i a_i \sum_{j=i}^n \int_{x-a_j}^{x-a_{j+1}} \phi_\tau(y) dy = \tau f_\tau(x) + \sum_{j=1}^{n-1} u_j \left(\int_{x-a_j}^{x-a_{j+1}} \phi_\tau(y) dy \right), \end{aligned}$$

since $u_n = 0$, so for each interval $I_j = (x - a_j, x - a_{j+1})$ we need to consider bounds on $\min_{y \in I_j} y^2$.

We write r for the index such that $a_r \leq x < a_{r-1}$.

First, we consider $x \geq 0$, where we can distinguish 3 cases: for $j < r$; $x - a_{j+1} < 0$, so for $y \in I_j$: $y^2 \geq (x - a_{j+1})^2 = (x - a_j)^2 + 2(a_j - a_{j+1})x + (a_{j+1}^2 - a_j^2) \geq (x - a_j)^2 - M$. For $j = r$; $y \in I_j$ means that $y^2 \geq 0 \geq (x - a_j)^2 - (a_{j-1} - a_j)^2 \geq (x - a_j)^2 - M$.

For $j > r$; $x - a_j > 0$, so for $y \in I_j$: $y^2 \geq (x - a_j)^2$.

Hence for all j , $\min_{y \in I_j} y^2 \geq (x - a_j)^2 - M$, so

$$\int_{x-a_j}^{x-a_{j+1}} \phi_\tau(y) dy \leq (a_j - a_{j+1}) \max_{y \in I_j} \phi_\tau(y) \leq (a_j - a_{j+1}) \phi_\tau(x - a_j) \exp\left(\frac{M}{2\tau}\right).$$

In Lemma 2.1, we prove two technical results, that $u_j(a_j - a_{j+1}) \leq \sigma^2$, and that $Mp \leq 2\sigma^2$. This allows us to deduce that for $x \geq 0$

$$\begin{aligned} \sum_{j=1}^{n-1} u_j \left(\int_{x-a_j}^{x-a_{j+1}} \phi_\tau(y) dy \right) &\leq \exp\left(\frac{M}{2\tau}\right) \sum_{j=1}^{n-1} u_j(a_j - a_{j+1}) \phi_\tau(x - a_j) \\ &\leq \exp\left(\frac{M}{2\tau}\right) \sum_{j=1}^{n-1} \sigma^2 \phi_\tau(x - a_j) \leq \exp\left(\frac{\sigma^2}{\tau p}\right) \sum_{j=1}^{n-1} \sigma^2 \left(\frac{p_j}{p}\right) \phi_\tau(x - a_j), \end{aligned}$$

as required.

Similarly, for $x \leq 0$, we deduce that for $y \in I_j$, $\min_{y \in I_j} y^2 \geq (x - a_{j+1})^2 - M$, and thus

$$\begin{aligned} \sum_{j=1}^{n-1} u_j \left(\int_{x-a_j}^{x-a_{j+1}} \phi_\tau(y) dy \right) &\leq \exp\left(\frac{M}{2\tau}\right) \sum_{j=1}^{n-1} u_j(a_j - a_{j+1}) \phi_\tau(x - a_{j+1}) \\ &\leq \exp\left(\frac{\sigma^2}{\tau p}\right) \sum_{j=1}^{n-1} \sigma^2 \left(\frac{p_{j+1}}{p}\right) \phi_\tau(x - a_{j+1}). \end{aligned}$$

Theorem 1.1 is proved.

Lemma 2.1. *Using the notation above: $u_j(a_j - a_{j+1}) \leq \sigma^2$ and $Mp \leq 2\sigma^2$.*

P r o o f. Note that $u_j = \sum_{i=1}^j p_i a_i = -\sum_{i=j+1}^n p_i a_i$.

For $a_{j+1} \geq 0$: $u_j(a_j - a_{j+1}) \leq u_j a_j \leq \sum_{i=1}^j p_i a_i a_j \leq \sum_{i=1}^j p_i a_i^2$.

For $a_j \leq 0$: $u_j(a_j - a_{j+1}) \leq -u_j a_{j+1} \leq \sum_{i=j+1}^n p_i a_i a_{j+1} \leq \sum_{i=j+1}^n p_i a_i^2$.

For $a_{j+1} \leq 0 \leq a_j$: $u_j(a_j - a_{j+1}) \leq \sum_{i=1}^j p_i a_i a_j u_j + \sum_{i=j+1}^n p_i a_i a_{j+1} \leq \sum_{i=1}^n p_i a_i^2$.

For the second part, we consider two cases, firstly where $M = a_s^2 - a_{s+1}^2$. In this case $\sigma^2 = \sum_i p_i a_i^2 \geq p_s a_s^2 = p_s(a_s^2 - a_{s+1}^2) + p_s a_{s+1}^2 \geq pM$. Alternatively, if $M = (a_s - a_{s+1})^2$, then $\sigma^2 = \sum_i p_i a_i^2 \geq p(a_s^2 + a_{s+1}^2) \geq p(a_s - a_{s+1})^2/2 \geq pM/2$.

3. Convergence of the Poincaré constant. We establish an explicit rate of convergence of the Poincaré constant using projection inequalities similar to those in [9].

Lemma 3.1. *Given independent random variables X, Y with Poincaré constants R_X, R_Y , for any function g :*

$$\mathbf{D}g(X+Y) \leq (R_X + R_Y) \mathbf{E}g'(X+Y)^2 - \frac{R_X}{R_X I(Y) + 1} \mathbf{D}g'(X+Y),$$

and hence $R_{X+Y} \leq R_X + R_Y$.

P r o o f. Without loss of generality, we can consider g such that $\mathbf{E}g(X+Y) = 0$, and define $h(u) = \mathbf{E}_Y g(u+Y)$, which thus also has mean zero. Now

$$\begin{aligned} \mathbf{D}g(X+Y) &= \mathbf{E}g^2(X+Y) = \mathbf{E}_X \left(\mathbf{E}g^2(X+Y) \mid X \right) \\ &= \mathbf{E}_X \mathbf{D}(g(X+Y) \mid X) + \mathbf{E}_X (\mathbf{E}(g(X+Y) \mid X))^2 \\ &\leq R_Y \mathbf{E}_X \left(\mathbf{E}_Y (g'(X+Y)^2 \mid X) \right) + \mathbf{E}h(X)^2 \leq R_Y \mathbf{E}g'^2(X+Y) + R_X \mathbf{E}h'(X)^2. \end{aligned}$$

To consider the second term, we use the score function ρ_Y and define $f(x) = \mathbf{E}_Y [(g'(x+Y) - h'(x)), \rho_Y(Y)]$, where by the Stein equation, $f(x) = -\mathbf{E}_Y g''(x+Y) = -h''(x)$. Further, by Cauchy-Schwarz, $\mathbf{E}h''(X)^2 = \mathbf{E}f(X)^2 \leq I(Y) \mathbf{E}(g'(X+Y) - h'(X))^2$, so that $\mathbf{D}h'(X) \leq R_X \mathbf{E}h''(X)^2 \leq R_X I(Y) (\mathbf{E}g'(X+Y)^2 - \mathbf{E}h'(X)^2)$, and writing $\mu = \mathbf{E}h'(X) = \mathbf{E}g'(X+Y)$, we obtain $\mathbf{E}h'(X)^2(1 + R_X I(Y)) \leq R_X I(Y) \mathbf{E}g'(X+Y)^2 + \mu^2$, which, rearranging, leads to

$$\mathbf{E}h'(X)^2 \leq \mathbf{E}g'(X+Y)^2 - \frac{\mathbf{D}g'(X+Y)}{R_X I(Y) + 1}.$$

Lemma 3.1 is proved.

Next we need a lemma which again uses the idea that if g' is nearly constant, then g is close to linear. We would like to apply it to the optimal g , which achieves the maximum in Definition 1.1. However, rather than use compactness arguments to show such a function exists, we can use a «good» g instead.

Lemma 3.2. For any random variable W with mean zero and any function g such that $R(t) = \mathbf{D}(g(W) + tW)/\mathbf{E}(g'(W) + t)^2$ has a local maximum at $t = 0$:

$$\left(\frac{\mathbf{D}g(W)}{\mathbf{E}g'(W)^2} - \mathbf{D}W \right) \leq 3R_W \sqrt{\frac{\mathbf{D}g'(W)}{\mathbf{E}g'(W)^2}}.$$

P r o o f. Without loss of generality assume that $\mathbf{E}g(W) = 0$ and write $\mu = \mathbf{E}g'(W)$ and $\delta_g = \mathbf{D}g'(W)/\mathbf{E}g'(W)^2 = 1 - \mu^2/\mathbf{E}g'(W)^2$. Then

$$\begin{aligned} \mathbf{E}g^2(W) &= \mathbf{E}(g(W) - \mu W)^2 + 2\mu\mathbf{E}W(g(W) - \mu W) + \mu^2\mathbf{E}W^2 \\ &\leq \mathbf{E}(g(W) - \mu W)^2 + 2|\mu|\sqrt{\mathbf{D}W}\sqrt{\mathbf{E}(g(W) - \mu W)^2} + \mu^2\mathbf{D}W \\ &\leq (\mathbf{E}g'(W)^2) \left(R_W \left(\delta_g + 2\sqrt{\delta_g(1 - \delta_g)} \right) + (1 - \delta_g)\mathbf{D}W \right) \\ &\leq (\mathbf{E}g'(W)^2) \left(3R_W\sqrt{\delta_g} + \mathbf{D}W \right) \end{aligned}$$

since $\mathbf{E}(g(W) - \mu W)^2 \leq R_W\mathbf{E}(g'(W) - \mu)^2 = R_W\mathbf{D}g'(W) \leq R_W\delta_g\mathbf{E}g'(W)^2$, and since by Lemma 1.1, part 3, $\mathbf{D}W \leq R_W$. Lemma 3.2 is proved.

Note, we can come up with tighter bounds: for example taking $h(W) = W$ in equation (1), $\mathbf{E}Wg(W) = R(0)\mu$. Hence $\mu(R(0) - \mathbf{D}W) = \mathbf{E}(g(W) - \mu W)W \leq \sqrt{\mathbf{D}W}\sqrt{\mathbf{E}(g(W) - \mu W)^2} \leq \sqrt{R_W\mathbf{D}W\mathbf{D}g'(W)}$. This implies that

$$R(0) - \mathbf{D}W \leq R_W \sqrt{\frac{\delta_g}{1 - \delta_g}}.$$

However, Lemma 3.2 is sufficient for our purposes.

P r o o f o f T h e o r e m 1.2. We consider convergence along the «powers of 2» subsequence $S_k = U_{2^k}$, which implies convergence for the whole sequence by subadditivity.

For all k , $1 \leq R_{S_k} \leq R$, and $I_{S_k} \leq I$. Taking $X = S_k/\sqrt{2}$ and $Y = S'_k/\sqrt{2}$ (an identical copy) in Lemma 3.1 implies (since $R_{S_k/\sqrt{2}} = R_{S_k}/2$ and $I(S_k/\sqrt{2}) = 2I(S_k)$) that for any g

$$\delta_g = \frac{\mathbf{D}g'(S_{k+1})}{\mathbf{E}g'(S_{k+1})^2} \leq 2 \left(\frac{1}{R_{S_k}} + I(S_k) \right) \left(R_{S_k} - \frac{\mathbf{D}g(S_{k+1})}{\mathbf{E}g'(S_{k+1})^2} \right).$$

Now, given $W = S_{k+1}$, we can find h such that $\mathbf{D}h/\mathbf{E}h'^2 \geq \max(R - \varepsilon, \mathbf{D}W)$. Since $\mathbf{D}(h(W) + tW)/\mathbf{E}(h'(W) + t)^2$ tends to $\mathbf{D}W$ at $\pm\infty$, and has one maximum t_0 and one minimum, we can find $g(W) = h(W) + t_0W$, which satisfies the conditions of Lemma 3.2:

$$\left((R_{S_{k+1}} - \varepsilon - 1)_+ \right)^2 \leq \left(\frac{\mathbf{D}g(S_{k+1})}{\mathbf{E}g'(S_{k+1})^2} - 1 \right)^2 \leq 9R_{S_{k+1}}^2\delta_g \leq C(R_{S_k} - R_{S_{k+1}} + \varepsilon),$$

where $x_+ = \max(x, 0)$ and $C = 18R_{S_{k+1}}^2(1/R_{S_k} + I(S_k)) \leq 18R(IR + 1)$. That is, since ε is arbitrary,

$$(R_{S_{k+1}} - 1)^2 \leq C(R_{S_{k+1}} - R_{S_k}). \quad (3)$$

Note that since R_{S_k} is decreasing and bounded below, successive differences tend to zero, and thus $R_{S_k} \rightarrow 1$.

To obtain a rate, write $u_k = (R_{S_k} - 1)/C$, equation (3) gives $u_k(1 + u_k) \leq u_{k-1}$. Since u_k are decreasing: $\dots \leq u_{n+2}^2 \leq u_{n+1}^2 \leq u_n^2 \leq u_{n-1} - u_n$, and hence $u_n^2 \leq u_{n-1} - u_n$, $u_n^2 \leq u_{n-2} - u_{n-1}$, $u_n^2 \leq u_{n-3} - u_{n-2}$, and so on. Summing, we obtain that for $m \leq n$: $(n - m)u_n^2 \leq u_m - u_n \leq u_m$. Taking $n = 2^r$, $m = 2^{r-1}$ implies that $u_{2^r} \leq \sqrt{u_{2^{r-1}}/2^{r-1}}$. Repeating this N times, we deduce that (since $u_k \leq u_1 \leq 1$)

$$u_{2^r} \leq 2^{-\sum_{j=1}^N (r-j)/2^j} = 2^{-r+2} 2^{(r-2-N)/2^N},$$

so if $N = r - 2$, then $u_{2^r} \leq 4/2^r$, and we can «fill in the gaps» by subadditivity, to show that $u_k \leq 16/k$ for all k . Theorem 1.2 is proved.

Acknowledgments. The author is a Fellow of Christ's College, Cambridge, who helped support a trip to Yale University during which many useful discussions with Andrew Barron took place. Yurii Suhov of Cambridge University and Alexander Holroyd of UCLA provided useful advice, and the anonymous referee provided several extremely useful improvements to the proofs.

REFERENCES

1. Barron A. R. Entropy and the central limit theorem. — Ann. Probab., 1986, v. 14, № 1, p. 336–342.
2. Биллингсли П. Сходимость вероятностных мер. М.: Наука, 1977, 352 с.
3. Боровков А. А., Утес С. А. Об одном неравенстве и связанной с ним характеристизации нормального распределения. — Теория вероятн. и ее примен., 1983, т. 28, № 2, с. 209–218.
4. Brown L. D. A proof of the central limit theorem motivated by the Cramér–Rao inequality. — Statistics and Probability: Essays in Honor of C. R. Rao. Ed. by G. Kallianpur, P. R. Krishnaiah, and J. K. Ghosh. New York: North-Holland, 1982, p. 141–148.
5. Cacoullos Th. On upper and lower bounds for the variance of a function of a random variable. — Ann. Probab., 1982, v. 10, № 3, p. 799–809.
6. Chen L. H. Y. An inequality for the multivariate normal distribution. — J. Multivariate Anal., 1982, v. 12, № 2, p. 306–315.
7. Chen L. H. Y., Lou J. H. Asymptotic normality and convergence of eigenvalues. — Stochastic Process. Appl., 1990, v. 34, № 2, p. 197.
8. Chernoff H. A note on an inequality involving the normal distribution. — Ann. Probab., 1981, v. 9, № 3, p. 533–535.
9. Johnson O. T., Barron A. R. Fisher information inequalities and the central limit theorem. — math. PR/0111020.
10. Klaassen C. A. J. On an inequality of Chernoff. — Ann. Probab., 1985, v. 13, № 3, p. 966–974.
11. Nash J. Continuity of solutions of parabolic and elliptic equations. — Amer. J. Math., 1958, v. 80, p. 931–954.
12. Utev S. A. An application of integrodifferential inequalities in probability theory. — Siberian Adv. Math., 1992, v. 2, № 4, p. 164–199.

Поступила в редакцию
5.I.2001

Исправленный вариант
24.VI.2002

© 2003 г.

MAUCH F.*

ON THE EXISTENCE OF PROBABILITY DISTRIBUTIONS WITH GIVEN MARGINALS

Пусть $X = \{0, \dots, n-1\}$ и $\Gamma := \{(x_1, \dots, x_s) \in X^s: \sum_{\sigma=1}^s x_{\sigma} = n-1\}$. Дается простая характеристизация маргинальных распределений вероятностных законов на Γ , обладающих тем дополнительным свойством, что они образуют s -кортеж убывающих вероятностей на X . Эта характеристизация имеет интересное применение к асимптотическим спектрам в смысле Штрассена [5], [6]. Некоторые смежные вопросы обсуждаются в приложении.

Ключевые слова и фразы: вероятностный закон, маргинальные распределения, асимптотические спектры в смысле Штрассена, необходимые и достаточные условия.

* University of Constance, Germany.