

TARTU ÜLIKOOLI  
TOIMETISED

---

УЧЕНЫЕ ЗАПИСКИ ТАРТУСКОГО УНИВЕРСИТЕТА  
ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS

---

893

STOCHASTIC MODELS

Matemaatika- ja mehaanikaalased tööd

Труды по математике и механике

TARTU ÜLIKOOLI TOIMETISED  
УЧЕНЫЕ ЗАПИСКИ ТАРТУСКОГО УНИВЕРСИТЕТА  
ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS

Alustatud 1893.a. VIHK 893 ВЫПУСК Основаны в 1893.g.

# STOCHASTIC MODELS

Matemaatika- ja mehaanikaalased tööd

Труды по математике и механике

TARTU 1990

Redaktsioonikolleegium

Ü.Lepik (esimees), L.Ainola, T.Arak, K.Kenk, M.Kilp,  
Ü.Lumiste, E.Reimers, E.Tiit, G.Vainikko  
Vastutav toimetaja: K.Pärna

Редакционная коллегия

Ю.Лепик (председатель), Л.Айнола, Т.Арак, Г.Вайникко,  
К.Кенк, М.Кильп, Ü.Лумисте, Э.Реймерс, Э.Тийт  
Ответственный редактор: К.Пярна

---

# ON LARGE DEVIATIONS FOR A SUM OF TYPE $\sum f(T^j t)$

G.Misevičius

## 1. Introduction and statement of results.

Let  $Tt$  be a mapping of  $[0,1]$  into itself defined by  $Tt = \{2t\}$  - the fraction of  $2t$ . If we express  $t$  in the form  $t = \frac{\varepsilon_1(t)}{2} + \dots + \frac{\varepsilon_n(t)}{2^n} = [\varepsilon_1(t), \varepsilon_2(t), \dots],$  then  $Tt = [\varepsilon_2(t), \varepsilon_3(t), \dots]$ . The coefficients  $\varepsilon_1, \varepsilon_2, \dots$  are independent r.v.s. with respect to the Lebesgue measure and they also are stationary. If the function  $f(t)$  is periodic with period 1, the relation  $f(Tt) = f(2t)$  takes place. We denote

$$S_n = S_n(t) = \sum_{j=1}^n f(2^j t), \quad Ef(t) = \int_0^1 f(t) dt = 0,$$

$$B_n^2 = DS_n^2 = \int_0^1 S_n^2(t) dt,$$

and let  $\text{mes}\{\dots\}$  signify the measure of the set of  $t \in [0,1]$  for which the condition in curly brackets is valid. As usually,  $\Phi(x)$  denotes the standard normal distribution function.  $\square$  indicates the end of the proofs.

The first limit theorem for  $S_n$  was proved by Fortet and Kac. I.A.Ibrahimov has given the optimal conditions for validity of that theorem. D.A.Moskvina established the first theorem for large deviations. In this paper some statements about large deviations for a class of unbounded functions will be proved.

**Theorem 1.** Let  $f(t)$  be a measurable function of bounded variation and

$$E|f(t)|^p = \int_0^1 |f(t)|^p dt \leq (p-2)! H_0 H_1^{p-2}, \quad (1)$$

where  $p = 3, 4, \dots, S+2$ ,  $H_0, H_1$  are the constants and  $S \leq (H_0 \sqrt{n} (\ln n)^{-1})^2 = \Delta_n$ .

Then, for  $1 \leq x \leq \sqrt{S}$ , the following relations for large deviations hold:

$$\frac{\text{mes}\{S_n > xB_n\}}{1 - \Phi(x)} = \exp\{(L(x)(1 + o_1(x) \frac{x+1}{\sqrt{S}})\}, \quad (2)$$

$$\frac{\text{mes}\{S_n < -xB\}}{\Phi(-x)} = \exp\{\tilde{L}(x)(1 + e_2 \tilde{f}_2(x) \frac{x+1}{\sqrt{S}})\},$$

where  $\tilde{f}_j(x)$ ,  $j=1,2$  is a bounded function, and power series  $\tilde{L}(x) = \sum_{k=0}^{\infty} x^{k+3}$  converges for  $|x| < \sqrt{2\Delta_n/2\sqrt{e}}$ . In this circle which the relation  $|\tilde{L}(x)| \leq 5|x|^3/4\Delta$  is valid. The coefficients  $\tilde{L}_k$ ,  $k=0,1,\dots$ , can be expressed by cumulants and for  $k \leq S-3$  these coefficients coincide with the coefficients of well known series of Cramer-Petrov.

**Theorem 2.** If for a function of bounded variation  $f(t)$  the estimation

$$\int |f(t)|^k dt \leq H_2(k!)^{1+\gamma} H^{-k-2} \quad (3)$$

is valid for  $k=3,4,\dots$ , then for  $0 \leq x < H(\sqrt{n})^{1/(1+2\gamma)}$  the following relations for large deviations hold

$$\frac{\text{mes}\{S_n > xB\}}{1-\Phi(x)} = \exp\{L_\gamma(x)(1 + e_1 f_1(x) \frac{x+1}{\Delta_\gamma})\}, \quad (4)$$

$$\frac{\text{mes}\{S_n < -xB\}}{1-\Phi(x)} = \exp\{L_\gamma(-x)(1 + e_2 f_2(x) \frac{x+1}{\Delta_\gamma})\},$$

where  $f_j(x)$  is a bounded function and

$$L_\gamma(x) = \sum_{3 \leq k < P} \lambda_k X^k + e(X/\Delta_\gamma)^3, \quad p = \begin{cases} (1/\gamma)-1, & \gamma > 0 \\ \infty, & \gamma = 0 \end{cases}$$

**Theorem 3.** If the conditions of the Theorem 2 hold, then

$$\sup_x |\text{mes}\{S_n < xB_n\} - \Phi(x)| \leq H_4 \left( \frac{\ln n}{\sqrt{n}} \right)^{1/(1+2\gamma)} \quad (5)$$

## 2. Auxiliary statements.

At first we shall introduce some notations. Put

$$[f]_j^{(u)}(t) = [f]_j^{(u)} = E\{[f]_j^{(u)} | e_{j+1}, \dots, e_{j+u}\} \quad (6)$$

(a conditional mean value) and

$$\eta_j^{(u)} = \frac{e_j}{2} + \frac{e_{j+1}}{2^2} + \dots + \frac{e_{j+u}}{2^u} \quad (7)$$

**Lemma 1.** The r.v.s.  $\eta_j^{(u)}$ ,  $j=1,2,\dots$  form a Markov chain.

Proof. It is evident that every  $\eta_j^{(u)}$  defines a unique set of  $\varepsilon_j, \dots, \varepsilon_{j+u}$ . Therefore, G-algebra  $F_u$  generated by variables  $\eta_{j_1}^{(u)}, \dots, \eta_{j_1}^{(u)}$  coincides with G-algebra generated by  $\varepsilon_{j_1}, \dots, \varepsilon_{j_1+u}$ . Using definition of the conditional probability and definition of Markov chain, now the lemma follows from the equality

$$\int_A \text{mes}\{\eta_1^{(u)} < \lambda - \left[ \frac{\varepsilon_j}{2} + \dots + \frac{\varepsilon_{j-j_1+u}}{2^{j-j_1+u}} \right]\} dt = \text{mes}\{t \in A, \eta_1^{(u)} < \lambda\}$$

which is valid for all  $A \in F$  and  $j_1 < j$ .  $\square$

Thus, variables  $[f]_j^{(u)}$  are connected into the Markov chain.

Further, let  $F_{ab}$  be the minimal G-algebra generated by  $\{\eta_m^{(u)}; a \leq m \leq b\}$  and  $\Omega_1$  the space of the values of  $\eta_1^{(u)}$ . The coefficient of ergodicity is equal to

$$\alpha_{1m} = 1 - \sup_{A \in F_m} \sup_{\eta_1^{(u)}, \eta_1 \in \Omega_1} |\text{mes}\{A | \eta_1^{(u)}\} - \text{mes}\{A | \eta_1^{(u)}\}|. \quad (8)$$

In this case we have

$$\alpha_{1m}^{(u)} = \begin{cases} 0 & \text{for } m-1 \leq u, \\ 1 & \text{for } m-1 > u+1, \end{cases} \quad (9)$$

because  $\varepsilon_1, \varepsilon_2, \dots$  are independent. At the same time  $\eta_1^{(u)}$  are u-independent. Put

$$\bar{\eta}_j^{(u)} = f(T^j t) - \eta_j^{(u)}, \quad S_n' = \sum_{j=1}^n [f]_j^{(u)}.$$

Lemma 2. The following relations are valid:

$$B_n^2 = DS_n = G^2 n + C_1, \quad (10)$$

$$B_n'^2 = DS_n' = G_u^2 n + C_2, \quad (11)$$

$$|G - G_u| \leq C_3/2^{u/2}. \quad (12)$$

It is easy to prove (see [2]) that

$$E|f(t) - \eta_1^{(u)}|^k \leq (\text{var } f)^k 2^{-u}, \quad (13)$$

in particular

$$E|f(t) - \eta_j^{(u)}|^2 \leq H_4^2 2^{-u}. \quad (14)$$

The last relation makes it possible to apply the calculus in [4] to obtain (10), (11) and (12).  $\square$

Examine the summ  $S'_n = \sum_{i=1}^n [f]_i^{(u)}$ . We must evaluate the cumulants  $\Gamma(S'_n)$  expressed by formula

$$\Gamma_k(S'_n) = \sum_{1 \leq t_1, \dots, t_k \leq n} \Gamma([f]_{t_1}^{(u)}, \dots, [f]_{t_k}^{(u)}), \quad (15)$$

where  $\Gamma([f]_{t_1}^{(u)}, \dots, [f]_{t_k}^{(u)})$  are mixed cumulants. For this purpose we use the formula of Statulevičius

$$\begin{aligned} \Gamma([f]_{t_1}^{(u)}, \dots, [f]_{t_k}^{(u)}) &= \\ &= \sum_{\nu=1}^k (-1)^{\nu-1} \sum_{\substack{\nu \\ \bigcup_{p=1}^{\nu} I_p = I}} N_{\nu}(I_1, \dots, I_{\nu}) \prod_{p=1}^{\nu} \hat{E}([f]_{I_p}^{(u)}) \end{aligned} \quad (16)$$

where the second summation is taken over all the partitions  $\{I_1, \dots, I_{\nu}\}$  of the set  $I$ . The integers  $N_{\nu}(I_1, \dots, I_{\nu})$ ,  $0 \leq \nu \leq N_{\nu}(I_1, \dots, I_{\nu}) \leq (\nu-1)!$ , depend on the set  $\{I_1, \dots, I_{\nu}\}$  and if  $N_{\nu}(I_1, \dots, I_{\nu}) > 0$ , then

$$\sum_{p=1}^{\nu} \max_{t_1, \dots, t_j \in I_p} (t_j - t_1) \geq \max_{1 \leq i, j \leq k} (t_j - t_i).$$

The symbol  $EY_1 \dots Y_k$  is defined recursively by

$$EY_1 \dots Y_k = EY_1 Y_2 \dots Y_k - \sum_{j=1}^{k-1} EY_1 \dots Y_j EY_{j+1} \dots Y_k \text{ for } k \geq 2$$

and  $EY_1 = EY_1$ .

If we put

$$\text{mes}_{t_{j-1}, t_j} \{x_{t_{j-1}}; A\} = \text{mes}\{\eta_{t_j}^{(u)} \in A | \eta_{t_{j-1}}^{(u)} = x_{t_{j-1}}\},$$

then

$$\begin{aligned} \hat{E}[f]_I^{(u)} &= \hat{E}([f]_{t_1}^{(u)} \dots [f]_{t_k}^{(u)}) = \\ &= \int \dots \int g_{t_1}(x_{t_1}) \text{mes}_{t_1} \{dx_{t_1}\} \times \\ &\times \prod_{j=2}^k g_{t_j}(x_{t_j}) (\text{mes}_{t_{j-1}, t_j} (x_{t_{j-1}}; dx_{t_j}) - \text{mes}_{t_j} (dx_{t_j})), \end{aligned}$$

where  $\text{mes}_t(B) = \text{mes}\{\eta_t^{(u)} \in B\}$  and  $g_t(x)$  is a  $F_t$ -measurable function.

As in [7], let

$$\Lambda_n(\alpha; w) = \max\{1, \max_{1 \leq s \leq w} \sum_{t=s}^n \alpha^{1/w}(s, t)\}.$$

**Lemma 3.** Under the conditions of Theorems 1 and 2 the estimation

$$\Gamma_k(S'_n) \leq (k!)^{1+\gamma} H_0^{\gamma} (H_1)^{k-2} u^{k-2} \quad (18)$$

takes place.

Proof. Following the reasoning on pages 94-95 of [7], we get from (9) that

$$\hat{E}[f]_{\mathbf{I}_p}^{(u)} \leq m_1^{(p)}! \dots m_p^{(p)}! H_0^{k-p-\varepsilon} \prod_{j=1}^p \alpha^{1/w} (l_j^{(p)}, l_{j+1}^{(p)})$$

where  $0 \leq \varepsilon \leq k_p$ ,  $w \geq 1$ ,  $H_0 \geq 0$ ,  $H_2 > 0$ ,  $1 \leq p \leq \nu$ ,  $1 < \nu < k$ , from which we deduce

$$\sum_{t \in \mathbf{n}} \Gamma(X_t) \leq nk! \delta^{k-1} H_0^{k-\varepsilon} \Delta_n^{k-1}(\alpha; w),$$

$$\mathbf{n} = (1, 2, \dots, n).$$

In our case, putting  $\varepsilon = 1$ ,  $w = 1$ , we see that

$$|\Gamma_k(S_n)| \leq n(k!)^{1+\gamma} H_1^k \Delta_n^{k-1}(\alpha, 1).$$

Due to (9),  $\Delta_n^k = u$ . Remembering the Lemma 2, we get assertion (18).  $\square$

Lemma 4 (Rudzkis, Saulis, Statulevičius, [7]). Let r.v.  $\xi$  with  $E\xi = 0$  and  $E\xi^2 = 1$  satisfy

$$|\Gamma_k(\xi)| \leq (k-2)!/\Delta^{k-2}, \quad k = 3, 4, \dots, S+2,$$

where  $S$  is even and  $S \leq 2\Delta^2$ . Then for  $x$ ,  $0 \leq x < \sqrt{S}/3\sqrt{e}$ , the following relations concerning large deviations are valid

$$\begin{aligned} \frac{1-F_\xi(x)}{1-\Phi(x)} &= \exp\{\tilde{L}(x)\} \left[ 1 + \Theta_1 \tilde{f}_1(x) \frac{x+1}{\sqrt{S}} \right], \\ \frac{F_\xi(-x)}{\Phi(-x)} &= \exp\{\tilde{L}(x)\} \left[ 1 + \Theta_2 \tilde{f}_2(x) \frac{x+1}{\sqrt{S}} \right], \end{aligned} \quad (19)$$

where

$$\tilde{f}_j(x) = \frac{117 + 96S \exp\{-\frac{1}{2}(1 - 3\sqrt{e} x/\sqrt{S})S^{1/4}\}}{(1 - 3\sqrt{e} x/\sqrt{S})},$$

$j=1, 2$ ;  $\tilde{L}(x) = \sum_{k=0}^{\infty} l_k x^{k+3}$  and this power series converges when

$|x| < \sqrt{2} \Delta / 3\sqrt{e}$ . In this circle  $|\tilde{L}(x)| \leq 5|x|^3/4\Delta$ . The coefficients  $l_k$ ,  $k = 0, 1, 2, \dots$  may be expressed by the first  $r_k = \min\{k+3, S\}$  cumulants of r.v.  $\xi$ , for  $k \leq S-3$  the coefficients being identical with classic series of Cramer-Petrov.

Lemma 5 (the same authors as in Lemma 4): If r.v.  $\xi$  with  $E\xi = 0$  and  $E\xi^2 = 1$  satisfies



$$|\Gamma_k(\xi)| \leq (k!)^{1+\gamma} / \Delta^{k-2}, \quad k = 3, 4, \dots \quad (20)$$

then in interval  $0 \leq x < \Delta_\gamma$  the following relations for large deviations hold

$$\frac{1 - F_\xi(x)}{1 - \bar{\Phi}(x)} = \exp\{L_\gamma(x)\} \left\{ 1 + \Theta_1 f_1(x) \frac{x+1}{\Delta_\gamma} \right\},$$

$$\frac{F_\xi(-x)}{\bar{\Phi}(-x)} = \exp\{L_\gamma(-x)\} \left\{ 1 + \Theta_2 f_2(x) \frac{x+1}{\Delta_\gamma} \right\},$$

where

$$f_j(x) = \frac{60(1 + 10\Delta_\gamma^2 \exp\{-(1-x/\Delta_\gamma)\sqrt{\Delta_\gamma}\})}{1 - x/\Delta_\gamma}, \quad j=1, 2,$$

$$L_\gamma(x) = \sum_{3 \leq k \leq p} \lambda_k X^k + \Theta(x/\Delta_\gamma)^3, \quad p = \begin{cases} 1/\gamma - 1, & \gamma > 0 \\ \infty, & \gamma = 0 \end{cases}$$

$$|\lambda_k| < \frac{2}{k} \left( \frac{16}{\Delta_\gamma} \right)^{k-2} ((k+1)!)^\gamma, \quad k = 3, 4, \dots, \quad |\Theta_1| \leq 1, \quad i=1, 2,$$

and

$$-x^3/\Delta_\gamma \leq L_\gamma(\pm x) \leq (x^2/2)(x/x + 8\Delta_\gamma).$$

Lemma 6. [7]. If for r.v.  $\xi$  the condition (20) holds, then

$$\sup_x |F_\xi(x) - \bar{\Phi}(x)| \leq 18/\Delta_\gamma, \quad \Delta_\gamma = c_\gamma \Delta^{1/(1+2\gamma)}.$$

### 3. Proofs of the Theorems.

If we put in (9)

$$u = \delta \ln n, \quad (21)$$

then the estimation (12) gives us

$$|G - G_u| \leq c/n^\delta \quad (22)$$

which is equivalent to

$$E^{1/k} |[f]_j^{(u)}|^k \leq E^{1/k} |f(t)|^k + \text{Var} f \cdot n^{-\delta},$$

$$E |[f]_j^{(u)}|^k \leq H H_n^{k-2} k!, \quad j=1, 2, \dots \quad (23)$$

Proof of the Theorem 1. Under the conditions of this theorem, in virtue of (18) and (23), we have

$$|\Gamma_k \left( \frac{S_n^-}{G_u \sqrt{n}} \right)| \leq (k!)^{1+\gamma} H_0 \left( \frac{\delta_1 H_1 \ln n}{G_u \sqrt{n}} \right)^{p-2} \quad (24)$$

Thus we can use the Lemma 4 with  $\Delta = \Delta_1 = \frac{G\sqrt{n}}{S_1 \ln n}$  and obtain the large deviation for  $S'_n/G\sqrt{n}$ . For the transition to  $S_n/G\sqrt{n}$  we evaluate the difference  $|S_n/G\sqrt{n} - S'_n/G\sqrt{n}|$ .

By the inequalities of Tchebychev and Hölder

$$\text{mes} \left\{ \left| \left( \frac{S_n}{G\sqrt{n}} \right) - \left( \frac{S'_n}{G\sqrt{n}} \right) \right| > \frac{1}{\delta_1} \right\} \leq \frac{c_3}{\delta_1^2} \quad (25)$$

where  $\delta_2 = \delta - \delta_1 - 2$ .

The obvious estimations are valid:

$$|\bar{\Phi}(x + \varepsilon) - \bar{\Phi}(x)| \leq c_3 \varepsilon_n \exp \left\{ -\frac{x^2}{2} \right\} \quad (26)$$

for  $\varepsilon > 0$  and  $x \geq 1$ ,

$$\frac{1}{4\sqrt{2n}(x+1)} \leq (1 - \bar{\Phi}(x)) \exp \left\{ \frac{x^2}{2} \right\} \quad (27)$$

for  $x \geq 0$ . If we choose  $\delta_1 > 2$  and  $\delta > 4$  (in (21)) we can state that

$$|\text{mes}\{S_n/G\sqrt{n} < x\} - \text{mes}\{S'_n/G\sqrt{n} < x + \varepsilon_n\}| \leq \frac{c^3}{n^2}, \quad (28)$$

where  $|\varepsilon_n| \leq c_2/n_2$ .

Having in mind that  $|L(x)| \leq 5|x|^3/4\Delta$ , the evaluations above give us the statement of the Theorem 1.  $\square$

Proof of the Theorem 2. The proof is based on the Lemma 5 and uses the same manipulations as in Theorem 1.  $\square$

Proof of the Theorem 3. The statement of this theorem follows from the Lemma 6 and (28). We get from Lemma 6 the estimation in the Central Limit Theorem for  $S'_n/G\sqrt{n}$ :

$$\sup_x |\text{mes}\{S'_n/G\sqrt{n} < x\} - \bar{\Phi}(x)| \leq c_7 \left( \frac{\ln n}{\sqrt{n}} \right)^{1/(1+2\gamma)}$$

The rest is evident.  $\square$

### References.

1. Heinrich, L. A Method for the derivation of limit theorems for sums of m-dependent random variables. Z. Wahrscheinlichkeitstheorie verw. Gebiete, 1982, **60**, 501-515.
2. Ibrahimov, I.A., Linnik, Yu.V. Independent and stationary connected variables. (Russ.); Nauka Publishers, Moscow, 1965.

3. Jakimavičius, D. On evaluation of cumulants and central moments for the mixing processes. (Russ.), LMR, 1988, 28, 614-626.
4. Misevičius, G. Asymptotic expansions for distribution functions of sums of the form  $\sum f(T^j t)$ . (Russ.). Annales Univ. Sci. Budapest. Ser. Math., 1971, 14, 77-92.
5. Moskvina, D.A. On large deviations of sums of the form  $\sum f(T^j t)$ . (Russ.), TVP, 1970, 15, 243-253.
6. Rudzakis, R., Saulis, L., Statulevičius, V. A general lemma about the probabilities of large deviations. (Russ.), LMR, 1978, 18, 99-116.
7. Saulis, L., Statulevičius, V. Limit theorems for large deviations. (Russ.). Mokslas Publishers, Vilnius, 1989.

О БОЛЬШИХ УКЛОНЕНИЙ ДЛЯ СУММ ТИПА  $\sum f(T^j t)$ .

Г.Мисявичюс

Р е з ю м е

В работе обобщаются результаты Д.А.Москвины и автора для больших отклонений сумм вида  $\sum f(T^j t)$ , а также приводится оценка в центральной предельной теореме.

Исследуются функции ограниченной вариации. В условиях (1) имеют место соотношения больших отклонений (2), а в условиях (3) — соотношения (4) и (5).

Received November 1989

# MOMENTS AND CUMULANTS OF MULTIVARIATE ELLIPTICAL DISTRIBUTION WITH SOME APPLICATIONS

I. Traat

Many statistical problems having simple solutions in the class of normal distributions, but being very complicated in the general case, appear to be quite easily solvable in the class of elliptical distributions too. This is caused by the fact that elliptical distributions have many common properties with multivariate normal distribution, which itself is a member of this class. For instance, the expressions of moments and cumulants of the elliptical distribution are similar in some sense. The fixed order central moments and cumulants of different elliptical distributions have the same functional relationship through second order cumulants. The difference appears in the constant multiplier only, which is determined by the concrete elliptical distribution. So the known expressions of central moments of multivariate normal distribution may be used for the representation of moments and cumulants of any multivariate elliptical distribution up to the constant multiplier.

The random  $p$ -vector  $x$  is said to have an elliptical distribution  $E_p(\mu, V)$  if its density function is of the form

$$f(x) = a_p |V|^{-1/2} h[(x-\mu)^T V^{-1} (x-\mu)] \quad (1)$$

for some function  $h$ , where  $V$  is a positive definite matrix,  $a_p$  is a normalizing constant.

The characteristic function of the elliptical distribution has the form

$$\varphi(t) = \exp(it^T \mu) \psi(t^T V t) \quad (2)$$

for some function  $\psi$ .

The expressions of cumulants of  $E_p(\mu, V)$  can be obtained by finding the partial derivatives of  $\ln \varphi(t)$ . After the differentiation we get the formulae (the cumulants up to the fourth order are given in [4])

$$E x = \mu, \quad \text{cov}(x) = (\kappa_{ij}) = \kappa_2 V, \quad (3)$$

$$\kappa_{ijkl} = \kappa_4 \sum_{s=1}^3 \kappa_{ij} \kappa_{kl} = \kappa_4 (\kappa_{ij} \kappa_{kl} + \kappa_{ik} \kappa_{jl} + \kappa_{il} \kappa_{jk}) \quad (4)$$

$$\mu_{ijklab} = k_5 \sum_{105} \mu_{ij} \mu_{kl} \mu_{ab}, \quad (5)$$

$$\mu_{ijklabcd} = k_6 \sum_{105} \mu_{ij} \mu_{kl} \mu_{ab} \mu_{cd}, \quad (6)$$

where the summation is carried out over all different products of  $\mu_{ij}$  like shown in (4). Note that the given sums present the central moments of order 4, 6, and 8 of the multivariate normal distribution with covariances  $\mu_{ij}$ .

The constants  $k_r$  are depending on the derivatives of  $\psi(u)$  with respect  $u$ ,  $u = t^T V t$ , in the following way:

$$k_2 = -2\psi'(0), \quad (7)$$

$$k_4 = \psi''(0)/\psi'^2(0) - 1, \quad (8)$$

$$k_6 = [\psi'''(0) - 3\psi''(0)\psi'(0)]/\psi'^3(0) + 2, \quad (9)$$

$$k_8 = [\psi^{(4)}(0) - 4\psi'''(0)\psi'(0) - 3\psi''^2(0) + 12\psi''(0)\psi'^2(0)]/\psi'^4(0) - 6. \quad (10)$$

The expressions of central moments of elliptical distribution can be found with the help of general relations between moments and cumulants (see [2],[3]), from which we get:

$$\mu_{ij} = \mu_{ij} = k_2 v_{ij}, \quad (11)$$

$$\mu_{ijkl} = c_4 \sum_{105} \mu_{ij} \mu_{kl}, \quad (12)$$

$$\mu_{ijklab} = c_5 \sum_{105} \mu_{ij} \mu_{kl} \mu_{ab}, \quad (13)$$

$$\mu_{ijklabcd} = c_6 \sum_{105} \mu_{ij} \mu_{kl} \mu_{ab} \mu_{cd}, \quad (14)$$

where

$$c_4 = k_4 + 1, \quad (15)$$

$$c_5 = k_5 + 3k_4 + 1, \quad (16)$$

$$c_6 = k_6 + 4k_5 + 3k_4^2 + 6k_4 + 1. \quad (17)$$

We can see that the central moments of an elliptical distribution depend on its second cumulants in the same way as its higher cumulants do. Hereby all the mixed central moments of order  $r$  differ from the corresponding mixed cumulants of the same order  $c_r/k_2$  times.

Using the expressions (7)-(10) it appears from (15)-(17) that the constants  $c_r$  have the following simple form in terms of derivatives of the characteristic function  $\psi(u)$ :

$$c_4 = \psi''(0)/\psi'^2(0), \quad (18)$$

$$c_6 = \psi'''(0)/\psi'^3(0), \quad (19)$$

$$c_8 = \psi^{(4)}(0)/\psi'^4(0). \quad (20)$$

If the distribution  $E_p(\mu, V)$  is fixed, then the constant  $c_r$  (or  $k_r$ ) is the same for all  $r$ -order mixed moments (or cumulants), i.e., the same for the  $r$ -order marginal moments too. This property allows us to find the constant  $c_r$  with the help of one-dimensional elliptical density function as described in Example 3.

Example 1. In the case of normal distribution with the characteristic function

$$\psi(u) = \exp(-u^2/2), \quad u = t^T V t,$$

we get from (7)-(10) and (15)-(17) the following values of constants:

$$k_2 = 1, \quad k_r = 0, \quad r \neq 2, \quad c_4 = c_6 = c_8 = 1.$$

Example 2. In the case of  $\epsilon$ -contaminated normal distribution with the characteristic function

$$\psi(u) = \epsilon \exp(-u^2/2) + (1-\epsilon) \exp(-\sigma^2 u^2/2),$$

the constants depend on  $\epsilon$  and  $\sigma$  in the following way:

$$k_2 = \epsilon + \sigma^2(1-\epsilon),$$

$$k_4 = [\epsilon + \sigma^4(1-\epsilon)]/k_2^2 - 1,$$

$$k_6 = [\epsilon + \sigma^6(1-\epsilon)]/k_2^3 - 3[\epsilon + \sigma^4(1-\epsilon)]/k_2^2 + 2,$$

$$k_8 = [\epsilon + \sigma^8(1-\epsilon)]/k_2^4 - 4[\epsilon + \sigma^6(1-\epsilon)]/k_2^3 - 3[\epsilon + \sigma^4(1-\epsilon)]^2/k_2^2 + \\ + 12[\epsilon + \sigma^4(1-\epsilon)]/k_2^2 - 6.$$

Example 3. Let us see the  $p$ -variate elliptical  $t$ -distribution on  $n$  degrees of freedom with the density function

$$f(x) = \frac{\Gamma(\frac{n+p}{2})}{\Gamma(\frac{n}{2})(n\pi)^{p/2}} |V|^{-1/2} \left[ 1 + \frac{1}{n}(x-\mu)^T V^{-1}(x-\mu) \right]^{-\frac{n+p}{2}} \quad (21)$$

The central moments of this distribution, if they exist, are expressed by the formulae (11)-(14), where the constants  $k_2$ ,  $c_4$ ,  $c_6$ ,  $c_8$  are found with the help of marginal density function of (21):

$$f_i(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n v_{ii}}} \left[ 1 + \frac{(x-(\mu)_i)^2}{n v_{ii}} \right]^{-\frac{n+1}{2}}$$

Denoting  $x_{ii}=x_2$ ,  $v_{ii}=v_2$ ,  $r$ -order marginal moment  $\mu_{ii\dots i}=\mu_r$ , the formulae (11)-(14) take the form:

$$\mu_2=x_2=k_2v_2, \mu_4=c_4 \cdot 3x_2^2, \mu_6=c_6 \cdot 15x_2^3, \mu_8=c_8 \cdot 105x_2^4.$$

Using these relations and, on the other hand, finding the integrals

$$\mu_r = \int_{-\infty}^{\infty} (x-\mu)^r f_i(x) dx, \quad r=2, \dots, 8$$

we get the constants  $c_r$  for the multivariate  $t$ -distribution:

$$c_2=k_2=\frac{n}{n-2}, \quad n>2,$$

$$c_4=\frac{n-2}{n-4}, \quad n>4,$$

$$c_6=\frac{(n-2)^2}{(n-4)(n-6)}, \quad n>6,$$

$$c_8=\frac{(n-2)^3}{(n-4)(n-6)(n-8)}, \quad n>8.$$

**Example 4.** Very often elliptical distributions are used to generalize the limiting distributions or Edgeworth expansions of some statistics obtained in the case of normal population. For this purpose the cumulants of observed statistic are necessary to be expressed for the elliptical population case. In the multivariate analysis the most usable statistics are various functions of sample covariance matrix. Hence with the help of the general expressions in [3] we give here the expressions of second and third cumulants of sample covariances  $s_{ij}$  when the sample is taken from the elliptical population:

$$E(s_{ij}-x_{ij})(s_{kl}-x_{kl})=n^{-1}[k_4x_{ij}x_{kl}+(k_4+1)(x_{ik}x_{jl}+x_{il}x_{jk})]+o(n^{-1}),$$

$$E(s_{ij}-x_{ij})(s_{kl}-x_{kl})(s_{ab}-x_{ab})=n^{-2}[k_6x_{ij}x_{kl}x_{ab}+ \\ + (k_6+3k_4+1)\sum x_{ik}x_{ja}x_{lb}+(k_6+2k_4)\sum x_{ij}x_{ka}x_{lb}]+o(n^{-2}).$$

**Example 5.** Fujikoshi [1] has given the Edgeworth expansion of the distribution function of  $\sqrt{n}(l_i-\lambda_i)/\sigma$ , where  $l_i$  and  $\lambda_i$  are the  $i$ -th largest latent roots of the sample and population covariance matrices. For elliptical population case the Fujikoshi's result has the following simple form:

$$P\{4n(1-\lambda_1)/\sigma^2 < y\} = \Phi(y) - n^{-1/2} \{a_1 \sigma^{-1} \Phi^{(1)}(y) + a_3 \sigma^{-3} \Phi^{(3)}(y)\} + \\ + n^{-1} \{ (b_2 + a_1^2/2) \sigma^{-2} \Phi^{(2)}(y) + (b_4 + a_1 a_3) \sigma^{-4} \Phi^{(4)}(y) + \\ + a_3^2 \sigma^{-6} \Phi^{(6)}(y)/2 \} + o(n^{-1}),$$

where  $\Phi^{(j)}(y)$  is the  $j$ -th derivative of the standard normal distribution function  $\Phi(y)$  and

$$a_1 = (k_4 + 1) \lambda_1 \sum_j \lambda_j (\lambda_1 - \lambda_j)^{-1}, \quad a_3 = (15k_6 + 36k_4 + 8) \lambda_1^3 / 6, \\ b_2 = -\lambda_1^2 [3k_4/2 - 3(k_6 - k_4^2 + k_4) \sum_j \lambda_j (\lambda_1 - \lambda_j)^{-1} + (k_4 + 1)^2 \sum_j \lambda_j^2 (\lambda_1 - \lambda_j)^{-2}], \\ b_4 = \lambda_1^4 (35k_6/8 + 15k_4 + 12k_4^2 + 18k_4 + 2), \quad \sigma^2 = (3k_4 + 2) \lambda_1^2,$$

with summation over  $j$ ,  $j \neq 1$ .

**Example 6.** The matrix form of the multivariate Edgeworth expansion, where the moments are organized into vectors, is given by Traat [5]. The vectors of the 4th and 6th order central moments of  $N_p(\mu, V)$  are:

$$\bar{\mu}_4 = J_1 (\text{vec} V \otimes \text{vec} V), \quad \bar{\mu}_6 = J_2 J_3 (\text{vec} V \otimes \text{vec} V \otimes \text{vec} V),$$

where  $J_1, J_2, J_3$  are the expressions from permutation matrices  $I_{p,q}$ :

$$J_1 = I_{p,4} + I_{p,p} \otimes I_p + I_{p,p} \otimes I_p, \quad J_3 = I_{p,6} + I_{p,p} \otimes I_{p,p} \otimes I_p + I_{p,p} \otimes I_{p,p} \otimes I_p, \\ J_2 = I_{p,6} + I_{p,p} \otimes I_{p,3} + I_{p,p} \otimes I_{p,3} + I_{p,p} \otimes I_{p,4} + I_{p,p} \otimes I_{p,5}.$$

From  $\bar{\mu}_4, \bar{\mu}_6$  follow the cumulants  $\bar{\kappa}_4^c, \bar{\kappa}_6^c$  and central moments  $\bar{\mu}_4^c, \bar{\mu}_6^c$  of elliptical distribution  $E_p(\mu, V)$  in the vector form:

$$\bar{\kappa}_4^c = k_4 \bar{\mu}_4, \quad \bar{\kappa}_6^c = k_6 \bar{\mu}_6, \quad \bar{\mu}_4^c = c_4 \bar{\mu}_4, \quad \bar{\mu}_6^c = c_6 \bar{\mu}_6$$

the elements of which are all 4th and 6th order mixed cumulants and moments respectively.

## References

1. Fujikoshi, Y. Asymptotic expansions for the distributions of the sample roots under nonnormality. *Biometrika*, 1980, 67, 45-51.
2. Kaplan, E. L. Tensor notation and the sampling cumulants of  $k$ -statistics. *Biometrika*, 1952, 39, 529-531.



3. Kendall, M.G., Stuart, A. Distribution theory. Moscow: Nauka Publishers, 1966 (Russ.).
4. Muirhead, R.J. Aspects of Multivariate Statistical Theory. New York: Wiley, 1982.
5. Traat, I. Matrix calculus for multivariate distributions. Acta Univ. Tartuensis, 1986, 733, 64-85.

МОМЕНТЫ И КУМУЛЯНТЫ МНОГОМЕРНОГО ЭЛЛИПТИЧЕСКОГО  
РАСПРЕДЕЛЕНИЯ С НЕКОТОРЫМИ ПРИМЕНЕНИЯМИ

И. Траат

Резюме

Приводятся выражения смешанных кумулянтов и центральных моментов до восьмого порядка многомерного эллиптического распределения. В примерах рассматриваются некоторые конкретные эллиптические распределения, кумулянты выборочных ковариаций и разложение распределения собственного значения выборочной ковариационной матрицы в случае эллиптически распределенной генеральной совокупности.

Received July 1989

ON THE EXISTENCE AND WEAK CONVERGENCE  
OF  $k$ -CENTRES IN BANACH SPACES

K. Pärna

Summary

Let  $P$  be a probability measure on a separable Banach space  $B$ . Any subset  $A = \{a_1, \dots, a_k\} \subset B$  is called the  $k$ -centre for  $P$  if it minimizes a criterion. The reflexivity of  $B$  is shown to be a sufficient condition for the existence of the  $k$ -centre for any  $P$  from a wide class of measures. Also, the weak convergence of  $k$ -centres, corresponding to certain weakly converging sequence of measures, has been studied.

1. Introduction

The problem of  $k$ -centres arises if there is a need for the discretization of a random variable. A well-known example is the quantization of continuous signal in order to transfer it through a discrete channel which is capable of admitting only  $k$  different values of the signal. Also, the optimal allocation of resources in order to meet a given distribution of needs may be regarded as a problem of that kind.

In recent years several papers have appeared where the problem is treated in terms of Banach spaces (see e.g. [1, 2]) or metric spaces [4, 5].

To be more precise, let  $B$  be a real separable Banach space,  $P$  a probability measure on  $B$  and  $k$  - a fixed positive integer number. Let us define a measure of goodness of the

approximation of the measure  $P$  by a finite set  $A = \{a_1, \dots, a_k\} \subset B$ :

$$W(A, P) = \int_B \min_{a_1 \in A} \varphi(\|x - a_1\|) P(dx). \quad (1)$$

We shall assume that the function  $\varphi$  satisfies the following requirements:

A1)  $\varphi: [0, \infty) \rightarrow [0, \infty)$ ,

A2)  $\varphi$  is continuous,

A3)  $\varphi$  is nondecreasing,

A4)  $\varphi(r) = 0$  iff  $r = 0$ ,

A5) there exists a constant  $\lambda$  such that  $\varphi(2r) \leq \lambda \cdot \varphi(r)$  for each  $r \geq 0$  ( $\Delta_2$ -property). Also, it is assumed that

$$\int \varphi(\|x\|) P(dx) < \infty, \quad (2)$$

which can be regarded as a restriction on the dispersion of the measure  $P$ . Further, let

$$W_k(P) = \inf \{W(A, P) : |A| = k\},$$

$$\mathcal{A}_k^*(P) = \{A : W(A, P) = W_k(P), |A| = k\}.$$

Any  $A \in \mathcal{A}_k^*(P)$  we shall call the  $k$ -centre of the measure  $P$ .

The first problem here is the problem of the existence of  $k$ -centres. We are revealing a class of spaces, as large as possible, where the existence of a  $k$ -centre can be proved for any measure with property (2). Show first that the class of metric spaces is too large to prove that. Indeed, consider the metric space  $T = \{x_1, x_2, x_3, y_1, y_2, \dots\}$  with distances  $d(x_1, x_j) = 1/(j-1)$ ,  $d(x_1, y_n) = \frac{1}{2} - 1/(n+10)$ ,  $d(x_2, y_n) = d(x_3, y_n) = \frac{1}{2} + 1/(n+10)$  ( $n \geq 1$ ), and  $d(y_n, y_m) = 0.8$  ( $n \neq m$ ). Put  $P(x_1) = P(x_2) = P(x_3) = 1/3$ . Then, defining  $W(a, P) = \int d(x, a) P(dx)$ ,  $a \in T$ , we have

$$W(x_1, P) = 2/3, \quad i=1, 2, 3, \dots$$

$$W(y_n, P) = 1/2 + 1/3(n+10), \quad n=1, 2, \dots$$

Hence  $W_1(P) = 1/2$ , but the infimum is not attainable in  $T$ , that is  $\mathcal{A}_1^*(P) = \emptyset$ .

It is easy to generalize the counterexample above to the case of  $k > 1$ . Simply copy the space  $T$   $k$  times defining the distance between points from different copies equal to 100 (say) and  $P$ -measure of each  $x$ -point equal to  $1/3k$ . Then  $\mathcal{A}_k^*(P) = \emptyset$ .

The counterexample given above shows that while study-

ing the existence problems of  $k$ -centres, it is reasonable to limit oneself with a more restrictive class of spaces, the Banach spaces. For  $k=1$ , the existence of 1-centres has been proved by Herrndorf [2] for a wide class of Banach spaces (so called IP-spaces). Still, it seems that his method of proof cannot be generalized to the case of arbitrary  $k$ . Cuesta and Matran [1] showed a way for proving the existence of  $k$ -centres for uniformly convex (u.c.) Banach spaces. Our aim here is to cover somewhat wider (as compared with u.c.) class of B-spaces, the class of reflexive B-spaces. Recall some examples of such spaces:  $R^m$ ,  $l_p$  ( $p>1$ ),  $L_p[0,1]$  ( $p>1$ ).

Besides the existence theorem we prove the convergence of  $k$ -centres of the measures  $P_n$ , assuming that  $\{P_n\}$  is a weakly convergent sequence,  $P_n \rightarrow P$ . This result generalizes a recent theorem from [1] (Th. 10), since 1) we cover more general spaces, 2) we do not assume the uniqueness of the  $k$ -centre of  $P$  and, 3) our sequence  $\{P_n\}$  is not necessarily empirical. Several results concerning the strong convergence of  $k$ -centres have been presented in [1,5,6].

The basic mathematical tool used in this paper is an existence theorem from the optimization theory. That theorem and some supporting lemmas will be presented in the next section. Such a preliminary work enables us to prove the main theorem (Theorem 2 in Section 3) very quickly.

## 2. Some preliminary results

This paper relies significantly on the following theorem from the theory of optimization (see e.g. [7], p.49).

We recall that a function  $J(u)$ , defined on a subset  $\mathcal{U}$  of Banach space  $B$  is called weakly lower semicontinuous, if for any sequence  $\{u_n\} \in \mathcal{U}$  converging weakly to some  $u \in \mathcal{U}$  (shortly,  $u_n \xrightarrow{w} u$ ) the inequality

$$\liminf_n J(u_n) \geq J(u)$$

holds.

We shall say that a sequence  $\{u_n\} \in B$  converges weakly to a subset  $\mathcal{U}_0 \subset B$  if each subsequence  $\{u_{n_k}\}$  contains a further subsequence  $\{u_{n_{k_l}}\}$  converging weakly to an element of  $\mathcal{U}_0$ .

Note that in the case when  $\mathcal{U}_0$  consists of a unique point  $u_0$  the definition above reduces to the weak convergence  $u_n \xrightarrow{w} u_0$  in common sense.

**Theorem 1.** Let  $\mathcal{U}$  be a weakly compact subset of Banach space  $B$  and let the function  $J(u)$  be defined, finite and weakly lower semicontinuous on  $\mathcal{U}$ . Then  $J_* = \inf_{u \in \mathcal{U}} J(u) > -\infty$ , the set  $\mathcal{U}_* = \{u \in \mathcal{U} : J(u) = J_*\}$  is nonempty, weakly compact and any minimizing <sup>1)</sup> sequence  $\{u_n\}$  converges weakly to  $\mathcal{U}_*$ .

In order to apply this theory, it is necessary to introduce a vector argument analogue for  $W(A, P)$ : for every  $\vec{A} = (a_1, \dots, a_k) \in B^k$  let

$$W(\vec{A}, P) = \int_B \min_{1 \leq i \leq k} \varphi(\|x - a_i\|) P(dx). \quad (3)$$

Evidently, this new function is invariant w.r.t. the permutations of the components of  $\vec{A}$ .

We shall show now that the function  $W(\vec{A}, P)$  satisfies all the assumptions of Th.1.

As a first step, we prove that  $W(\vec{A}, P)$  is finite on the whole  $B^k$ . Indeed, for any  $\vec{A} = (a_1, \dots, a_k) \in B^k$  we have

$$\begin{aligned} W(\vec{A}, P) &= \int_B \min_{1 \leq i \leq k} \varphi(\|x - a_i\|) P(dx) \\ &\leq \int_B \varphi(\|x - a_1\|) P(dx) \leq \int_B \varphi(\|x\| + \|a_1\|) P(dx) \\ &\leq \int_B \varphi(2\|a_1\|) P(dx) + \int_{\|x\| \geq \|a_1\|} \varphi(2\|x\|) P(dx) \\ &\leq \varphi(2\|a_1\|) + \lambda \cdot \int_B \varphi(\|x\|) P(dx) < \infty, \end{aligned}$$

due to (2) and A5). Secondly, we verify the following

**Lemma 1.** The function  $W(\vec{A}, P)$  is weakly lower semicontinuous on  $B^k$ .

**Proof.** We have to show that from  $\vec{A} \xrightarrow{w} \vec{A} \in B^k$  (which means that  $a_i^n \xrightarrow{w} a_i$  for  $i=1, \dots, k$ ) it follows that

$$\lim_n W(\vec{A}_n, P) \geq W(\vec{A}, P). \quad (4)$$

If  $a_i^n \xrightarrow{w} a_i$ , then for each  $x \in B$  we also have  $x - a_i^n \xrightarrow{w} x - a_i$ .  
<sup>1)</sup> A sequence  $\{u_n\}$  is called minimizing if  $\lim_n J(u_n) = J_*$ .

$\|x - a_1\|$ , and a known property of weak convergence entails

$$\liminf_n \|x - a_1^n\| \geq \|x - a_1\|, \quad i=1, \dots, k. \quad (5)$$

As the elements of any weakly converging sequence are uniformly bounded ([3], p.167), the limit on the left hand side of (5) is finite. Then, since  $\varphi$  is monotonic, it follows that

$$\min_i \varphi(\liminf_n \|x - a_i^n\|) \geq \min_i \varphi(\|x - a_i\|) \quad (6)$$

and then, by A2) and A3),

$$\liminf_n \min_i \varphi(\|x - a_i^n\|) \geq \min_i \varphi(\|x - a_i\|). \quad (7)$$

After the integration we have

$$\int_B \liminf_n \min_i \varphi(\|x - a_i^n\|) P(dx) \geq W(\tilde{A}, P). \quad (8)$$

As a final step, we apply Fatou's Lemma to the left hand side of (8). So we have

$$\liminf_n \int_B \min_i \varphi(\|x - a_i\|) P(dx) \geq W(\tilde{A}, P),$$

which is equivalent to (4). The proof is completed.

Now we show a suitable weakly compact subset  $\mathcal{U}$  of Theorem 1.

**Lemma 2.** Let  $P$  be not concentrated at any  $k-1$  (or fewer) points of  $B$ . Then, for any  $\varepsilon$ ,  $0 < \varepsilon < W_{k-1}(P) - W_k(P)$ , all the  $\varepsilon$ -optimal  $k$ -vectors  $\tilde{A}$  are contained in the set <sup>2)</sup>

$$\mathcal{U} = (B[x_0, M])^k \subset B^k. \quad (9)$$

If  $B$  is reflexive, then  $\mathcal{U}$  is weakly compact in  $B^k$ .

**Proof.** It is known (see [5], Lemma 2) that all the  $\varepsilon$ -optimal  $k$ -centres, i.e.  $k$ -sets  $A$  satisfying  $W(A, P) < W_k(P) + \varepsilon$ , are contained in some  $B[x_0, M]$  with  $M$  depending on  $\varepsilon$ , provided that  $0 < \varepsilon < W_{k-1}(P) - W_k(P)$ . (It will be proved in Appendix that the strict inequality  $W_k(P) < W_{k-1}(P)$  holds as long as  $P$  is not concentrated at any  $k-1$  or fewer points; hence we can choose a positive  $\varepsilon$ ). Clearly, all  $\varepsilon$ -optimal vectors  $\tilde{A}$  then belong to  $(B[x_0, M])^k$ , and any such an  $\tilde{A}$  contains exactly  $k$  different components.

<sup>2)</sup>  $B[x_0, M]$  is the closed ball with centre  $x_0$  and radius  $M$ .

Further, it is seen that the set  $\mathcal{U}$  in (9) is a closed, bounded, convex subset of the reflexive Banach space  $B^k$ . Since all such subsets are known to be weakly compact (see [7], p.51), Lemma 2 follows.

### 3. The main results

As we prefer to formulate the weak convergence of  $k$ -centres in terms of  $k$ -sets rather than  $k$ -dimensional vectors, it is necessary to say what the weak convergence of  $k$ -sets is.

**Definition 1.** We say that a sequence of  $k$ -sets  $A_n = \{a_1^n, a_2^n, \dots, a_k^n\}$  converges weakly to a  $k$ -set  $A = \{a_1, \dots, a_k\}$  (and we shall write  $A_n \xrightarrow{w} A$ ) if for some labeling  $a_{11}^n, \dots, a_{1k}^n$  of the points in  $A_n$ ,  $n = 1, 2, \dots$ , it happens that  $a_{1j}^n \xrightarrow{w} a_j$ , for  $j = 1, \dots, k$ .

Obviously, if the coordinate-wise convergence  $\tilde{A}_n \xrightarrow{w} \tilde{A}$  for certain  $k$ -vectors  $\tilde{A}_n, \tilde{A} \in B^k$  (provided they all have  $k$  distinct components) takes place, then also  $A_n \xrightarrow{w} A$  where  $A_n$  and  $A$  are  $k$ -sets consisting of the components of  $\tilde{A}_n$  and  $\tilde{A}$ , respectively.

**Definition 2.** We say that the sequence of  $k$ -sets  $\{A_n\}$  converges weakly to  $\mathcal{A}$ , a class of  $k$ -sets (and we write  $A_n \xrightarrow{w} \mathcal{A}$ ) if every subsequence  $\{A_{n_i}\}$  admits a further subsequence  $\{A_{n_{i_j}}\}$  converging weakly to a  $k$ -set from  $\mathcal{A}$ .

Once again, if  $\mathcal{A}$  consists of a single  $k$ -set  $A$  the latter definition is equivalent to the Definition 1 (see Lemma 1 in [1]).

We formulate the main result of this paper.

**Theorem 2.** Let  $B$  be a real separable reflexive Banach space. Then for each  $P$  satisfying (2)

- 1) the class  $\mathcal{A}_k^*(P)$  is not empty,
- 2) if, in addition,  $P$  is not concentrated at any  $l$ ,  $1 \leq l < k$ , points, then any minimizing sequence  $\{A_n\}$  converges weakly to  $\mathcal{A}_k^*(P)$ .

**Proof.** First suppose that  $P$  is concentrated on some subset  $A = \{a_1, \dots, a_l\} \subset B$ ,  $l < k$ . Then the assertion 1) holds, since  $\mathcal{A}_k^*(P)$  consists of all  $k$ -sets which include the points  $a_1, \dots, a_l$ . Obviously, in such a case  $W_k(P) = \dots = W_1(P) = 0$ .

If  $P$  is not concentrated on any  $l$ -set,  $1 < k$ , then Lemma 2 implies that the global infimum  $W_k(P)$  coincides with  $\inf\{W(\tilde{A}, P) : \tilde{A} \in \mathcal{U}\}$ , where  $\mathcal{U}$  is given by (9). According to Lemma 1 the (finite) function  $W(\tilde{A}, P)$  is weakly lower semi-continuous on  $\mathcal{U}$ . Both statements of the theorem now follow directly from Theorem 1. The proof is completed.

**Remark.** The reflexivity of  $B$  cannot be ignored, at least totally. It is seen from the following counterexample concerning the space  $c_0$ . Let  $k=1$ ,  $W(a, P) = \int \|x - a\|^2 P(dx)$ , and let  $P$  be concentrated at points  $2e_1, 2e_2, \dots$  with  $P(2e_1) = k^{-1}$ . Then  $W_1(P) = 1$ , a minimizing sequence is  $a_n = (\underbrace{1, 1, \dots, 1}_n, 0, 0, \dots)$ ,  $n = 1, 2, \dots$ , but the infimum is not attainable in  $c_0$ .

Observe that the second statement of Th.2 leaves the nature of minimizing sequences  $\{A_n\}$  open. We show now a special class of such sequences, defining  $A_n$  as a  $k$ -centre for the measure  $P_n$  from a weakly convergent sequence.

**Corollary 1.** Let  $B$  be a real separable reflexive Banach space and let  $P$  be a probability measure on  $B$ , not concentrated at any  $l$  ( $1 < k$ ) points. If the sequence  $\{P_n\}$  satisfies

c1)  $P_n \rightarrow P$ ,

c2) the function  $\varphi(\|x\|)$  is uniformly integrable w.r.t.

$\{P_n\}$ ,

then for any  $A_n \in \mathcal{A}_k^*(P_n)$  we have

$$A_n \xrightarrow{W} \mathcal{A}_k^*(P), \quad n \rightarrow \infty. \quad (10)$$

**Proof.** Due to Lemma 3 (see Appendix) the measure  $P$  verifies

$$W_1(P) > \dots > W_k(P). \quad (11)$$

Then, according to Corollary 1 from [5], each sequence  $A_n \in \mathcal{A}_k^*(P_n)$ ,  $n = 1, 2, \dots$ , is minimizing:

$$\lim_n W(A_n, P) = W_k(P).$$

Hence Theorem 2 applies and we obtain (10). The proof is completed.

It is of worthy to point out that assumptions c1), c2) have been shown to be weak enough to include the important case of empirical measures  $P_n$ , correspond-



ing to the measure  $P$  (see Section 3 in [5]). Some other interesting particular cases of  $\{P_n\}$  also can be given.

#### 4. Appendix

In this section we give sufficient conditions that ensure inequalities (11) to hold. These conditions are significantly milder as compared with those in our previous result (see [4], Lemma 2): no more we need  $\varphi$  being strictly monotonic, nor the existence of a 1-centre for  $l = 1, 2, \dots, k-1$  is assumed. (In fact, all this enables us to reduce assumptions to our results given in [4], [5]).

The following lemma considers the spaces more general than Banach ones. Let  $(T, d)$  be a separable metric space. Define

$$W(A, P) = \int_T \min_{a_1 \in A} \varphi(d(x, a_1)) P(dx), \quad A = \{a_1, \dots, a_k\} \subset T,$$

- a generalization of (1). The Hausdorff distance between two subsets of  $T$  is given by

$$h(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} d(a, b), \sup_{b \in B} \inf_{a \in A} d(a, b) \right\}.$$

**Lemma 3.** Let  $(T, d)$  be a separable metric space,  $P$  a probability measure on  $T$ , not concentrated at any  $k-1$  or fewer points. Assume that  $\int \varphi(d(x, y_0)) P(dx) < \infty$  for some  $y_0 \in T$ . Then

$$W_1(P) > W_2(P) > \dots > W_k(P).$$

**Proof.** We prove this lemma for the case of  $k=3$  only. Other values of  $k$  can be treated similarly.

Consider a sequence  $\{A_n\}$ ,  $A_n = \{a_1^n, a_2^n\}$  being an  $(1/n)$ -optimal 2-set for the measure  $P$ , i.e.

$$W_2(P) \leq W(A_n, P) < W_2(P) + \frac{1}{n}. \quad (12)$$

There are two possibilities:

- a)  $\{A_n\}$  converges in Hausdorff metrics to some 2-set,
- b)  $\{A_n\}$  does not converge in H.m. to any 2-set.

**The case a).** If, for some  $A_0 = \{a_1^0, a_2^0\}$ ,  $h(A_n, A_0) \rightarrow 0$ , then, due to the continuity of  $W(\cdot, P)$  (see [5], Lemma A2), we also have  $W(A_n, P) \rightarrow W(A_0, P)$ . On the other hand, (12) implies  $W(A_n, P) \rightarrow W_2(P)$ . Consequently,

$$W(A_0, P) = W_2(P), \quad (13)$$

which means that  $A_0 \in \mathcal{A}_2^*(P)$ .

Show now that there exist a point  $b \in T$ ,  $b \in A_0$ , and an  $s$ ,  $0 < s < (1/3) \cdot \min\{d(a_1^0, b), d(a_2^0, b)\}$ , such that the open sphere  $B(b, s)$  has positive  $P$ -measure. The idea is that if this  $b$  will be added to  $A_0$  one gets a triple  $A_0^+$  which is strictly better than  $A_0$  itself, in sense of  $W(\cdot, P)$ . Then the needed result follows immediately.

Indeed, since  $P$  is assumed to be not concentrated at any 2 points, we have  $P(A_0) < 1$ . Then there exists an  $r > 0$  such that the closed set  $A_0^r = \{x \in T: \min\{d(x, a_1^0), d(x, a_2^0)\} \leq r\}$  also satisfies  $P(A_0^r) < 1$ . (If for each  $r > 0$   $P(A_0^r) = 1$ , then by the continuity of  $P$  it follows that  $P(A_0) = 1$  - the contradiction). Hence, the open set  $T \setminus A_0^r$  is  $P$ -positive. As  $T$  is separable, there exists, for any  $s > 0$ , a countable system of spheres with centres in  $T \setminus A_0^r$  and of radius  $s$ , say the system  $\mathcal{S}_s$ , covering  $T \setminus A_0^r$ . Choose an  $s$  so that  $0 < s < r/3$  and  $\varphi(2s) > \varphi(s)$ . (It is an easy exercise to show, using A1) to A5), that the latter inequality holds for arbitrary small  $s$ ). Now from  $P(T \setminus A_0^r) > 0$  it follows that at least one sphere from  $\mathcal{S}_s$ , say  $B(b, s)$ , also satisfies  $P(B(b, s)) > 0$ .

Show the set  $A_0^+ = \{a_1^0, a_2^0, b\}$  is strictly 'better' than  $A_0$ . Let

$$S_b = \{x \in T: \varphi(d(x, b)) \leq \varphi(d(x, a_1^0)), \quad i=1,2\}.$$

Observing that  $S_b$  contains the ball  $B(b, s)$  and that  $d(x, a_1^0) \geq 2s$  for all  $x \in B(b, s)$ , we have

$$\begin{aligned} W(A_0, P) - W(A_0^+, P) &= \int_{S_b} [\min_{i=1,2} \varphi(d(x, a_i^0)) - \varphi(d(x, b))] P(dx) \\ &\geq \int_{B(b, s)} [\varphi(2s) - \varphi(s)] P(dx) \\ &= [\varphi(2s) - \varphi(s)] P(B(b, s)) > 0. \end{aligned} \quad (14)$$

Now, combining (13) and (14) with

$$W_3(P) \leq W(A_0^+, P),$$

we have

$$W_3(P) < W_2(P),$$

which completes the proof in the case a).

The case b). Assume that  $\{A_n\}$  does not converge to any 2-set in Hausdorff metrics. Then for each  $D = \{b, c\} \subset T$  there exists an  $r_D > 0$  such that at least one ball,  $B(b, r_D)$  or  $B(c, r_D)$ , does not contain any point from  $A_n$  for infinitely many values of  $n$ . Say this ball is  $B(b, r_D)$ . Since for each  $b$  the radius  $r_D$  depends also on its pair-mate  $c$ , we fix an arbitrary value of these  $r_D$ 's and denote it by  $r_b, r_b > 0$ .

Thus we have

$$B(b, r_b) \cap A_{n'} = \emptyset \quad (15)$$

along some subsequence  $\{n'\}$ .

It turns out that the points  $b$  with property (15) cover all  $T$  except, perhaps, a single point  $c_0$ . Indeed, if there were two such points, say  $c_0$  and  $c_1$ , then after putting  $D = \{c_0, c_1\}$  one immediately reaches the contradiction.

Further discussion exploits the ideas already used in the case a). First define a covering for  $T \setminus \{c_0\}$ , the system of open balls

$$\mathcal{B} = \{B(b, s_b) : b \in T \setminus \{c_0\}, 0 < s_b \leq r_b/3, \varphi(2s_b) > \varphi(s_b)\},$$

with  $r_b$  satisfying (15). Again, since  $T$  is separable,  $\mathcal{B}$  contains a countable subsystem also covering  $T \setminus \{c_0\}$ . But  $T \setminus \{c_0\}$  is  $P$ -positive and hence at least one sphere from that countable system, the sphere  $B(b, s)$ , is also  $P$ -positive,  $P(B(b, s)) > 0$ . Put  $A_n^+ = \{a_1^{n'}, a_2^{n'}, b\}$ . By the same way as in (14) we obtain that

$$W(A_n^+, P) - W(A_n^+, P) \geq [\varphi(2s) - \varphi(s)]P(B(b/s)) \equiv \alpha > 0, \quad (16)$$

for all  $n' \rightarrow \infty$ . Now it suffices to choose an  $n' > 1/\alpha$  and recall that  $A_n$  is  $(1/n')$ -optimal to write

$$W_3(P) \leq W(A_n^+, P) \leq W(A_n, P) - \alpha < W_2(P) + \frac{1}{n} - \alpha < W_2(P).$$

Thus Lemma 3 is proved.

## References

1. C u e s t a, J. A., M a t r a n, C. The strong law of large numbers for k-means and best possible nets of Banach valued random variables. Probab. Th. Rel. Fields, 1988, 78, 523-534.
2. H e r r n d o r f, N. Approximation of vector-valued random variables by constants. J. Approx. Th., 1983, 37, 175-181.
3. L y u s t e r n i k, L. A., S o b o l e v, V. I. A Short Course in Functional Analysis. Moscow, 1982. (Russ.)
4. P ä r n a, K. Strong consistency of k-means clustering criterion in separable metric spaces. Tartu Riikl. Ülik. Toimetised, 1986, 733, 86-96.
5. P ä r n a, K. On the stability of k-means clustering in metric spaces. Tartu Riikl. Ülik. Toimetised, 1988, 798, 19-36.
6. P o l l a r d, D. Strong consistency of k-means clustering. Ann. Statist., 1981, 2, 135-140.
7. V a s s i l j e v, F. P. Methods of Solution For Extremal Problems. Moscow, Nauka Publishers, 1981. (Russ.)

## О СУЩЕСТВОВАНИИ И СЛАБОЙ СХОДИМОСТИ К-ЦЕНТРОВ В БАНАХОВЫХ ПРОСТРАНСТВАХ

К.Пярна

### Р е з ю м е

Рассматривается задача дискретной (конечной) аппроксимации вероятностных распределений, заданных на банаховых пространствах. Пусть  $P$  является вероятностной мерой на сепарабельном банаховом пространстве  $B$ . Любое подмножество  $A = \{a_1, \dots, a_k\} \subset B$  называется  $K$ -центром меры  $P$ , если оно минимизирует следующий критерий средней ошибки аппроксимации:

$$W(A, P) = \int_B \min_{1 \leq i \leq k} \varphi(\|x - a_i\|) P(dx) \rightarrow \min_{|A|=k} . \quad (1)$$

Относительно функции  $\varphi$  предполагается, что она  $1^\circ$  непрерывна,  $2^\circ$  не убывает,  $3^\circ \varphi: [0, \infty) \rightarrow [0, \infty)$ ,  $4^\circ \varphi(r) = 0 \Leftrightarrow r = 0$ ,  $5^\circ$  найдется  $\lambda > 0$  такое, что  $\varphi(2r) \leq \lambda \cdot \varphi(r)$ ,  $r > 0$ .

Нам представляют интерес вопросы существования  $k$ -центров, а также вопросы сходимости последовательности  $k$ -центров, соответствующих мерам из слабо сходящейся последовательности.

Теорема 2. Пусть  $B$  - вещественное сепарабельное рефлексивное банахово пространство. Тогда для каждой меры  $P$ , заданной на  $B$  и удовлетворяющей условию  $\int \varphi(\|x\|) P(dx) < \infty$ , существует хотя бы один  $k$ -центр ( $k = 1, 2, \dots$ ).

Следствие 1. Пусть  $B$  - вещественное сепарабельное рефлексивное банахово пространство и пусть мера  $P$ , заданная на  $B$ , не сконцентрируется на каких-либо  $1, 1 < k$ , точках  $B$ . Если последовательность мер  $\{P_n\}$  удовлетворяет условиям 1)  $P_n \rightarrow P$  (слабо), 2)  $\varphi(\|x\|)$  интегрируема равномерно по  $\{P_n\}$ , то любая последовательность  $\{A_n\}$   $k$ -центров для мер  $P_n$  сходится слабо к множеству всевозможных  $k$ -центров меры  $P$ :

$$A_n \xrightarrow{w} \mathcal{A}_k^*(P), \quad n \rightarrow \infty.$$

Received December 1989

Exact samples for testing ANOVA procedures.

M.Vähi

Today there exist many statistical packages the users of which want to be sure that the programs work accurately and correctly. A convenient possibility to check up statistical programs is given by the method of the "exact sample" described in papers [1], [2] and [3].

The exact sample is an array of data with a special structure the dimensions and identifying parameters of which may be chosen freely. The values of necessary statistics are calculated not by the usual algorithms but analytically by the help of the identifying parameters.

Lower the rules for construction of exact samples for checking algorithms of variance analysis will be constructed. Only the balanced cross-models will be considered.

We shall construct the exact sample step by step.

1. One-way analysis of variance

At first we shall learn the simplest model - the model with one factor. We describe the construction of the exact sample and give the formulas for the calculation of the necessary statistics.

In that case the data  $y_{ij}$  (the  $j$ -th measurement on  $i$ -th level of factor) are presented by the model

$$y_{ij} = \mu + \alpha_i + e_{ij},$$

where  $\mu$  - the general mean,

$\alpha_i$  - the change caused by  $i$ -th level of factor,

$e_{ij}$  - the random error of measurement.

The factor has  $k$  levels ( $i=1,2,\dots,k$ ) and on each level  $n$  measurements ( $j=1,2,\dots,n$ ) are carried out. Hence we have  $nk$  measurements in all. In an essential way the data are divided into  $k$  groups - in the same group there are data that which are measured on the same level of factor.

The necessary statistics for one-way analysis are the following [4]:

a) the means

$$\bar{y}_{..} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^n y_{ij}.$$

$$\bar{y}_{1.} = \frac{1}{n} \sum_{j=1}^n y_{1j};$$

b) the variance components

$$S_1^2 = \sum_{i=1}^k n(\bar{y}_{i.} - \bar{y}_{...})^2,$$

$$S^2 = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2;$$

c) the F-ratio

$$F = \frac{S_1^2(N-k)}{S^2(k-1)}.$$

For the exact sample we shall presume that  $n$  is an odd number. The structure of the exact sample only "imitates" the described model. We introduce the "random error" with the basic structure of the group and with a "variance parameter". For the basic structure of the group we select the following sequence with  $n$  elements:

$$-\frac{n-1}{2}, -\frac{n-1}{2} + 1, \dots, -\frac{n-1}{2} + (n-1).$$

To get an exact sample, we determine the number of measurements in group ( $n$ ) and the number of the factor's levels ( $k$ ). We form the first group from the basic structure by multiplying all its elements by the arbitrarily chosen variance parameter  $u$  ( $u > 0$ ). We get the second group by adding to all elements of the first group the freely chosen parameter  $d$ . By this parameter we imitate the change of the group mean. We get all following groups in the same way: by adding parameter  $d$  to all elements of the preceding group.

The elements of such exact sample are determined by the formula:

$$y_{ij} = -\frac{n-1}{2} u + u(j-1) + d(i-1), \quad (1)$$

where  $i=1, \dots, k$ ;  $j=1, \dots, n$ .

**Example 1.** Let the factor have 3 levels ( $k=3$ ) and in the group there are 5 elements ( $n=5$ ). The variance parameter  $u=2$  and the parameter of the mean's change  $d=4$ . Then the basic structure has the form:

-2, -1, 0, 1, 2 ;

the first group is

-4, -2, 0, 2, 4 ;

and the exact sample is following

-4, -2, 0, 2, 4, 0, 2, 4, 6, 8, 4, 6, 8, 10, 12 .

It is easy to see that in such case the necessary statis-

tics can be calculated by the parameters  $k$ ,  $n$ ,  $u$  and  $d$  in the following way:

a) means

$$\bar{y}_{1..} = (1-1)d ,$$

$$\bar{y}_{..} = \frac{1}{k} \sum_{i=1}^k \bar{y}_{i.} = \frac{d(k-1)}{2} ;$$

b) the variance components

$$S_1^2 = \frac{d^2 N(k^2-1)}{12} ,$$

$$S^2 = \frac{u^2 N(n^2-1)}{12} ;$$

c) the F-ratio

$$F = \frac{d^2 k(k+1)}{u^2 (n+1)} .$$

Example 2. For the data given in Example 1 the statistics have the following values

$$\bar{y}_{1.} = 0, \bar{y}_{2.} = 4, \bar{y}_{3.} = 8, \bar{y}_{..} = 4,$$

$$S_1^2 = 160, S^2 = 120, F = 8 .$$

## 2. Two-way analysis of variance

In case of the two-way analysis we use the word "group" for data corresponding to the certain combinations of levels of factors. So  $y_{hij}$  is the  $j$ -th measurement on the  $h$ -th level of the first and on the  $i$ -th level of the second factor. Let the first factor have  $k$  levels, the second factor have  $m$  levels and in the group be  $n$  measurements. The necessary statistics for the two-way analysis are the following:

a) means

$$\bar{y}_{...} = \frac{1}{N} \sum_{h=1}^k \sum_{i=1}^m \sum_{j=1}^n y_{hij} ,$$

$$\bar{y}_{h..} = \frac{1}{kn} \sum_{i=1}^m \sum_{j=1}^n y_{hij} ,$$

$$\bar{y}_{.i.} = \frac{1}{mn} \sum_{h=1}^k \sum_{j=1}^n y_{hij} ,$$

$$\bar{y}_{hi.} = \frac{1}{n} \sum_{j=1}^n y_{hij} ;$$

b) the variance components

$$S_1^2 = mn \sum_{i=1}^m (\bar{y}_{.i.} - \bar{y}_{...})^2 ,$$

$$S_2^2 = kn \sum_{h=1}^k (\bar{y}_{h..} - \bar{y}_{...})^2 ,$$



$$S_{1,2}^2 = n \sum_{h=1}^m \sum_{i=1}^k (\bar{y}_{hi.} - \bar{y}_{h..} - \bar{y}_{.i.} + \bar{y}_{...})^2,$$

$$S^2 = \sum_{h=1}^m \sum_{i=1}^k \sum_{j=1}^n (y_{hij} - \bar{y}_{hi.})^2;$$

c) the F-ratios

$$F_1 = \frac{S_{mk(n-1)}^2}{S^2(k-1)},$$

$$F_2 = \frac{S_{mk(n-1)}^2}{S^2(m-1)},$$

$$F_{1,2} = \frac{S_{1,2}^2 mk(n-1)}{S^2(m-1)(n-1)}.$$

To get an exact sample, we determine the number of measurements in group  $n$  (let it be an odd number) and the number of levels for each factor  $k$  and  $m$ . Then we choose the parameters  $u$  and  $d$  (the meaning of these parameters is the same as in the preceding case) and generate by formula (1) data for each group where the second factor has the first level. Then we choose the parameter  $c$ , by this parameter we imitate the change of the group mean caused by the second factor. The elements for all the groups where the second factor has  $h$ -th level we get by adding  $(h-1)c$  to elements of the first-level groups.

The elements of such an exact sample are determined by the formula

$$y_{hij} = -\frac{n-1}{2} u + u(j-1) + d(i-1) + c(h-1). \quad (2)$$

**Example 3.** Let us choose the following values of the parameters:  $k=3$ ,  $m=2$ ,  $n=3$ ;  $u=1$ ,  $d=4$  and  $c=5$ . Then we get the following exact sample

the 2-nd factor level	the 1-st factor level		
	1	2	3
1	-1, 0, 1	3, 4, 5	7, 8, 9
2	4, 5, 6	8, 9, 10	12, 13, 14

It is easy to see that for such an exact sample the variance components are calculated by the formulas

$$S_1^2 = \frac{d^2 N(k^2-1)}{12},$$

$$S_2^2 = \frac{c^2 N(m^2-1)}{12},$$

$$S_{1,2}^2 = 0,$$

$$S^2 = \frac{u^2 N(n^2 - 1)}{12}$$

For such an exact sample all the interactions of factors are equal to zero. For generating some interactions we add to the elements of first and last group a parameter b. The first is the group where all the factors have the first level, the last is the group, where all the factors have the highest level.

**Example 4.** Consider the data from Example 3. Let  $b=6$ . Then we get the following exact sample:

the 2-nd factor level	the 1-st factor level		
	1	2	3
1	5, 6, 7	3, 4, 5	7, 8, 9
2	4, 5, 6	8, 9, 10	18, 19, 20

The elements of the exact sample are determined by the formula

$$y_{hij} = -\frac{n-1}{2} u + u(j-1) + d(i-1) + c(h-1) + b(\delta_{h1} \delta_{i1} + \delta_{hm} \delta_{ik}), \quad (3)$$

where

$$\delta_{ij} = \begin{cases} 1, & i=j, \\ 0, & i \neq j. \end{cases}$$

It is easy to see that for such an exact sample the necessary statistics may be calculated by formulas

a) means

$$\bar{y}_{...} = \frac{1}{2} (d(k-1) + c(m-1)) + \frac{2b}{mk},$$

$$\bar{y}_{h..} = \frac{d(k-1)}{2} + (h-1)c + \frac{b}{k} (\delta_{h1} + \delta_{hm}),$$

$$\bar{y}_{.i.} = d(i-1) + \frac{c(m-1)}{2} + \frac{b}{m} (\delta_{i1} + \delta_{ik}),$$

$$\bar{y}_{hi.} = (i-1)d + (h-1)c + b(\delta_{h1} \delta_{i1} + \delta_{hm} \delta_{ik});$$

b) the variance components

$$S_1^2 = \frac{d^2 N(k^2 - 1)}{12} + \frac{2b^2 n(k-2)}{km},$$

$$S_2^2 = \frac{c^2 N(m^2 - 1)}{12} + \frac{2b^2 n(m-2)}{km},$$

$$S_{1,2}^2 = \frac{2b^2 n^2 ((m-1)(k-1)+1)}{N},$$

$$S^2 = \frac{u^2 N(n^2 - 1)}{12};$$

c) the F-ratios

$$F_1 = \frac{d^2 mk(k+1)}{u^2(n+1)} + \frac{24b^2(k-2)}{u^2 mk(k-1)(n+1)}$$

$$F_z = \frac{c^2 mk(m+1)}{u^2(n+1)} + \frac{24b^2(m-2)}{u^2 mk(m-1)(n+1)}$$

$$F_{1,2} = \frac{24b^2 n((m-1)(k-1)+1)}{u^2 N(m-1)(k-1)(n+1)}$$

Example 5. In the case of the data given in Example 4 the necessary statistics have the following values

a) means

$$\bar{y}_{...} = 8,5 ; \quad \bar{y}_{1..} = 6 ; \quad \bar{y}_{2..} = 11 ;$$

$$\bar{y}_{.1.} = 5,5 ; \quad \bar{y}_{.2.} = 6,5 ; \quad \bar{y}_{.3.} = 13,5 ;$$

$$\bar{y}_{11.} = 6 ; \quad \bar{y}_{12.} = 4 ; \quad \bar{y}_{13.} = 8 ;$$

$$\bar{y}_{21.} = 5 ; \quad \bar{y}_{22.} = 9 ; \quad \bar{y}_{23.} = 13 ;$$

b) the variance components

$$S_1^2 = 228 ; \quad S_2^2 = 112,5 ;$$

$$S_{1,2}^2 = 108 ; \quad S^2 = 12 ;$$

c) the F-ratios

$$F_1 = 114 ; \quad F_2 = 112,5 ; \quad F_{1,2} = 54 .$$

### 3. Multi-way analysis of variance

Usually in practice one limits himself with the three-way analysis. But as in principle the number of factors is not limited, we construct the exact sample for  $p$  factors.

In case of  $p$ -way analysis the group will mean the data corresponding to a certain combination of levels of all the factors. Let the number of measurements in group be  $n$ , the number of levels of factors be  $k_1, \dots, k_p$ . As in preceding cases,  $u$  is a parameter of variance and  $d_1, \dots, d_p$  are the parameters of the influence of the factors.

The exact sample is generated step by step. At the first step we generate the exact sample for two factors by formula (2). On the next step we include the third factor: the already generated data corresponding to the first level of the factor. We get the groups corresponding to the  $i$ -th level by adding  $(i-1)d$  to that array. In the same way we continue until all the factors are included. At last we add the parameter  $b$  to all the elements of the first and the last group.

The elements of such an exact sample are determined by the formula

$$y_{i_1 \dots i_p j} = -\frac{n-1}{2} u + u(j-1) + \sum_{h=1}^p (i_h - 1) d_h + b \left[ \prod_{h=1}^p \delta_{i_h 1} + \prod_{h=1}^p \delta_{i_h k_h} \right], \quad (4)$$

where  $j=1, \dots, n$ ;  $i_1=1, \dots, k_1$ ;  $i_2=1, \dots, k_2$ ; ...;  $i_p=1, \dots, k_p$ .

For such an exact sample the necessary statistics can be calculated by formulas:

a) means

$$\bar{y} = \frac{1}{2} \left[ \sum_{h=1}^p (k_h - 1) d_h \right] + \frac{2bn}{N}$$

The group mean is calculated by the formula

$$\bar{y}_{i_g} = \frac{1}{2} \left[ \sum_{h=1}^p (k_h - 1) d_h \right] + (i_g - 1) d_g + \frac{bnk}{N} \left[ \delta_{i_g 1} + \delta_{i_g k_g} \right],$$

where  $g=1, \dots, p$ ;

and the mean on the fixed level some factors is calculated by the formula

$$\begin{aligned} \bar{y}_{i_{f_1} \dots i_{f_v}} &= \frac{1}{2} \left[ \sum_{h=1}^p (k_h - 1) d_h \right] + \sum_{j=1}^v (i_{f_j} - 1) d_{f_j} + \\ &+ \frac{bn}{N} \sum_{j=1}^v k_{f_j} \left[ \sum_{i=1}^{i_{f_j}} \delta_{i_{f_j} i} + \sum_{i=k_{f_j}}^{i_{f_j} k_{f_j}} \delta_{i_{f_j} i} \right], \end{aligned}$$

where  $f_1, \dots, f_v$  are the fixed levels.

b) the variance components

$$S_h^2 = \frac{1}{12} (Nd_h^2 (k_h^2 - 1)) + \frac{2b^2 n^2 (k_h - 2)}{N},$$

$$S_{f_1 \dots f_v}^2 = \frac{2b^2 n^2}{N} \left[ \sum_{j=1}^v (k_{f_j} - 1) + (-1)^v \right],$$

$$S^2 = \frac{1}{12} u^2 N (n^2 - 1);$$

c) the F-ratios

$$F_h = \frac{Nd_h^2 (k_h + 1)}{u^2 n (n+1)} + \frac{24b^2 n (k_h - 2)}{u^2 N (n+1) (k_h - 1)},$$

$$F_{f_1 \dots f_v} = \frac{24b^2 n \left[ \sum_{j=1}^v (k_{f_j} - 1) + (-1)^v \right]}{u^2 N (n+1) \sum_{j=1}^v (k_{f_j} - 1)}.$$

**Example 6.** Let us choose the following values for the parameters:  $k_1=3$ ,  $k_2=k_3=2$ ,  $d_1=1$ ,  $d_2=4$ ,  $d_3=5$ ,  $n=3$ ,  $u=1$ ,  $b=2$ . Then we get the following exact sample

the 3-rd factor level	the 2-nd factor level	the 1-st factor level		
		1	2	3
1	1	1, 2, 3	0, 1, 2	1, 2, 3
	2	3, 4, 5	4, 5, 6	5, 6, 7
2	1	4, 5, 6	5, 6, 7	6, 7, 8
	2	8, 9, 10	9, 10, 11	12, 13, 14

The means have the values:

$$\begin{aligned}
\bar{y}_{...} &= 5,833 ; & \bar{y}_{1..} &= 3,333 ; & \bar{y}_{2..} &= 8,333 ; \\
\bar{y}_{.1.} &= 3,833 ; & \bar{y}_{.2.} &= 7,833 ; & \bar{y}_{.1.1} &= 5 ; \\
\bar{y}_{.1.2} &= 5,5 ; & \bar{y}_{.1.3} &= 7 ; & \bar{y}_{11.} &= 1,667 ; \\
\bar{y}_{12.} &= 5 ; & \bar{y}_{21.} &= 6 ; & \bar{y}_{22.} &= 10,667 ; \\
\bar{y}_{2.1} &= 3 ; & \bar{y}_{1.2} &= 3 ; & \bar{y}_{1.3} &= 4 ; \\
\bar{y}_{2.1} &= 7 ; & \bar{y}_{2.2} &= 8 ; & \bar{y}_{2.3} &= 10 ; \\
\bar{y}_{.11} &= 3,5 ; & \bar{y}_{.12} &= 3,5 ; & \bar{y}_{.13} &= 4,5 ; \\
\bar{y}_{.21} &= 6,5 ; & \bar{y}_{.22} &= 7,5 ; & \bar{y}_{.23} &= 9,5 ; \\
\bar{y}_{111} &= 2 ; & \bar{y}_{112} &= 1 ; & \bar{y}_{113} &= 2 ; \\
\bar{y}_{121} &= 4 ; & \bar{y}_{122} &= 5 ; & \bar{y}_{123} &= 6 ; \\
\bar{y}_{211} &= 5 ; & \bar{y}_{212} &= 6 ; & \bar{y}_{213} &= 7 ; \\
\bar{y}_{221} &= 9 ; & \bar{y}_{222} &= 10 ; & \bar{y}_{223} &= 13 .
\end{aligned}$$

The variance components:

$$\begin{aligned}
S_1^2 &= 225 ; & S_2^2 &= 144 ; & S_3^2 &= 26 ; \\
S_{1.2}^2 &= 4 ; & S_{1.3}^2 &= 6 ; & S_{2.3}^2 &= 6 ; \\
S_{1.2.3}^2 &= 2 ; & S^2 &= 24 .
\end{aligned}$$

The F-ratios:

$$\begin{aligned}
F_1 &= 225 ; & F_2 &= 144 ; & F_3 &= 13 ; \\
F_{1.2} &= 4 ; & F_{1.3} &= 3 ; & F_{2.3} &= 3 ; & F_{1.2.3} &= 1 .
\end{aligned}$$

### References

1. Tiit, E.-M. Exact samples of multivariate distributions and their exploitation in statistical algorithm s testing. Tartu Riikliku Ülikooli Toimetised, 1986, 733, p. 40-62.
2. Tiit, E.-M. Testing algorithms and programs of multivariate statistical procedures - necessary assumption of

- building expert systems. International Journal of Pattern Recognition and Artificial Intelligence. 1989, 3, 121-133.
3. Парринг, А.-М., Тийт, Э.-М. Применение точной выборки для тестирования алгоритмов дискриминантного анализа. - III Всесоюзная школа-семинар "Программно-алгоритмическое обеспечение прикладного многомерного статистического анализа". Тезисы докладов. Ч. II. М., 1987, 234-237.
4. Шеффе, П. Дисперсионный анализ. М., 1980.

### КОНСТРУИРОВАНИЕ ТОЧНЫХ ВЫБОРОК ДЛЯ ТЕСТИРОВАНИЯ ПРОЦЕДУР ДИСПЕРСИОННОГО АНАЛИЗА

М.Вяхи

#### Р е з ю м е

Все более широкое применение вычислительных машин и пакетов статистической обработки данных создает необходимость контроля корректности и точности применяемых программ. Одним методом тестирования статистических процедур является метод точных выборок.

В настоящей статье приведена общая схема построения точной выборки подходящей для контроля процедур дисперсионного анализа и формула вычисления статистик, исходя из параметров этой выборки.

При конструировании подходящей выборки используют следующие произвольно задаваемые параметры: количество факторов; количество измерений в группе; количество уровней факторов; параметр внутригрупповой расхожимости; параметры влияния факторов; параметр совлиания. Общий член конструкции вычисляется по предписанию (4).

Received June 1989.

## C O N T E N T S

G. M i s e v i č i u s. On large deviations for a sum of type $\sum f(T^j t)$ .....	3
I. T r a a t. Moments and cumulants of multivariate elliptical distribution with some applications ..	11
K. P á r n a. On the existence and weak convergence of k-centres in Banach spaces .....	17
M. V a h i. Exact samples for testing ANOVA procedures ..	29

## Р Е З Ю М Е

Г. М и с я в и ч ю с. О больших отклонениях для сумм типа $\sum f(T^j t)$ .....	10
И. Т р а а т. Моменты и кумулянты многомерного эллип- тического распределения с некоторыми примени- ями .....	16
К. П я р н а. О существовании слабой сходимости k-центров в банаховых пространствах .....	28
М. В я х и. Конструирование точных выборок для тес- тирования процедур дисперсионного анализа .....	37

Ученые записки Тартуского университета.  
Выпуск 893.  
**СТОХАСТИЧЕСКИЕ МОДЕЛИ.**  
Труды по математике и механике.  
На английском языке.  
Резюме на русском языке.  
Тартуский университет.  
ЭР, 202400, г.Тарту, ул.Лийкооли, 18.  
Vastutav toimetaja K. Pärna.  
Paljundamisele antud 16.05.1990.  
Formaat 60x90/16.  
Kirjutusepaaber.  
Masinakiri. Rotaprint.  
Arvestuspoognaid 2,28. Trükipoognaid 2,5.  
Trükiarv 350.  
Tell. nr. 343.  
Hind 60 kop.  
TÜ trükikoda. ЭВ, 202400 Tartu, Tiigi t. 78.