ARKIV FÖR MATEMATIK Band 7 nr 7

1.66 08 3 Communicated 11 May 1966 by O. FROSTMAN and U. GRENANDER

Multi-dimensional integral limit theorems for large deviations

By BENGT VON BAHR

1. Introduction

The problem of large deviations in the central limit theorem was first treated by Khintchine [3] in a special case and later by Cramér [2] in a more general one-dimensional case. His results were slightly improved by Petrov [4], who studied the distribution of sums of independent but not necessarily identically distributed random variables. Richter has proved local central limit theorems in the one-dimensional case [5] and in the multi-dimensional case [6], when the distribution of the sum is either absolutely continuous or of lattice type. He has also stated theorems of integral type [7], but, as he pointed out, he was obliged to restrict himself to the above-mentioned special cases, mainly because the ordinary integral limit theorems were lacking.

Here I want to use the results obtained in [1] to generalize Richter's results in [7] and one of Cramér's results [2] to the multi-dimensional case. I shall only treat the case of a sum of independent and identically distributed random vectors (r.v.'s), the generalization to non-identically distributed r.v.'s being straightforward but somewhat cumbersome.

2. Statement of the problem

Let $X=(X_1, ..., X_k)$ be a r.v. in R_k, $k>1$, with the distribution function (d.f.) $F(x)$, $x=(x_1, ..., x_k)$, with zero mean and non-singular covariance matrix M. Furthermore let, for some $h_0>0$, the moment generating function (m.g.f.) of X,

$$R(t)=\int_{R_k} e^{(t,x)}dF(x), \quad (t, x)=\sum_{j=1}^{k} t_j x_j$$

exist for all $t=(t_1, \dots, t_k)$ with $|t|=(\sum_{j=1}^{k} t_j^2)^{\frac{1}{2}}<h_0$

If $X^{(1)}, ..., X^{(n)}$ is a sequence of independent r.v.'s with the same d.f.'s as X, and $Y_n=(1/\sqrt{n})\sum_{\nu=1}^{n} X^{(\nu)}$, the problem is to estimate the probability $P(Y_n \in B)$, where B is a Borel set of a type specified in section 4. In Theorem 1, B is contained in a sphere with its center in the origin and of radius $R \leqslant \varepsilon_0\sqrt{n}$, and in Theorem 2, B is contained in the complement of such a sphere. In Theorems 3 and 4, I give applications to the d.f.'s of $|Y_n|$ and Y_n respectively.

3. Transformation of the distribution function

Following Cramér, we introduce for a fixed $h \in R_k$, $|h| < h_0$, the d.f. $F(x, h)$ defined by

$$dF(x, h) = \frac{e^{(h, x)} dF(x)}{R(h)}.$$

Let $X(h) = (X_1(h), \ldots, X_k(h))$ be a r.v. with the d.f. $F(x, h)$, with the mean $m = m(h)$ and the non-singular covariance matrix $M(h)$. $X(h)^{(1)}, \ldots, X(h)^{(n)}$ is a sequence of independent r.v.'s with the same d.f. as $X(h)$ and $Y_n(h) = (1/\sqrt{n})(\Sigma_{\nu=1}^{n} X(h)^{(\nu)} - nm)$ is its normed sum. The m.g.f. of $X(h)$ is

$$R(t, h) = \frac{R(t+h)}{R(h)}$$

and that of $\sum_{\nu=1}^{n} X(h)^{(\nu)}$ is

$$R_n(t, h) = (R(t, h))^n = \frac{R_n(t+h)}{R_n(h)} \tag{1}$$

where $R_n(t) = (R(t))^n$ is the m.g.f. of $\sum_{\nu=1}^{n} X^{(\nu)}$. If $G_n(x)$ and $G_n(x, h)$ are the d.f.'s of $\sum_{\nu=1}^{n} X^{(\nu)}$ and $\sum_{\nu=1}^{n} X(h)^{(\nu)}$ respectively, then according to (1)

$$dG_n(x, h) = \frac{e^{(h, x)} dG_n(x)}{R_n(h)}.$$

This relation can be written

$$dF_n(x) = R^n(h) e^{-\sqrt{n}(h, x)} dF_n(x - m\sqrt{n}, h) \tag{2}$$

where $F_n(x)$ and $F_n(x, h)$ are the d.f.'s of Y_n and $Y_n(h)$ respectively. We shall use (2) to estimate the probability that Y_n will fall into a set in the neighbourhood of the point $m\sqrt{n}$. Now, $m = m(h)$ is, for $|h| < h_0$, given by

$$m = \frac{\int x e^{(h, x)} dF(x)}{R(h)} = Mh + O(|h|^2)$$

where Mh is a vector, the ith component of which is $\Sigma_{j=1}^{k} M_{ij} h_j$. The Jacobian of the transformation $h \to m(h)$ is $|M(h)| = det(M(h)) > 0$ and thus the transformation is invertible, and we obtain

$$h = h(m) = \Lambda m + O(|m|^2) \tag{3}$$

where $\Lambda = M^{-1}$ (the inverse matrix of M). Consequently, there exists an $\varepsilon_0 > 0$ such that, for every $x \in R_k$ with $|x|/\sqrt{n} < \varepsilon_0$, h can be chosen so that $m = x/\sqrt{n}$.

According to the central limit theorem [1], $F_n(x, h)$ is approximatively a normal d.f., and therefore we shall approximate $dF_n(x)$ by $dW_n(x) = w_n(x)dx$, where dx is the volume element of R_k, and

$$w_n(x) = (2\pi)^{-k/2} |M|^{-\frac{1}{2}} R^n(h(x/\sqrt{n})) e^{-\sqrt{n}(h(x/\sqrt{n}), x)}$$

We put $d(v)=(h(v), v)-\log R(h(v))$ for $|v|\leqslant\varepsilon_0$, and thus obtain

$$w_n(x)=(2\pi)^{k/2}|M|^{-\frac{1}{2}}e^{-nd(x/\sqrt{n})}.$$

The function $d(v)$ is analytic for $|v|<\varepsilon_0$, and a simple calculation of the MacLaurin expansion gives

$$d(v)=\tfrac{1}{2}(v, \Lambda v)-\sum_{\nu=3}^{\infty}Q_\nu(v) \tag{4}$$

where $Q_\nu(v)$ are homogeneous polynomials of degree ν, the coefficients of which are functions of the semi-invariants of $F(x)$ of order not greater than ν.

4. A class of Borel sets

The central limit theorem in R_k is proved for a class $\mathcal{B}_1$ of Borel sets [1], and it is probable that the following estimations may be carried out with the appropriate modifications in this class. However, I shall confine myself to the class $\mathcal{D}$ of logical differences between convex Borel sets, that is, $D\in\mathcal{D}$ if $D=A_1\cap A_2'$, where A_1 and A_2 are convex Borel sets. Without loss of generality, we can assume that A_1 is the convex hull of D and $A_2\subset A_1$, and thus we write $D=A_1-A_2$. We define for every $\delta>0$ the exterior parallel set B_δ of a Borel set B by $B_\delta=\bigcup_{|u|<1}(B+\delta u)$, where $B+\delta u$ is the translate of B by δu, and the union is taken over all $u\in R_k$ with $|u|<1$. We denote by $V(B)$ the k-dimensional volume of the set B, and by $S(B)$ the $(k-1)$-dimensional area of the boundary points of the set B, both being defined for $B\in\mathfrak{C}=$ the class of all convex Borel sets.

5. Two lemmas

We first prove the following lemma, which gives an estimate of $F_n(D)$ for a small set $D\in\mathcal{D}$ belonging to the sphere $\{x:|x|\leqslant\varepsilon_0\sqrt{n}\}$, where $\varepsilon_0>0$ is independent of n and sufficiently small.

Notations. C and c are unspecified positive finite constants, θ satisfies $|\theta|\leqslant C$, and $O(z)$ stands for a function satisfying $|O(z)|\leqslant Cz$ for $z>0$.

Lemma 1. *If $D=A_1-A_2\in\mathcal{D}$ and D is a subset of both the spheres $\{x:|x|\leqslant R\}$ and $\{x:|x-a|\leqslant 1/R\}$ for some $a\in R_k$ and $1\leqslant R\leqslant\varepsilon_0\sqrt{n}$, then*

$$F_n(D)=W_n(D)\,(1+O(R/\sqrt{n}))+(\theta/\sqrt{n})e^{-nd(a/\sqrt{n})}S((A_1)_{c/\sqrt{n}}).$$

Proof. Putting $h=h(a/\sqrt{n})$ we get $m\sqrt{n}=a$, and thus form (2)

$$F_n(D)=R^n(h(a/\sqrt{n}))\int_D e^{-\sqrt{n}(x,\,h(a/\sqrt{n}))}dF_n(x-a,\ h(a\sqrt{n}))=e^{-nd(a/\sqrt{n})}I,$$

where
$$I=\int_{D-a}e^{-\sqrt{n}(x,\,h)}dF_n(x, h)$$

According to the central limit theorem [1],

$$F_n(x, h) = \Phi(x, h) + H_n(x, h) \tag{5}$$

where $\Phi(x, h)$ is the normal d.f. with zero mean and convariance matrix $M(h)$, and

$$|H_n(A, h)| \leqslant \frac{C}{\sqrt{n}}\{V(A_{c/\sqrt{n}}) + S(A_{c/\sqrt{n}})\} \tag{6}$$

for every $A \in \mathfrak{C}$.

From (5) we get $I = I_1 + I_2$, where

$$I_1 = (2\pi)^{-k/2}|M(h)|^{-\frac{1}{2}} \int_{D-a} e^{-\sqrt{n}(x,\, h) - \frac{1}{2}(x,\, \Lambda(h)x)}dx$$

$$= (2\pi)^{-k/2}|M|^{-\frac{1}{2}} \int_{D-a} e^{-\sqrt{n}(x,\, h) - \frac{1}{2}(x,\, \Lambda x)}dx(1 + O(|h|))$$

for the components of $\Lambda(h)$ are $\Lambda_{ij}(h) = \Lambda_{ij} + O(|h|)$, and $|x| \leqslant 1$ when $x \in D - a$.

We shall compare this expression for I_1 with

$$W_n(D) = (2\pi)^{-k/2}|M|^{-\frac{1}{2}}e^{-nd(a/\sqrt{n})} \int_{D-a} e^{-n[d((a+x)/\sqrt{n}) - d(a/\sqrt{n})]}dx.$$

From (4) we obtain if, $|u| + |v| \leqslant \varepsilon_0$ and ε_0 is sufficiently small

$$d(v + u) - d(v) = \tfrac{1}{2}(u, \Lambda u) + (u, \Lambda v) + O(|u| \cdot |v|^2)$$

and thus, because of (3),

$$n[d((a + x)/\sqrt{n}) - d(a/\sqrt{n}] = \tfrac{1}{2}(x, \Lambda x) + \sqrt{n}(x, h) + O(|x|\,|a|^2/\sqrt{n}).$$

Since $|x|\,|a|^2/\sqrt{n} = O(R/\sqrt{n})$ and $|h| = O(R/\sqrt{n})$, we get

$$W_n(D) = e^{-nd(a/\sqrt{n})}I_1(1 + O(R/\sqrt{n})).$$

It now remains to estimate I_2. We put

$$I_2 = \int_{D-a} = \int_{A_1 - a} - \int_{A_2 - a} = I_{21} - I_{22}.$$

For $\nu = 1$ and 2 respectively, we put

$$A_\nu(z) = (A_\nu - a) \cap \{x: n(h, x) \leqslant z\} \quad \text{and} \quad Q_\nu(z) = H_n(A_\nu(z), h).$$

Now, $A_\nu(z)$ is convex, and thus the inequality (6) holds for $Q_\nu(z)$ with $A = A_\nu(z)$ but since $V(A) \leqslant c_k R_A S(A)$, $A \in \mathfrak{C}$, where R_A is the radius of the sphere circumscribed A, we get, with a new value of C,

$$Q_\nu(z) \leqslant \frac{C}{\sqrt{n}} S((A_\nu)_{c/\sqrt{n}}) \leqslant \frac{C}{\sqrt{n}} S((A_1)_{c/\sqrt{n}}).$$

If
$$\alpha_\nu = \inf_{x \in A_\nu - a} \sqrt{n}(h, x)$$

and
$$\beta_\nu = \sup_{x \in A_\nu - a} \sqrt{n}(h, x)$$

then
$$|I_{2\nu}| = \left| \int_{\alpha_\nu}^{\beta_\nu} e^{-z} dQ_\nu(z) \right| \leqslant 2e^{-\alpha_\nu} \sup_{\alpha_\nu \leqslant z \leqslant \beta_\nu} |Q_\nu(z)|$$

and since $\alpha_\nu = O(\sqrt{n}|h| \cdot |x|) = O(1)$, we get

$$|I_2| \leqslant |I_{21}| + |I_{22}| \leqslant \frac{C}{\sqrt{n}} S((A_1)_{c/\sqrt{n}}).$$

The lemma is proved.

Lemma 2. *There exists a positive consant ζ such that*

$$P(|Y_n| > \varepsilon_0 \sqrt{n}) \leqslant Ce^{-\zeta n}.$$

Proof. It suffices to prove that for every j, $1 \leqslant j \leqslant k$ the component Y_{nj} of Y_n satisfies

$$P(|Y_{nj}| > \varepsilon_0 \sqrt{n/k}) \leqslant e^{-\zeta n}.$$

Since
$$R(h) = 1 + \tfrac{1}{2}(h, Mh) + O(|h|^3) \leqslant e^{c|h|^2}$$

for $|h|$ sufficiently small, we have

$$E(e^{\sqrt{n}\, h_j Y_{nj}}) = (R(0, \ldots, h_j \ldots, 0))^n \leqslant e^{cnh_j^2}$$

and thus from Chebyshev's inequality

$$P\,(|Y_{nj}| > \varepsilon_0 \sqrt{n/k}) \leqslant \frac{e^{cnh_j^2}}{e^{\varepsilon_0 h_j n/\sqrt{k}}} = e^{-\zeta n}$$

if h_j is sufficiently small.

6. Main theorems

The following two theorems are the fundamental limit theorems for large deviations in R_k, and can now easily be proved by summing estimates of the probabilities of small sets obtained in Lemma 1.

Theorem 1. *If $D \in \mathfrak{D}$ is a subset of the sphere $\{x: |x| \leqslant R\}$, where $1 \leqslant R \leqslant \varepsilon_0 \sqrt{n}$, then*

$$F_n(D) = W_n(D) + \theta \frac{R}{\sqrt{n}} W_n(D_{2/R}).$$

Proof. We divide D into a disjoint union of sets $D \cap K_\nu$, where K_ν are congruent half-open cubes with the edges parallel to the coordinate axes and with the edge length $d = 2/(R\sqrt{k})$.

Now $D \cap K_\nu \in \mathcal{D}$ and $D \cap K_\nu \subset \{x: |x - a_\nu| \leq 1/R\}$, if a_ν is the centre of K_ν, and we thus obtain from Lemma 1:

$$F_n(D) = W_n(D)\,(1 + O(R/\sqrt{n})) + \frac{\theta}{\sqrt{n}} \sum_\nu e^{-nd(a_\nu/\sqrt{n})} S((K_\nu)_{c/\sqrt{n}}).$$

We shall compare the terms in the sum with $W_n(K_\nu)$. We have

$$W_n(K_\nu) = (2\pi)^{-k/2} |M|^{-\frac{1}{2}} e^{-nd(a_\nu/\sqrt{n})} \int_{K_\nu - a_\nu} e^{-\sqrt{n}(x,\, h_\nu) - \frac{1}{2}(x,\, \Lambda x)} dx (1 + O(R/\sqrt{n}))$$

in the same way as in the proof of Lemma 1 ($h_\nu = h(a_\nu/\sqrt{n})$). Because the exponent in the above integrand is bounded, we get

$$W_n(K_\nu) \geq C R^{-k} e^{-nd(a_\nu/\sqrt{n})}$$

and thus because

$$S((K_\nu)_{c/\sqrt{n}}) \leq C R^{-k+1}$$

the sum is

$$\theta \frac{R}{\sqrt{n}} \sum_{\nu=1}^{N} W_n(K_\nu) \leq \theta \frac{R}{\sqrt{n}} W_n(D_{2/R}).$$

The theorem is proved.

Remark. As mentioned in Section 1, Richter [7] has studied the same problem when $F(x)$ is a lattice d.f. and when $F_m(x)$ is absolutely continuous, some $m \geq 1$. He considers sets B of the type

$$B = \{x : t_1 < |x| \leq t_2,\ x/|x| \in \Omega\}$$

where $0 \leq t_1 < t_2 = o(\sqrt{n})$ and Ω is a subset of the surface $\Omega_0 = \{x : |x| = 1\}$ with positive Lebesque measure. His proposition is

$$F_n(B) = W_n(B)\,(1 + O(t_2/\sqrt{n})).$$

This is obviously not correct if $F(x)$ is a lattice d.f.

For, let $\Omega = \Omega_1 \cup \Omega_2$, where Ω_1 has positive Lebesque measure, and Ω_2 is the denumerable set of points $x/|x|$ $(x \neq 0)$ corresponding to all points $x \in R_k$, with $F_m(\{x\}) > 0$ for some $m \geq 1$. Then

$$W_n(B) = \int_{\substack{x/|x| \in \Omega_1 \\ t_1 < |x| \leq t_2}} w_n(x)\,dx$$

but $F_n(B) = P(t_1 < |Y_n| \leq t_2)$, independently of Ω_1.

Theorem 2. *If* $D \in \mathcal{D}$ *and* $D \subset \{x: R \leqslant |x| \leqslant \varepsilon_0 \sqrt{n}\}$, *then*

$$F_n(D) = W_n(D) + \frac{\theta}{\sqrt{n}} \int_{D_{1/R}} |x| w_n(x) dx.$$

Proof. As in the proof of the preceding theorem, we divide D into a union of sets $D \cap K_\nu$ where K_ν are cubes with the edges parallel to the coordinate axes, but here the edge length of K_ν must depend on the distance to the origin. Consider a rectangular grid in R_k with the edge length $d = 1/(R\sqrt{k})$, and take out those cubes which lie in the sphere $\{x: |x| \leqslant 2R\}$. Divide each of the remaining cubes into 2^k congruent cubes with the edge length $d/2$ and take out those which lie in the sphere $\{x: |x| \leqslant 4R\}$, and so on. In this way, we obtain a finite number of cubes K_ν intersecting D, and we can apply Lemma 1 for each $D \cap K_\nu$. If a_ν is the centre and d_ν the edge-length of K_ν, then $1 < |a_\nu| d_\nu \sqrt{k} < 2$, and thus

$$F_n(D) = W_n(D) + \frac{\theta}{\sqrt{n}} \sum_\nu |a_\nu| W_n(D \cap K_\nu) + \frac{\theta}{\sqrt{n}} \sum_\nu e^{-nd(a_\nu/\sqrt{n})} d_\nu^{-k+1}.$$

The theorem follows after simple calculations.

The magnitude of the remainder terms, in proportion to the main terms in Theorem 1 and Theorem 2 depends on the relative difference in size (volume) between D and $D_{1/R}$. If this is negligible, as is the case if the dimensions of D are very large compared with $1/R$, we obtain the relation

$$F_n(D) = W_n(D)(1 + O(R/\sqrt{n}))$$

in both theorems.

We shall obtain results of this type in the following sections.

7. Applications to the distribution of $|Y_n|$

The following theorem was stated in a slightly different form by Richter [7] in the two special cases mentioned earlier. His proof is, however, not satisfactory.

Theorem 3. *There exists a constant* $\delta_0 > 0$ *such that, if* $1 \leqslant t \leqslant \delta_0 \sqrt{n}$, *then*

$$P(|Y_n| > t) = (2\pi)^{-k/2} |M|^{-\frac{1}{2}} \int_{u \in \Omega_0} \exp\left(n \sum_{\nu=3}^{\infty} (t/\sqrt{n})^\nu Q_\nu(u)\right) dS$$
$$\times \int_t^\infty e^{-(u,\,\Lambda u) y^2/2} y^{k-1} dy (1 + O(t/\sqrt{n}))$$

where dS is the surface element of $\Omega_0 = \{u: |u| = 1\}$.

Proof. Putting $D = \{x: t < |x| \leqslant \varepsilon_0 \sqrt{n}\}$, we immediately obtain from Theorem 2

$$F_n(D) = W_n(D) + \frac{\theta}{\sqrt{n}} \int_{D_{1/t}} |x| w_n(x) dx. \tag{7}$$

We have

$$W_n(D) = (2\pi)^{-k/2}|M|^{-\frac{1}{2}}\int_{u\in\Omega_0} dS$$
$$\times \int_t^{\varepsilon_0\sqrt{n}} \exp\left(-(u, \Lambda u)\, y^2/2 + n\sum_{\nu=3}^{\infty}(y/\sqrt{n})^{\nu}Q_{\nu}(u)\right) y^{k-1}dy \tag{8}$$

and we shall show that

$$W_n(D) = (2\pi)^{-k/2}|M|^{-\frac{1}{2}}\int_{u\in\Omega_0} \exp\left(n\sum_{\nu=3}^{\infty}(t/\sqrt{n})^{\nu}Q_{\nu}(u)\right)$$
$$\times \int_t^{\varepsilon_0\sqrt{n}} e^{-(u,\Lambda u)y^2/2}y^{k-1}dy(1+O(t/\sqrt{n})). \tag{9}$$

For that purpose, we form the absolute value of the difference between (8) and the main part of (9). It is at most equal to

$$I_1 = (2\pi)^{-k/2}|M|^{-\frac{1}{2}}\int_{u\in\Omega_0} \exp\left(n\sum_{\nu=3}^{\infty}(t/\sqrt{n})^{\nu}Q_{\nu}(u)\right) dS$$
$$\times \int_t^{\varepsilon_0\sqrt{n}} e^{-(u,\Lambda u)y^2/2}\left|\exp\left(n\sum_{\nu=3}^{\infty}(y^{\nu}-t^{\nu})n^{-\nu/2}Q_{\nu}(u)\right)-1\right|y^{k-1}dy. \tag{10}$$

We denote the inner integral by I_2, and obtain after simple estimations of the exponent, if ε_0 is sufficiently small,

$$I_2 \leqslant \int_t^{\varepsilon_0\sqrt{n}} e^{-(u,\Lambda u)y^2/2}(e^{cy^2(y-t)/\sqrt{n}}-1)\,y^{k-1}dy$$
$$= e^{-(u,\Lambda u)t^2/2}t^{k-2}\int_0^{nv(\varepsilon_0-v)} e^{-(u,\Lambda u)(z+z^2/2nv^2)}\,(1+z/nv^2)^{k-1}(e^{cvz(1+z/nv^2)^2}-1)dz$$

where we have put $y=t+z/t$ and $t=v\sqrt{n}$.

It is elementary to show that this integral is $O(v)$ for $1/\sqrt{n}\leqslant v\leqslant\varepsilon_0$, that is

$$I_2 \leqslant Ct^{k-1}e^{-(u,\Lambda u)t^2/2}/\sqrt{n} \leqslant C\frac{t}{\sqrt{n}}\int_t^{\varepsilon_0\sqrt{n}} e^{-(u,\Lambda u)y^2/2}y^{k-1}dy.$$

This result, introduced into (10), proves (9).

The second term of (7) is treated in a similar way and the result is that $F_n(D)$ is given by exactly the same formula (9) as $W_n(D)$. Clearly, we can also change the upper limit $\varepsilon_0\sqrt{n}$ to $+\infty$ in the second integral of (11), without breaking down the equality.

It remains to show that $F_n(\{x: |x| > \varepsilon_0\sqrt{n}\})$ is negligible compared with $F_n(D)$, but since $F_n(D)\sim e^{-ct^2}$, this follows from Lemma 2, if $t<\delta_0\sqrt{n}$ and δ_0 is sufficiently small. The proof is concluded.

By simple calculations, we obtain from Theorem 3 the following results, also stated by Richter [7] in slightly different forms:

$$P(|Y_n|>t+g/t)=t^{k-2}\int_{u\in\Omega_0} e^{-g(u,\Lambda u)}w_n(tu)(u,\Lambda u)^{-1}dS$$
$$\times(1+O((1+g^2)/t^2)+O((1+g)t/\sqrt{n}))$$

and if $M=E_k$ (unit matrix of order $k\times k$).

$$\frac{P(t<|Y_n|\leqslant t+g/t)}{P(|Y_n|>t)}=1-e^{-g}+O((1+g^2)/t^2)+O((1+g)t/\sqrt{n})$$

for $t\geqslant 1$, $0\leqslant g\leqslant t^2/2$ and $t+g/t\leqslant\delta_0\sqrt{n}$.

The last relation shows that the distribution of $|Y_n|$ asymptotically satisfies the same functional equation as the distribution of a one-dimensional Gaussian random variable with unit standard deviation. This is a generalization of a result obtained by Khintchine [3] and Cramér [2] in the one-dimensional case.

8. Application to the distribution function of Y_n

We now return to the relation (2) and shall use it to estimate $P(Y_{nj}>a_j,\ 1\leqslant j\leqslant k)$, where $1\leqslant a_j=o(\sqrt{n})$, when the components of X are uncorrelated. With no loss of generality, we may thus assume that $M=E_k$. The result is a direct generalization of one obtained by Cramér [2] in the one-dimensional case.

Theorem 4. *If $1\leqslant a_j=o(\sqrt{n})$ and $a_j\geqslant\alpha|a|$, $1\leqslant j\leqslant k$, for some positive constant α, then, if $M=E_k$,*

$$P(Y_{nj}>a_j,\ 1\leqslant j\leqslant k)\Big/\prod_{j=1}^{k}(1-\Phi(a_j))=\exp\left(n\sum_{\nu=3}^{\infty}Q_\nu(a/\sqrt{n})\right)\left(1+O\left(\frac{|a|}{\sqrt{n}}\right)\right)$$

where $\Phi(z)$, $z\in R_1$, is the normalized normal d.f.

Remark. The theorem cannot be true in an equivalent form for every covariance matrix $M\neq E_k$. For, according to Theorem 2 and Lemma 2 the probability concerned is approximated by

$$(2\pi)^{-k/2}|M|^{-\frac{1}{2}}\int_{\substack{\text{all } x_j>a_j\\ |x|<\varepsilon_0\sqrt{n}}}\exp\left((-x,\Lambda x)/2+n\sum_{\nu=3}^{\infty}Q_\nu(x/\sqrt{n})\right)dx$$

and this cannot for all a be almost equal to

$$\exp\left(n\sum_{\nu=3}^{\infty}Q_\nu(a/\sqrt{n})\right)(2\pi)^{-k/2}|M|^{-\frac{1}{2}}\int_{\text{all } x_j>a_j}e^{-(x,\Lambda x)/2}dx$$

unless the maximum of $e^{-(x,\Lambda x)/2}$ in $\{x:\text{all } x_j\geqslant a_j\}$ is attained in the point $x=a$.

Proof. Putting $h=h(a/\sqrt{n})$ in (2), we obtain

$$P(Y_{nj}>a_j,\ 1\leqslant j\leqslant k)=e^{-nd(a/\sqrt{n})}\int\limits_{\text{all } x_i>0} e^{-\sqrt{n}(h,x)}dF_n(x,h)$$

We denote the integral by I, and divide it according to (5) into I_1+I_2, where

$$I_1=(2\pi)^{-k/2}|M(h)|^{-\frac{1}{2}}\int\limits_{\text{all } x_j\geqslant 0} e^{-\sqrt{n}(h,x)-(x,\Lambda(h)x)/2}dx$$

From (3) we get

$$a_j/\sqrt{n}=h_j+O(|h|^2)$$

but $a_j>\alpha|a|$ implies $|h|=O(|h_j|)$, and thus we have

$$h_j\sqrt{n}=a_j(1+O(|a|/\sqrt{n}))\geqslant c$$

By using methods similar to those used to obtain (9) out of (8), we get

$$I_1=(2\pi)^{-k/2}\int\limits_{\text{all } x_i\geqslant 0} e^{-(a,x)-|x|^2/2}(1+O(|h|))=e^{|a|^2/2}\prod_{j=1}^{k}(1-\Phi(a_j))\left(1+O\left(\frac{|a|}{\sqrt{n}}\right)\right). \quad (11)$$

In order to estimate

$$I_2=\int\limits_{\text{all } x_j>0} e^{-\sqrt{n}(h,x)}dH_n(x,h)$$

we form for every $z>0$ the polyhedron

$$P(z)=\{x:\ \sqrt{n}(h,x)<z,\ \text{all } x_j>0\}$$

and put

$$K(z)=H_n(P(z),h)$$

We then get

$$I_2=\int_0^\infty e^{-z}dK(z)=\int_0^\infty e^{-z}K(z)\,dz.$$

Since $P(z)$ is convex, $K(z)$ satisfies an inequality of the type (6). Simple calculations give

$$|I_2|\leqslant\frac{C}{\sqrt{n}}\left(\prod_{j=1}^{k}(h_j\sqrt{n})\right)^{-1}\left(\sum_{j=1}^{k}h_j\sqrt{n}\right).$$

From (11) we get

$$I_1\geqslant C\left(\prod_{j=1}^{k}a_j\right)^{-1}$$

and thus $I_2/I_1=O(|h|)$. The theorem follows.

REFERENCES

1. von Bahr, B., Multi-dimensional integral limit theorems. *Arkiv för Matematik*, *7*, pp. 71–88 (1967).
2. Cramér, H., Sur un nouveau théorème-limite de la théorie des probabilités. Actualités Scientifiques et industrielles, Paris, *736*, s. 23 (1938).
3. Khintchine, A., Über einen neuen Grenzwertsatz der Wahrscheinlichkeitsrechnung. Math. Annalen *101*, 745–752 (1929).
4. Petrov, V. V., A generalization of Cramér's limit theorem, Uspekhi Mat. Nauk. *9* Vol. 4, 195–202 (1959) (in Russian).
5. Richter, V., Local limit theorems for large deviations. Theory of Probability and its Applications, *2*, 206–220 (1957).
6. Richter, V., Multi-dimensional local limit theorems for large deviations. Theory of Probability and its Applications, *3*, 100–106 (1958).
7. Richter, V., Multi-dimensional limit theorems for large deviations and their application to the χ^2-distribution. Theory of Probability and its Applications, *9* No. 1, 28–37 (1964).

Tryckt den 22 mars 1967

Uppsala 1967. Almqvist & Wiksells Boktryckeri AB