

Poččet pravděpodobnosti a aplikovaná matematika

Časopis pro pěstování matematiky a fyziky, Vol. 74 (1949), No. 4, 281--310

Persistent URL: <http://dml.cz/dmlcz/122481>

## Terms of use:

© Union of Czech Mathematicians and Physicists, 1949

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

## 5. A sekce:

### Počet pravděpodobnosti a aplikovaná matematika.

#### THE APPROXIMATION OF THE LIMIT PROBABILITY LAW.

HALINA MILICER-GRUŻEWSKA, Warszawa.

The conditions under which the asymptotic expansion theorem is fulfilled,<sup>1)</sup> if the random variables are equivalent,<sup>2)</sup> are examined.

We prove that the central limit theorem is false, if the equivalent variables are strongly correlated and uniformly bounded, although, it is known<sup>3)</sup> that this theorem is true if the moments of these variables fulfil some special conditions (1).

Even if we suppose that the conditions (1) are satisfied, the asymptotic expansion theorem may be false, if some further moments of the same variables do not fulfil additional conditions (2). However, if we suppose that the conditions (1) and (2) are satisfied, we obtain the following approximation for the distribution function  $F_n(x)$  of the standardized sum of  $n$  equivalent variables:

$$F_n(x) = \Phi(x) + o(n^{-\alpha} \log^{-1} n), \quad 0 < \alpha < 1,$$

where  $\Phi(x)$  is the Gauss-Laplace integral. The properties of the distribution function  $F_n(x)$  and of the corresponding characteristic function  $\varphi_n(t)$  depend on the properties of the joined distribution  $F_n(x_1, \dots, x_n)$  of the equivalent variables  $X_1, X_2, \dots, X_n$  we describe.

Conditions (1) and (2) essentially differ from the analogous conditions for the case of independent random variables. They consist not only

<sup>1)</sup> H. CRAMER: *Random Variables and Probability Distributions*, Cambridge, ch. VII, Th. 25 (1937), 81—83.

<sup>2)</sup> B. de FINETTI: *Classi di numeri aleatori equivalenti*, Rendiconti della Reale Accademia Nazionale dei Lincei, 6a vol. XVIII (1933), 203—207.

<sup>3)</sup> H. MILICER-GRUŻEWSKA: *Sulla legge limite delle variabili casuali equivalenti*, Atti d. Acc. Naz. d. Lincei, S. VIII-v. II, ser. I<sup>a</sup>, 2 (1948), Theorema 1, pp. 28—29.

in the existence of the moments and the estimation of their gradual increase, but also in the fulfilment of an enumerable number of equalities.

The summary report will appear in the *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie* (cl. III) entitled *On the Distribution Law and the Characteristic Function of the Standardized Sum of the Equivalent Variables*.

\*

Streszczenie. — Summary.

### Aproksymacja prawa granicznego.

HALINA MILICER-GRUŻEWSKA, Warszawa.

Autorka bada warunki, przy których spełnione jest asymptotyczne twierdzenie dla zmiennych ewentualnych ekwiwalentnych.

## REPRESENTATIVNÍ VÝBĚR SKUPIN METHODOU DVOU FÁZÍ.

JAROSLAV HÁJEK, Praha.

Při aplikaci reprezentativní metody často nelze vybírat jednotlivé prvky, nýbrž pouze celé jejich skupiny. Jiří Neyman podává řešení, při němž má každá kombinace stejnou naději, že bude vybrána. Potom však nelze výběrový průměr

$$\bar{y} = \frac{r_1 y_1 + \dots + r_k y_k}{r_1 + \dots + r_k},$$

( $r_i$  a  $y_i$  jsou rozsahy a průměry skupin vybraných) užít jako odhad průměru základního, protože má skreslení a je nutno volit veličinu jinou, která však má obvykle větší rozptyl. Aby skreslení výběrového průměru zmizelo, stačí pozměnit výběrový postup, zvolit metodu, již jsem nazval metodou dvou fází. V první fázi totiž vybereme jednu ze skupin tak, aby pravděpodobnost vybrání byla pro každou skupinu úměrná jejímu rozsahu (což lze pomocí tabulek náhodných čísel snadno zařídit). Ve druhé fázi vybereme ( $k - 1$ ) zbývajících skupin obyčejným způsobem, při kterém má každá kombinace skupin stejnou naději.

\*

Summary. — Výťah.

### The cluster sampling by the two-stage method.

JAROSLAV HÁJEK, Praha.

During applications of the sampling method, it is often impossible to select single elements, but only whole groups. J. Neyman presents a solution in which each combination of groups has the same chance

of being selected. Then, however, the sample mean

$$\bar{y} = \frac{r_1 y_1 + \dots + r_k y_k}{r_1 + \dots + r_k}$$

( $r_i$  and  $y_i$  denote the size resp. mean of the selected groups) cannot be used as an estimate of the population average since it has a bias, and it becomes necessary to choose some other quantity which has usually, however, a larger dispersion. In order that the bias of the sample average should vanish, it suffices to alter the sampling procedure: we choose a method which the author has called the two-stage method. The first stage consists of selecting one of the groups in such a way that the probability of selection will be, for each group, proportional to its size. (This can be easily arranged with the help of tables of random numbers). In the second stage, we select the  $k - 1$  remaining groups in the usual way, where each combination of groups has the same chance.

## ZÁKON ČETNOSTI CHYB FYSIKÁLNÍCH MĚŘENÍ.

ZDENĚK HORÁK, Praha.

Podle dosavadních zkušeností lze mít za to, že rozdělení chyb při fyzikálních měřeních splňuje zhruba normální zákon, avšak o jeho přesné platnosti nelze bezpečně rozhodnout, ježto řady fyzikálních měření nejsou obvykle dosti početné. Za účelem ověření normálního zákona chyb, byly provedeny v ústavu technické fyziky na Českém vysokém učení technickém dvě řady můstkových měření po 500 členech a bylo zjištěno, že v obou případech jsou křivky četnosti ostřejší, než křivky Gaussovy, t. j. mají kladný excés. Tento experimentální výsledek pokusil jsem se vysvětlit trojím způsobem, čemuž odpovídají tyto tři alternativy:

a) *Normální zákon četnosti platí přesně pro homogenní soubory měření.* Delší řady měření, které potřebujeme k spolehlivému ověření zákona chyb, není snadno provést za přesně stejných podmínek a proto je musíme pokládat za souhrn jistého počtu částečných souborů o různých středních chybách. Superposicí těchto částečných normálních křivek vzniknou skutečně křivky, které mají charakter pozorovaných frekvenčních křivek.

b) *Normální zákon platí pro skutečná měření jen přibližně, kdežto přesný zákon chyb má obecnější tvar, obsahující normální zákon jako zvláštní případ.* K zobecnění normálního zákona lze dojít na základě kvantové statistiky, která se osvědčila v moderní fyzice. Je známo, že klasická statistika odpovídá zákonu Gaussovu, podobně lze kvantové statistice přiřadit zákon chyb:

$$\eta(x) = \frac{a}{e^{a^2 x^2} - b}, \quad b < 1, \quad (1)$$

kde  $\eta(x)$  je relativní četnost chyby  $x$  a  $\alpha$ ,  $a$ ,  $b$  jsou konstanty. Pro  $b > 0$  má křivka kladný exces, pro  $b < 0$  má exces záporný. První případ odpovídá statistice Bose-Einsteinově, druhý statistice Fermi-Diracově.

c) *Rozdělení skutečných chyb není normální a lze je vyjádřit zákonem jiného tvaru.* S tohoto hlediska by přicházely v úvahu především symetrické typy Pearsonových křivek. Ukázalo se však, že lépe vyhovuje křivka

$$\eta = \frac{\alpha}{\pi \cosh(\alpha x)} \text{ po př. } \eta = \frac{\alpha}{2 \cosh^2(\alpha x)}. \quad (2)$$

Druhá z těchto křivek se dosti málo liší od křivky normální. Tak jsem byl veden k obecnému typu souměrných křivek četnosti

$$\eta = a \cosh^n(\alpha x), \quad (3)$$

které pro  $n < 0$  mají zhruba tvar křivky normální.

V tomto případě lze vyjádřit integrály chyb elementárními transcendentami ( $gd(\alpha x)$ ,  $tgh \alpha x$ ).

K stanovení konstant frekvenčních křivek daných předpoklady  $a$ ,  $b$ ,  $c$ , jsem vypracoval jednoduché početní metody, v nichž jsem užil kromě obvyklých momentů  $\Sigma x^n \eta$  lineárních v  $\eta$  také momentů *absolutních*  $\Sigma |x^n| \eta$  a momentů *vyšších stupňů*, totiž  $\Sigma |x^n| \eta^n$ , na př. kvadratických  $\Sigma |x^n| \eta^2$  a kubických  $\Sigma |x^n| \eta^3$ .

V případě jediné neznámé konstanty — u normální křivky  $a$  u křivek (2) — je výhodno určit ji na základě kumulativních četností, t. j. z průběhu součtové (integrální) křivky. Při tomto postupu lze určit nejpravděpodobnější hodnotu konstanty methodou nejmenších čtverců.

České vysoké učení technické, Ústav technické fyziky.

\*

Summary. — Výtah.

### Law of Error in Physical Measurements.

ZDENĚK HORÁK, Praha.

In order to verify the normal law of error, two series of 500 measurements were made and their distribution functions derived experimentally. It was found that the actual frequencies of small errors were greater than those given by the normal distribution law. To explain this result I suggested three possible assumptions as follows.

a) *The normal law is valid only if the accuracy of all the measurements is the same.* We see that in the case of a long series all measurements cannot be made under precisely equal conditions. We can then obtain a frequency curve by the superposition of several normal curves giving a resulting curve of the type actually observed.

b) *The normal law is merely a special case of a more general law of error still to be found out.* I adopted the function (1) analogous to that

assumed in quantum statistics as a generalized form of the normal distribution function.

c) *The mathematical form of the true law of error differs from that of the normal law.* Searching for a new and more appropriate distribution formula I found, that functions (2) can be admitted as a satisfactory reproduction of the error distribution observed. More generally one can use the function (3).

The method of moments is a general and useful method of finding the constants but the higher the moments the more the accuracy is reduced. For this reason I used, in addition to the usual moments, both the *absolute* moments  $\Sigma|x^n|\eta$  and those of *higher degrees* (e. g. quadratic moments  $\Sigma|x^n|\eta^2$  or cubic ones  $\Sigma|x^n|\eta^3$ ). In the case of one unknown constant, its most probable value is best calculated from the cumulative frequencies (given by the Error integral) using the least squares method.

Technical University of Prague,  
Department of Technical Physics.

## PRUŽNÁ ŘETĚZOVKA.

VÁCLAV HRUŠKA, Praha.

Předmětem sdělení bylo zjednodušení některých důkazů uvedených v níže uvedených pojednáních, v nichž zavedené označení bude i níže použito. Zjednodušení spočívá v tom, že bylo upuštěno od diskontinuitního činitele  $\varepsilon = \pm 1$  (viz [2; vzorec 2, 27]) a celý důkaz proveden pouze s použitím parametrů  $v, u$  [2; § 3 a § 4], jak bylo částečně provedeno již v onom pojednání.

### Literatura.

- [1] V. HRUŠKA: Venkovní elektrická vedení počítaná jako pružná řetězovka (Elektrotechnický Svaz Československý, Praha, 1940).  
[2] V. HRUŠKA: Pružná řetězovka (Rozpravy II. třídy české Akademie v Praze, 54 (1944), čís. 10).

\*

Résumé. — Výtah.

### La chaîne élastique.

VÁCLAV HRUŠKA, Praha.

L'objet de cette communication était la simplification de quelques démonstrations du mémoire [2]. La simplification consiste, d'une part, dans l'omission du facteur de discontinuité  $\varepsilon = \pm 1$  [2; formule 2,27] et, d'autre part, dans le fait que toutes démonstrations ont été basées exclusivement sur l'emploi des paramètres  $v, u$  [2; § 3 et § 4], ce qui était fait déjà en partie dans le mémoire [2].

$$\text{O FUNKCII } T_k = \sum_{s=0}^{k-1} (-1)^s \binom{m}{s} (n-sq)^l. * \quad (1)$$

ANTON HUŤA, Bratislava.

Ako je známe, vyskytujú sa v praxi i v teorii aplikovanej matematiky, a to hlavne v mat. štatistike, vzorce ako

$$P = \frac{1}{l!} \sum_{i=0}^l (-1)^i \binom{l}{i} \left(\frac{n}{q} - i\right)^l \quad (\text{Moiivre-ov problém})$$

$$v(1-g)^{v-1} - \binom{v}{2}(1-2g)^{v-1} + \dots + (-1)^{m-1} \binom{v}{m}(1-mg)^{v-1} \quad (\text{„Fisherov test“})$$

$\frac{n^n}{(n-1)!} \sum_{r=0}^k (-1)^r \binom{n}{r} \left(m - \frac{r}{n}\right)^{n-1}$  (rozdelenie priemerov výberov u pravouholníkového rozloženia).

Všetky tieto vzorce sú špeciálnym prípadom vzorca  $T_k$ . Ide o to vyšetriť vlastnosti tejto funkcie. Ak do (1) zavedieme substitúciu  $n = (k-1)q + r$ , môžeme (1) písať

$$T_k = T(k, m, q, r, l) = \sum_{s=0}^{k-1} (-1)^s \binom{m}{s} [(k-1-s)q + r]^l. \quad (2)$$

Hodnoty  $T_k$  sú koeficienty členov radu určitej vytvorujúcej funkcie  $f(x)$ , ktorú, po určitom počte transformácií a úprav dostaneme vo tvare

$$\begin{aligned} f(x) &= (1-x)^m \varphi_l(x) = \\ &= (1-x)^{m-l-1} \sum_{k=0}^{l+1} x^k \sum_{s=0}^{k-1} (-1)^s \binom{l+1}{s} [(k-1-s)q + r]^l \end{aligned} \quad (3)$$

Funkciu  $f(x)$  si môžeme rozviesť do Laurentovho radu, ktorého koeficient  $T_c$  pri  $x^c$  bude daný vzorcom

$$T_c = \frac{1}{2\pi i} \int_{(0)} \frac{1}{x^{c+1}} (1-x)^{m-l-1} \sum_{k=1}^l x^k \sum_{s=0}^{k-1} (-1)^s \binom{l}{s} [(k-1-s)q + r]^{l-1} \quad (4)$$

Ako zo vzorca (3) plynie

$$\varphi_l(x) = \frac{P_{l+1}(x)}{(1-x)^{l+1}}, \quad (5)$$

kde  $P_{l+1}$  je polynóm  $(l+1)$ -ho stupňa v  $x$ .

Polynómy  $P$  prejavujú určitú symetrickosť, ktorú môžeme ukázať vzorcami

\*) Sdělení nebylo předneseno, protože se autor nemohl zúčastniti sjezdu.

$$P_{2v}(x) = \sum_{t=1}^v g_t(u, r)x^{2v-t+1} + g_t(r, u)x^t, \quad (6)$$

$$P_{2v+1}(x) = g_{v+1}(u, r)x^{v+1} + \sum_{t=1}^v g_t(u, r)x^{2v-t+2} + g_t(r, u)x^t. \quad (7)$$

Špecializáciou týchto vzorcov dostaneme individuálne vlastnosti jednotlivých vzorcov v praxi sa vyskytujúcich. Ako príklad uvažme funkciu

$$T_n^{(l)} = \sum_{t=0}^N (-1)^t \binom{n}{t} (n-2t)^l \quad \frac{n}{2} - 1 \leq N < \frac{n}{2}. \quad (8)$$

Zo vzorca (4) môžeme dostať hodnoty tejto funkcie. Ak je  $n$  aj  $l$  súčasne sudé alebo liché, potom  $T_n^{(l)} = 0$ , zo (6) a (7) dostaneme špecializáciou reciproké polynómy atď.

\*

Summary. — Výťah.

**Concerning the function**  $T_k = \sum_{s=0}^{k-1} (-1)^s \binom{m}{s} (n-sq)^l$ .\*) (1)

ANTON HUŤA, Bratislava.

It is a well-known fact that in practical work and in the theory of applied mathematics, especially in mathematical statistics, there very often occur expressions of the form

$$P = \frac{1}{l!} \sum_{i=1}^l (-1)^i \binom{l}{i} \left(\frac{n}{q} - i\right)^l \quad (\text{de Moivre's problem})$$

$$\nu(1-g)^{\nu-1} - \binom{\nu}{2}(1-2g)^{\nu-1} + \dots + (-1)^{m-1} \binom{\nu}{m}(1-mg)^{\nu-1}$$

(„Fisher's test“)

$$\frac{n^n}{(n-1)!} \sum_{r=0}^k (-1)^r \binom{n}{r} \left(m - \frac{r}{n}\right)^{n-1} \quad (\text{distribution of sample}$$

means for a rectangular distribution).

All these expressions are special cases of the relation  $T_k$ . We wish to investigate the properties of this function. If we introduce into (1) the substitution  $n = (k-1)q + r$ , we can write (1) in the form

$$T_k = T(k, m, q, r, l) = \sum_{s=0}^{k-1} (-1)^s \binom{m}{s} [(k-1-s)q + r]^l. \quad (2)$$

\*) The author being prevented from participating in the congress did not read this communication.



The values of  $T_k$  are the coefficients of the members of the series of a certain generating function  $f(x)$  which, after several transformations and rearrangements, can be written

$$f(x) = (1-x)^m \varphi_l(x) = \\ = (1-x)^{m-l-1} \sum_{k=0}^{l+1} x^k \sum_{s=0}^{k-1} (-1)^s \binom{l+1}{s} [(k-1-s)q+r]^l. \quad (3)$$

The function  $f(x)$  can be expanded into the form of a Laurent series whose coefficient  $T_c$  of  $x^c$  will be given by the expression

$$T_c = \frac{1}{2\pi i} \int_{(0)} \frac{1}{x^{c+1}} (1-x)^{m-l-1} \sum_{k=1}^l x^k \sum_{s=0}^{k-1} (-1)^s \binom{l}{s} [(k-1-s)q+r]^{l-1} \quad (4)$$

From (3) it follows

$$\varphi_l(x) = \frac{P_{l+1}(x)}{(1-x)^{l+1}}, \quad (5)$$

where  $P_{l+1}$  is a polynomial of degree  $(l+1)$  in  $x$ .

The polynomials  $P$  display a certain symmetry, which can be demonstrated by means of the expressions

$$P_{2v}(x) = \sum_{t=1}^v g_t(u, r) x^{2v-t+1} + g_t(r, u) x^t, \quad (6)$$

$$P_{2v+1}(x) = g_{v+1}(u, r) x^{v+1} + \sum_{t=1}^v g_t(u, r) x^{2v-t+2} + g_t(r, u) x^t. \quad (7)$$

For special cases of these expressions we obtain the individual properties of the expressions occurring in practical work. As an example, let us consider the function

$$T_n^{(l)} = \sum_{t=0}^N (-1)^t \binom{n}{t} (n-2t)^l, \quad \frac{n}{2} - 1 \leq N < \frac{n}{2}. \quad (8)$$

From (4) we can obtain the values of this function. If  $n$  and  $l$  are simultaneously odd or even, then  $T_n^{(l)} = 0$ ; for special cases, we obtain from (6) and (7) reciprocal polynomials, etc.

## TWIERDZENIE GRANICZNE O WYRAZACH WYŻSZEGO RZĘDU W ZAGADNIENIU BAYESA.

WŁODZIMIERZ KRYSICKI, Łódź.

Wiadomo, że zagadnienie BAYESA polega na wyznaczeniu prawdopodobieństwa, że prawdopodobieństwo  $p$  zajścia danego zdarzenia w poszczególnym doświadczeniu zawiera się w danym przedziale, przy

czym wiemy, że w ciągu  $n$  jednakowych i niezależnych doświadczeń dane zdarzenie powtórzyło się  $\alpha$  razy oraz, że gęstość prawdopodobieństwa a priori zmiennej  $p$  ma daną wartość  $f(p)$ . Gęstość powyższego prawdopodobieństwa a posteriori wyraża się wzorem

$$P_n(p) = \frac{f(p)p^\alpha q^{n-\alpha}}{\int_0^1 f(p)p^\alpha q^{n-\alpha} dp}, \quad (1)$$

gdzie  $q = 1 - p$ .

Przy założeniu, że stosunek  $\frac{\alpha}{n} = p_0$  pozostaje stały, gdy jednocześnie  $\alpha$  i  $n$  rosną nieograniczenie, asymptotyczną wartością (1) jest

$$P_n(p) \approx \sqrt{\frac{n}{2\pi p_0 q_0}} e^{-\frac{n}{2p_0 q_0} (p - p_0)^2} \quad (2)$$

Następnie załóżmy, że wielkości  $n$  i  $p$  zmieniają się w taki sposób, że wyrażenie

$$\sqrt{\frac{n}{2p_0 q_0}} (p - p_0) = t^2 \quad (3)$$

pozostaje stałe. Nasuwa się wówczas zagadnienie wyznaczenia takich dwóch liczb rzeczywistych  $\gamma$  i  $K \neq 0$ , aby

$$\lim_{n \rightarrow \infty} \left\{ n^\gamma \left[ \frac{P_n(p)}{\sqrt{n}} - \frac{1}{\sqrt{2\pi p_0 q_0}} e^{-t^2} \right] \right\} = K,$$

gdzie na podstawie (3) jest

$$p = p_0 + \sqrt{\frac{2p_0 q_0}{n}} t.$$

Na podstawie wzoru (2) widoczne jest, że gęstość prawdopodobieństwa przy  $n \rightarrow \infty$  w każdym punkcie  $p \neq p_0$  dąży do zera, w punkcie zaś  $p = p_0$  rośnie nieograniczenie. Obierając odpowiedni przedział otaczający punkt  $p_0$ , przez stopniowe szacowanie prawej strony wzoru (1) — szczególnie całki występującej w mianowniku — i zważanie wybranego przedziału, otrzymuje się wynik:

$$\gamma = \frac{1}{2},$$

$$K = \left[ \frac{2(q_0 - p_0)}{3p_0 q_0} t^2 + \frac{f'(p_0)}{f(p_0)} \right] \frac{te^{-t^2}}{\sqrt{\pi}}. \quad (4)$$

1) R. MISES: *Wahrscheinlichkeitsrechnung u. ihre Anwendung in der Statistik u. theoretischen Physik*, Leipzig, 1931, str. 158.

2) W. POGORZELSKI: *Zarys rachunku prawdopodobieństwa i teorii błędów*, Warszawa, 1948, str. 28.

Jeżeli funkcja  $f(p) = \text{constans}$ , to jako szczególny przypadek (4) otrzymuje się

$$K = \frac{2(q_0 - p_0)}{3p_0q_0\sqrt{\pi}} t^3 e^{-t^2}.$$

Ten ostatni wynik można również otrzymać ze wzoru (1) — po uproszczeniu ułamka przez  $f(p)$  — stosując przekształcenia oparte o wzór STIRLINGA.

(Cała praca (rozszerzona) ukaże się w jednym z polskich czasopism naukowych.)

\*

Résumé. — Streszczenie.

### Un théorème limite sur les termes d'ordre supérieur dans le problème de Bayes.

WŁODZIMIERZ KRYSICKI, Łódź.

On sait que le problème de Bayes consiste dans la recherche de probabilité pour que la probabilité  $p$  d'un événement dans une expérience est compris dans un intervalle, étant donnée que dans la suite de  $n$  expériences identiques et indépendantes l'événement se produisait  $\alpha$  fois; en outre on admet que la densité de probabilité a priori de la variable  $p$  est une fonction donnée  $f(p)$ . La densité de probabilité a posteriori qui est la solution du problème de Bayes est donnée par la formule suivante.

$$P_n(p) = \frac{f(p)p^\alpha q^{n-\alpha}}{\int_0^1 f(p)p^\alpha q^{n-\alpha} dp}, \quad (1)$$

où  $q = 1 - p$ .

En supposant que la fréquence  $\frac{\alpha}{n} = p_0$  est constante, quand  $\alpha$  et  $n$  tendent simultanément vers l'infini, une expression asymptotique de  $P_n(p)$  est donnée par la formule

$$P_n(p) \approx \sqrt{\frac{n}{2\pi p_0 q_0}} e^{-\frac{n}{2p_0 q_0} (p - p_0)^2}. \quad (2)$$

Supposons maintenant que  $n$  et  $p$  varient de façon que

$$\sqrt{\frac{n}{2p_0 q_0}} (p - p_0) = t \quad (3)$$

soit une constante. Il faut trouver deux nombres réels  $\gamma$  et  $K \neq 0$  tels que

$$\lim_{n \rightarrow \infty} \left\{ n^\gamma \left[ \frac{P_n(p)}{\sqrt{n}} - \frac{1}{\sqrt{2\pi p_0 q_0}} e^{-t^2} \right] \right\} = K,$$

où  $p$  est donné par l'expression (3).

Le calcul donne la réponse suivante:

$$\gamma = \frac{1}{2},$$

$$K = \left[ \frac{2(q_0 - p_0)}{3p_0 q_0} t^2 + \frac{f'(p_0)}{f(p_0)} \right] \frac{te^{-t^2}}{\sqrt{\pi}}. \quad (4)$$

Dans le cas où la fonction  $f(p)$  est constante, on obtient

$$K = \frac{2(q_0 - p_0)}{3p_0 q_0 \sqrt{\pi}} t^3 e^{-t^2}.$$

On peut obtenir aussi ce dernier résultat de la formule (1) à l'aide de la formule de STIRLING.

## ON THE ERGODIC THEOREM AND THE LAW OF LARGE NUMBERS.

EDWARD MARCZEWSKI, Wrocław.

Remarks on BIRKHOFF'S ergodic theorem, KOLMOGOROFF'S law of large numbers and DE FINETTI'S notion of equivalent random variables.<sup>1)</sup>

Some of the results will be published in *Colloquium Mathematicum*.

\*

Streszczenie. — Summary.

### O twierdzeniu ergodycznym i prawie wielkich liczb.

EDWARD MARCZEWSKI, Wrocław.

## THE ESTIMATION OF STRUCTURAL PARAMETERS OF PARTIALLY CONSISTENT SERIES.

BOHUMIL PARDUBSKÝ, Praha.

We start out from the assumption that we are given a certain random sample from a given basic population, possessing a cumulative distribution function whose mathematical form is known, but which

<sup>1)</sup> Only after the Congress I had a possibility to peruse the paper by J. L. DOOB, *The law of large numbers for continuous stochastic processes*, *Duke Math. Journal*, **6** (1940), 290—306, which contains (Theorem 1, p. 291) the main results of my communication.

contains a certain number of parameters which are unknown. The problem is to estimate these parameters as functions of the values of the random sample. The number of such functions is infinitely large. The question arises which of these estimates is the best.

Let  $T_i(x_1, \dots, x_n)$  for  $i = 1, 2, \dots, \nu$  be a system of estimates of the parameters  $\Theta_i (i = 1, 2, \dots, \nu)$ . We shall consider the best system of estimates to be the one whose simultaneous distribution of probability density of the estimates has maximum concentration. On the assumption that certain conditions of regularity are fulfilled, HARALD CRAMÉR established in the Skandinavisk Aktuarietidskrift in 1946 the equation of the ellipsoid of maximum concentration belonging to a given simultaneous distribution of probability density of estimates, in the form

$$\sum_{ij} \kappa_{ij} (T_i - \Theta_i)(T_j - \Theta_j) = \nu + 2,$$

where

$$\kappa_{ij} = E \left[ \frac{\partial \log f}{\partial \Theta_i}, \frac{\partial \log f}{\partial \Theta_j} \right], \quad f = f(x_1, \dots, x_n | \Theta_1, \dots, \Theta_n)$$

is the simultaneous distribution of probability density of the variables  $x_1, \dots, x_n$ . The random variables of a sequence  $\{\xi_i\}$  are said to be consistent if the number of parameters involved in their distribution functions of probability density is finite, and if each parameter occurs in all or an infinite number of distributions of probability density. The random variables of a sequence  $|\xi_i|$  are said to be partially consistent if the parameters involved in their distributions of probability density can be divided into two groups. The first group consists of a finite number of parameters and these parameters are contained in all distributions of probability density. The second group of parameters is composed of an infinite number of parameters of which each occurs in only a finite number of distributions of probability density. The parameters of the first group are called structural parameters; the parameters of the second group are called incidental parameters.

Because estimates of structural parameters of partially consistent sequences determined by the method of maximum likelihood are not necessarily consistent, new systematic methods were sought which would give consistent estimates. J. NEYMAN and ELIZABETH SCOTT established five conditions which the system of equations

$$F_{si}(x_1, \dots, x_n, T_1, \dots, T_\nu) = 0, \quad i = 1, 2, \dots, \nu$$

must satisfy i. e. the system of estimates of the structural parameters  $\Theta_i (i = 1, 2, \dots, \nu)$ .

Because of the difficulty of fulfilling these conditions it was decided to abandon this general method and to look for a method for a more special, but frequently occurring, case. The method may be called the

method of modified equations of maximum likelihood, where we need not limit ourselves to only one incidental parameter, always occurring in the function of one random variable, but where there may be any number. A necessary condition is that the mathematical chance of the partial derivative of the logarithm of the appropriate distribution of probability density, after elimination of the incidental parameters, should be a constant, or should be a function only of the structural parameters.

The squares of the standard deviations of the estimates of individual parameters have lower limits given by the coefficients of the quadratic form which is the inverse to the quadratic form of the corresponding ellipsoid of maximum concentration.

\*

Výtah. — Summary.

### Stanovení odhadu strukturálních parametrů parciálně konsistentních serií.

BOHUMIL PARDUBSKÝ, Praha.

System odhadů parametrů možno klasifikovati pomocí elipsoidu koncentrace. Za „nejlepší“ systém odhadů pokládáme pak takový systém, jemuž přísluší elipsoid o maximální koncentraci. Nalezení systému odhadů strukturálních parametrů provádí se na podkladě metody největší věrohodnosti k tomu účelu modifikované. Klasifikací odhadů lze se přesvědčit, zda takto získaný systém odhadů je „nejlepší“.

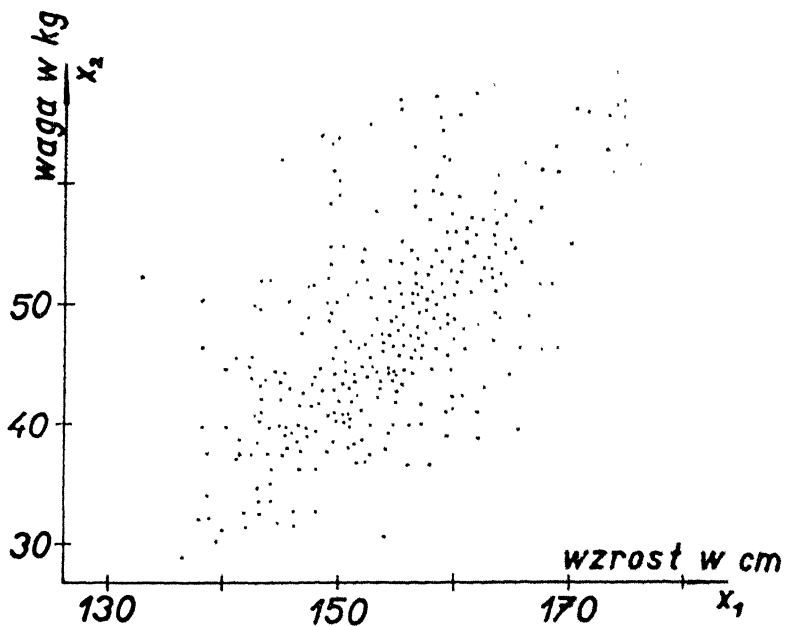
### O PEWNYCH KORELACJACH OBSZAROWYCH.

JULIAN PERKAL, Wrocław.

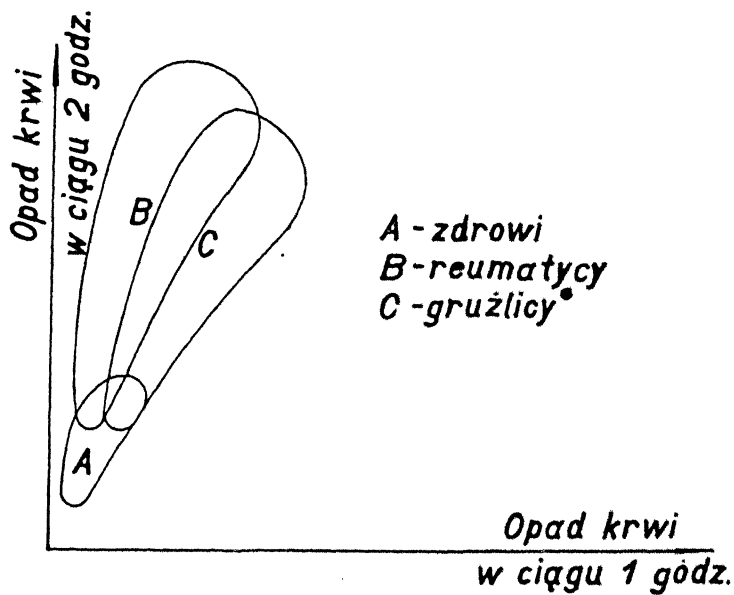
Przez korelację między zmiennymi  $x$  i  $y$  rozumie się zazwyczaj zależność typu

$$y = f(x), \text{ lub } g(x, y) = 0. \quad (1)$$

Przypuśćmy, że w wyniku 1000 doświadczeń otrzymamy 1000 par liczb  $x, y$ . Można je sobie wyobrazić jako 1000 punktów na płaszczyźnie  $x, y$ . Jeśli te punkty leżą w pobliżu jakiejś prostej, paraboli, lub innej krzywej, to mówimy o korelacji liniowej, parabolicznej, lub innej krzywoliniowej. Jednakże często się zdarza, że punkty otrzymane w wyniku doświadczeń nie leżą w pobliżu żadnej krzywej, lecz wypełniają pewien obszar dwuwymiarowy. Tak np. ułożą się punkty, których współrzędnymi będą wzrost i waga 14-letnich chłopców. Innym przykładem będzie następujący: Charakterystycznym przy pewnych chorobach jest tzw. „Objaw Biernackiego“ (O. B.) wyznaczony przez dwie liczby opadu



Rys. 1. Punkty o współrzędnych  $x_1$  = wzrost,  $x_2$  = waga 14-letnich chłopców.



Rys. 2. Obszary odpowiadające różnym grupom badanych na „O. B.”

krwi (po godzinie i po dwóch godzinach). Jeśli te dwie liczby traktować jako współrzędne punktów płaszczyzny, to okaże się, że odpowiednie punkty rozrzucone są w pewnych obszarach płaszczyzny i to w obszarach odpowiadających różnym chorobom. Tak np. można wyodrębnić obszar odpowiadający ludziom zdrowym, obszar reumatyków, obszar gruźlików itd. Badania nad tymi zjawiskami (prowadzone wspólnie z dr T. Nowakowskim) są w toku. Niestety nie udało się nam dotychczas osiągnąć wyników tak ostrych, by miały one praktyczne zastosowanie w medycynie.

Podobna sytuacja zachodzi przy badaniu zależności między większą niż 2 liczbą zmiennych. Takie wieloczynnikowe korelacje odgrywają specjalnie dużą rolę w badaniach psychotechnicznych związanych z analizą testów. Wyobraźmy sobie, że pewna metoda badania testowego polega na wypełnieniu 10 testów. Oceny ich stanowią 10 liczb charakteryzujących danego osobnika. Metodę taką należy zrelatywizować do reprezentacyjnej populacji złożonej np. z 1000 osobników. Oceny testów wypełnionych przez członków tej reprezentacyjnej populacji można ustawić w macierz o 10 kolumnach i 1000 wierszach. Każdy wiersz odpowiada jednemu osobnikowi i może być traktowany w przestrzeni 10-cio wymiarowej jako punkt. Badanie korelacji pierwszego stopnia, czy wyższych stopni sprowadza się do badania, czy wyznaczone punkty leżą w pobliżu jakiejś hiperpłaszczyzny, czy hiperpowierzchni w rozważanej przestrzeni 10-cio wymiarowej. Szczególnie wiele badań tego typu prowadzi się różnymi metodami w Ameryce.<sup>1)</sup> Podobnie jak w przypadku dwuwymiarowym, często się zdarza, że punkty odpowiadające układom zmiennych nie leżą w pobliżu jakiegoś tworzywa o niższej liczbie wymiarów niż przestrzeń, lecz wypełniają jakieś obszary tej przestrzeni. Tak np. się zdarza, gdy do naszych prób postawienia diagnozy przy pomocy badania położenia punktu odpowiadającego choremu wprowadzić prócz dwóch współrzędnych — liczb opadu krwi — jeszcze trzecią liczbę dermatografizmu (jest to długość okresu występowania śladu podrażnienia skóry). Punkty przestrzeni trójwymiarowej odpowiadające poszczególnym osobnikom grupują się w pewnych obszarach odpowiadających znowu poszczególnym chorobom.

Nierówności, które muszą spełniać współrzędne punktu na to, by punkt ten leżał w wyróżnionym obszarze, można nazwać korelacjami obszarowymi.

Okazuje się, że można wyróżnić pewien układ współrzędnych związanych z dowolnym takim zbiorem punktów, oraz można wyróżnić pewną rodzinę minimalnych obszarów, wewnątrz których zawierają się określone części (ilości) punktów rozważanego zbioru. Rozwiążemy kolejno te dwa zagadnienia.

<sup>1)</sup> Por. np. monografię L. L. THURSTONE: *Multiple — Factor Analysis*, Chicago Press, 1947.



Niech będzie macierz prostokątna

$$(x_j^i). \quad (2)$$

Numer wiersza  $i = 1, 2, \dots, N$  jest numerem doświadczenia, a numer kolumny  $j = 1, 2, \dots, n$  jest numerem danej w każdym doświadczeniu. Każdy wiersz wyznacza punkt w przestrzeni  $n$ -wymiarowej. Dla znalezienia układu współrzędnych wyróżnionego ze względu na zbiór  $Z$  punktów wyznaczonych przez wiersze macierzy (2), rozwiążemy następujące zadanie: Znaleźć hiperpłaszczyznę  $n - 1$  wymiarową

$$L(x_1, x_2, \dots, x_n) \equiv A_1 x_1 + A_2 x_2 + \dots + A_n x_n + A_{n+1} = 0 \quad (3)$$

taką, by suma kwadratów odległości wszystkich punktów zbioru  $Z$  od tej hiperpłaszczyzny osiągała minimum. Zakładając

$$A_1^2 + A_2^2 + \dots + A_n^2 = 1, \quad (4)$$

można warunek na minimum wyrazić:

$$\frac{1}{N} \sum_{i=1}^{i=N} L^2(x_1^i, x_2^i, \dots, x_n^i) = \min. \quad (5)$$

Na mnożnik Lagrange, a  $\lambda$  otrzymamy, po znanych operacjach,  $n$  wartości — pierwiastków równania sekularnego:

$$\begin{vmatrix} c_{11} - \lambda, & c_{12}, & c_{13}, & \dots, & c_{1n} \\ c_{21}, & c_{22} - \lambda, & c_{23}, & \dots, & c_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ c_{n1}, & c_{n2}, & c_{n3}, & \dots, & c_{nn} - \lambda \end{vmatrix} = 0, \quad (6)$$

gdzie  $c_{ij} = c_{ji}$  jest współczynnikiem kowariancji  $i$ -tej i  $j$ -tej współrzędnych, a w szczególności  $c_{ii}$  jest wariancją (kwadratem dyspersji)  $i$ -tej współrzędnej. Znalezione  $\lambda_i$  są jednocześnie ekstremalnymi wartościami lewej strony wyrażenia (5).

Każde  $\lambda_i$  wyznacza jedną hiperpłaszczyznę (3) i jest równe średniemu kwadratowi odległości od niej punktów zbioru  $Z$ . Hiperpłaszczyzny te są parami do siebie prostopadłe i tworzą układ prostokątnych współrzędnych.

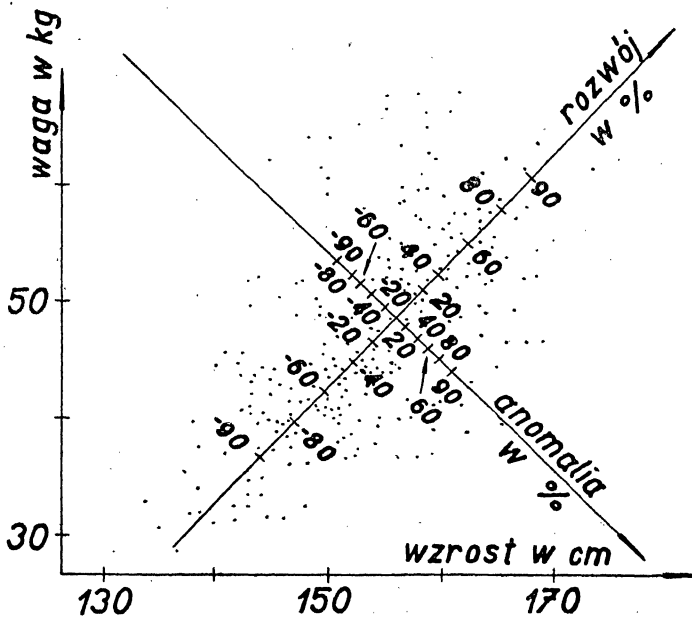
**Twierdzenie I.** Jeśli istnieje układ współrzędnych prostokątnych, niezależnych ze względu na zbiór  $Z$ , to jest nim skonstruowany powyżej układ wyróżniony.

Układ współrzędnych wyróżniony ze względu na zbiór  $Z$  oznaczmy przez  $y_1, y_2, \dots, y_n$ . Współrzędne niezależne posiadają następującą własność: przy dowolnym ustaleniu dowolnych  $n - 1$  współrzędnych średnia pozostałej zmiennej współrzędnej jest stała. A zatem istnieje  $n - 1$  wymiarowa hiperpłaszczyzna prostopadła do osi tej zmiennej współrzędnej, której każdy punkt jest środkiem tej części zbioru  $Z$ , która leży na prostej prostopadłej (dla  $n$  skończonych — wewnątrz pewnego walca o tworzących prostopadłych) do hiperpłaszczyzny i prze-

chodzącej przez ten punkt. Hiperpłaszczyzna taka spełnia warunek (5), a zatem jest jedną z płaszczyzn  $y_i = 0$ .

W dalszym ciągu będziemy zakładali, że zbiór  $Z$  posiada układ współrzędnych niezależnych.

Pierwiastki  $\lambda_1, \lambda_2, \dots, \lambda_n$  są wariancjami zbioru  $Z$  względem hiperpłaszczyzn układu. Może się okazać, że jedna, lub  $s$  tych wariancyj jest



Rys. 3. Współrzędne niezależne (rozwój i anomalia) zbioru pokazanego na Rys. 1.

dostatecznie mała w stosunku do tolerancji z jaką wyznaczamy współrzędne  $x_i$ . Wówczas można zbiór  $Z$  zastąpić rzutem  $Z$  na odpowiednią hiperpłaszczyznę  $n - 1$ , lub  $n - s$  wymiarową, bez uszczerbku dla dokładności wyników. Wróćmy do przykładu z testami. Przypuśćmy, że każdy z 10-ciu testów jest oceniany liczbą rzeczywistą zawartą między 0 i 100, przyczem dokładność tej oceny nie przewyższa jedności. Przypuśćmy dalej, że  $\lambda_1, \lambda_2, \lambda_3$  i  $\lambda_4$  są mniejsze od 0.1. Wynika stąd, że wszystkie punkty zbioru  $Z$  leżą z prawdopodobieństwem większym niż 0,998 bliżej niż o jednąć od hiperpłaszczyzny 6-ciuwymiarowej  $y_1 = y_2 = y_3 = y_4 = 0$ . Jeśli więc zbiór  $Z$  zastąpimy w tym przypadku jego rzutem na tę hiperpłaszczyznę, to żadna ze współrzędnych  $x_i$  nie zmieni się więcej niż o 1. Wynika stąd, że te same informacje, jakie mamy z 10-ciu można uzyskać z 6-ciu tylko testów. Usunąć można oczywiście dowolne 4 testy. W praktyce usunąć należy testy o numerach

znalezionych w następujący sposób: jeśli  $\lambda_1$  jest dostatecznie małe to obliczamy kąty między osią  $y_1$  a wszystkimi osiami  $x_i$ ; jeśli najmniejszy z tych kątów tworzy np. os  $x_7$ , to właśnie test Nr 7 należy usunąć.

Pozostałe 6 pierwiastków  $\lambda_5, \lambda_6, \dots, \lambda_{10}$  mogą być użyte jako jednostki miary na osiach  $y_5, y_6, \dots, y_{10}$ . Osi te można odpowiednio nazwać np. osią pamięci, osią szybkiej orientacji, osią staranności, cierpliwości itd., co zresztą jest sprawą psychologów. Jeśli trzeba zbadać osobnika nie należącego do rozważanej populacji reprezentacyjnej, to można wyliczyć odpowiadające mu współrzędne  $y_5 \dots y_{10}$  (w jednostkach  $\lambda_5 \dots \lambda_{10}$ ) i stąd określić jego cechy.

Przejdźmy do drugiego zagadnienia, znalezienia obszaru, który przy danych wymiarach (hiperobjętości) posiadałby największe prawdopodobieństwo zawierania punktów zbioru  $Z$ .

Jasne jest, że kształt tego obszaru zależy od rozkładu zbioru  $Z$ . Sposób tej zależności omówię na przykładzie dwuwymiarowym. Oznaczmy przez  $x_1$  wzrost, a przez  $x_2$  wagę 14-letnich chłopców. Zbiór  $Z$  będzie leżał w płaszczyźnie  $x_1, x_2$  (p. rys. 1). Rozkład tego zbioru reprezentuje jakaś powierzchnia nad płaszczyzną  $x_1, x_2$ . Warstwica tej powierzchni ograniczająca pole  $P$  z góry zadane posiada żadaną własność zawierania punktów zbioru  $Z$  z największym prawdopodobieństwem. Rzeczywiście, każda inna krzywa ograniczająca takie samo pole będzie się od warstwicy różniła polem  $p$ , na którym gęstość prawdopodobieństwa będzie mniejsza niż wewnątrz warstwicy.<sup>2)</sup>

Przechodząc do przypadku ogólnego, zbiór  $Z$  musimy rozpatrywać w przestrzeni  $n$ -wymiarowej, a hiperpowierzchnię rozkładu można nadbudować dołączając następny,  $n + 1$ -szy wymiar (mnożąc kartezjańsko przez os  $y_{n+1}$ ). Hiperwarstwicami nazwiemy rzuty na przestrzeń  $y_{n+1} = 0$  przekrojów hiperpowierzchni rozkładu przez hiperpłaszczyzny  $y_{n+1} = \text{const}$ . Powtarzając poprzednie rozumowanie mamy:

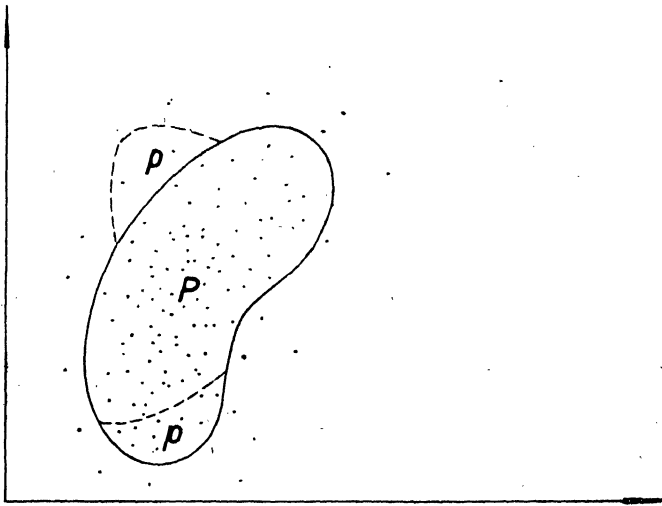
**Twierdzenie 2.** Obszary ograniczone przez hiperwarstwice posiadają, przy danych wymiarach, największe prawdopodobieństwo zawierania punktów rozważanego zbioru.

Szczególnie prosto wypadają rachunki, jeśli rozkład zbioru  $Z$  jest normalny. Wówczas jako warstwice otrzymuje się elipsy, elipsoidy i hiperelipsoidy o równaniach

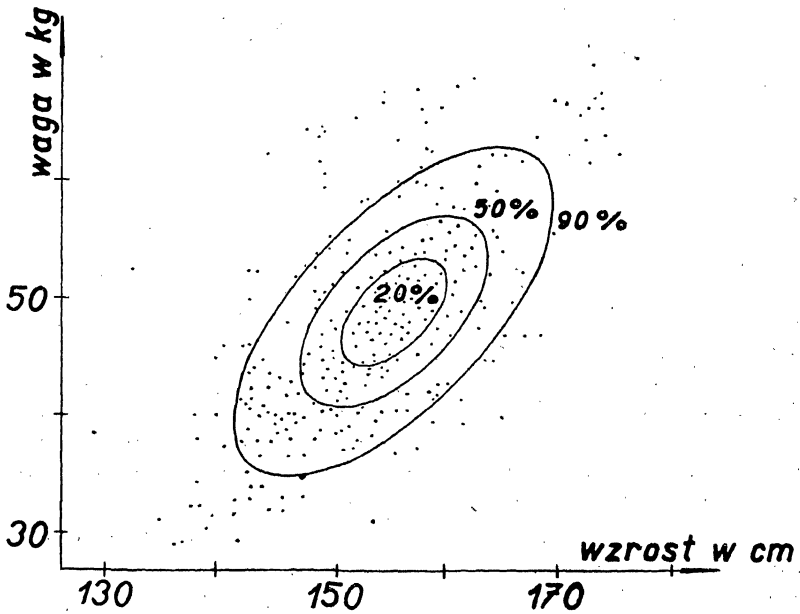
$$\frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} + \dots + \frac{y_n^2}{\lambda_n} = R^2, \quad (7)$$

gdzie  $\lambda_i$  są pierwiastkami równania (6), a  $R$  parametrem rodziny hiperwarstwic. Prawdopodobieństwo tego, że  $y_i$  zawiera się w przedziale  $y_i - \frac{1}{2} dy_i, y_i + \frac{1}{2} dy_i$  wyraża się wobec normalności rozkładu wielkością

<sup>2)</sup> Rozumowanie to, jak zresztą wiele innych uwag i sam temat zawdzięczam p. H. STEINHAUSOWI.



Rys. 4. Warstwica pow. rozkładu zawiera więcej punktów niż inna krzywa o tym samym polu.



Rys. 5. Najmniejsze elipsy zawierające 20%, 50% i 90% punktów zbioru pokazanego na Rys. 1.

$$(2\pi\lambda_i)^{-1} e^{-\frac{y_i^2}{2\lambda_i}} dy_i \quad (8)$$

a wobec tego, że współrzędne  $y_i$  są niezależne, prawdopodobieństwo tego, że punkt znajdzie się wewnątrz kostki elementarnej o wymiarach  $dy_1, dy_2, \dots, dy_n$  będzie iloczynem takich wyrażeń; czyli

$$(2\pi)^{-in} (\lambda_1 \lambda_2 \dots \lambda_n)^{-1} e^{-\frac{1}{2} \left( \frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} + \dots + \frac{y_n^2}{\lambda_n} \right)} dy_1 \dots dy_n. \quad (9)$$

Prawdopodobieństwo zawierania się punktu wewnątrz obszaru (7) będzie całką wyrażenia (9) po tym obszarze. Całka ta przy użyciu podstawień hipersferycznych uogólnionych da się w przypadku  $n$  parzystego wyrazić funkcjami elementarnymi, w przypadku zaś  $n$  nieparzystego sprowadza się do całki funkcji Gaussa.

Korzystając z tych wyników można do omówionej poprzednio metody badań testowych dodać jeszcze wskaźnik normalności badanego osobnika. Jeśli jego współrzędne  $y_1 \dots y_{10}$  wtawić do lewej strony wyrażenia (7) — otrzymamy  $R$  tj. czynnik przez jaki należy pomnożyć dyspersje  $(\lambda_i)^{-1}$ , aby punkt odpowiadający badanemu osobnikowi znalazł się na hiperelipsoidzie (7) o wyliczonych półośiach. Można ułożyć tablice podające dla różnych  $R$  prawdopodobieństwo zawierania się punktów wewnątrz, lub na powierzchni odpowiedniej hiperelipsoidy, a stąd stopień normalności badanego osobnika.

\*

Резюме: — Streszczenie.

### О некоторых корреляциях в областях.

ЮЛНАИ ПЭРКАЛЬ, Вроцлав.

Корреляциями в областях между двумя переменными называем неравенства, которым удовлетворяют эти переменные. Аналогически понимаются корреляции между  $n$  переменными. Такие корреляции получаем, если некоторую совокупность предметов изучаем по их  $n$  признакам. Эти предметы можно изобразить в  $n$ -мерном пространстве точками, а всю совокупность множеством точек.

В настоящей статье разрешаются две задачи: 1. Конструкция независимых координат для такого множества точек. 2. Конструкция минимальных областей содержащих определенную долю всех точек такого множества.

## O ANAMORFOSE V NOMOGRAFII.

VÁCLAV PLESKOT, Praha.

Obsahem sdělení byly stručně shrnuté výsledky o anamorfose.

1. Abychom mohli zobrazit vztah

$$F_{123} = 0 \quad (1)$$

mezi třemi proměnnými  $x_1, x_2, x_3$  spojnicovým nomogramem, je třeba a stačí, abychom jej přepsali na tvar s Massauovým determinantem

$$\begin{vmatrix} f_1, & g_1, & h_1 \\ f_2, & g_2, & h_2 \\ f_3, & g_3, & h_3 \end{vmatrix} = 0, \quad (2)$$

v němž jsou rozloučeny funkce jednotlivých proměnných (v  $i$ -tém řádku jsou funkce pouze  $i$ -té proměnné, při čemž obecně  $f_1 \neq f_2 \neq f_3$  a podobně u  $g_i$  a  $h_i$ ;  $i = 1, 2, 3$ ).

Položíme-li totiž

$$\xi_i = \frac{f_i}{h_i}, \quad \eta_i = \frac{g_i}{h_i} \quad (i = 1, 2, 3), \quad (3)$$

dostaneme pro trojice  $(x_1, x_2, x_3)$ , které vyhovují rovnici (1), z rovnic (3) souřadnice tří bodů, které leží na téže přímé spojnici.

Měníme-li okrouhle hodnoty téže proměnné, vytvoří příslušné body s připsanými hodnotami proměnné stupnici této proměnné.

2. Podaří-li se nám rovnici (1) přepsat na tvar (2), pak rovnice, které dostaneme z (2) vynásobením nenulovým determinantem (podle d'Ocagne)

$$\begin{vmatrix} a_1, & b_1, & c_1 \\ a_2, & b_2, & c_2 \\ a_3, & b_3, & c_3 \end{vmatrix} \neq 0, \quad (4)$$

t. j. rovnice tvaru

$$|a_1 f_i + b_1 g_i + c_1 h_i, a_2 f_i + b_2 g_i + c_2 h_i, a_3 f_i + b_3 g_i + c_3 h_i| = 0,$$

připouštějí zobrazení spojnicovými nomogramy o zobrazovacích rovnicích

$$\bar{\xi}_i = \frac{a_1 f_i + b_1 g_i + c_1 h_i}{a_3 f_i + b_3 g_i + c_3 h_i}, \quad \bar{\eta}_i = \frac{a_2 f_i + b_2 g_i + c_2 h_i}{a_3 f_i + b_3 g_i + c_3 h_i}. \quad (5)$$

Nákresy sestojené podle zobrazovacích rovnic (3) a (5) jsou v kolinearci.

3. Transformaci, po níž lze přepsat vztah (1) na tvar (2), nazýváme anamorfosou.

Obecné řešení problému anamorfosy musí zodpovědět dvě otázky:

- a) jaké jsou podmínky pro vztah, aby byl schopný anamorfosy,
- b) je-li vztah schopný anamorfosy čili anamorfosabilní, připouští-li anamorfosy od sebe kolinearně rozdílné.

SOREAU (1921) podal řešení pro vztahy nomograficky uspořádané t. j. takové, v nichž lze vždy funkce jedné proměnné separovatí od funkcí ostatních dvou. (Rovnici  $F_{12}f_3 + G_{12}g_3 + H_{12}h_3 = 0$  nazýváme nomograficky uspořádanou vzhledem k třetí proměnné.)

Soreau uvedl věty, v nichž osvětlil, kdy vztah připouští

- a) nekonečně mnoho anamorfos algebraických i transcendentních,\*)
- b) jednu anamorfosu algebraickou,
- c) žádnou anamorfosu a konečně,
- d) kdy je podmíněně schopena anamorfosy.

Zcela obecné řešení problému anamorfosy podal GRONWALL (1912), které požaduje vyhledání společného integrálu dvou parciálních diferenciálních rovnic druhého řádu.

GRONWALL doprovodil své řešení úplnou diskusí případu

$$\frac{\partial^2 \log M}{\partial x \partial y} = 0.$$

Autor sdělení provedl diskusí případu  $\frac{\partial^2 \log M}{\partial x \partial y} \neq 0$ ;  $\frac{\partial C}{\partial y} = 0$ ,  $\frac{\partial D}{\partial x} = 0$  v článku Příspěvek k anamorfose, Věstník královské české společnosti nauk, IV (1949).

\*

Résumé. — Výtah.

### L'anamorphose en nomographie.

VÁCLAV PLESKOT, Praha.

Pour qu'une équation  $F_{123} = 0$  soit représentable par un nomogramme à points alignés il faut et il suffit qu'elle prenne la forme

$$\begin{vmatrix} f_1, g_1, h_1 \\ f_2, g_2, h_2 \\ f_3, g_3, h_3 \end{vmatrix} = 0. \quad (1)$$

La tâche très importante dans la nomographie consiste à trouver quelles sont les conditions pour qu'une équation soit capable de revêtir la forme (1) et ensuite, ces conditions étant remplies, si elle peut être représentée par des nomogrammes homographiquement distincts. Tel est le problème de l'anamorphose en nomographie.

\*) Anamorfosa je algebraická nebo transcendentní, jestli rovnice transformovaná obsahuje algebraické funkce případně transcendentní funkce funkcí  $f_n$ , kde  $f_n$  jsou funkce v rovnici dané. Na př. prepíšeme-li rovnici  $f_1 \cdot f_2 = f_3$  na  $\log f_1 + \log f_2 = \log f_3$ , říkáme, že byla podrobena anamorfose logaritmické a tedy transcendentní.

SORBAU a résolu ce problème pour les équations nomographiquement ordonnées, GRONWALL a résolu le cas général. Gronwall a discuté le cas où  $\frac{\partial^2 \log M}{\partial x \partial y} = 0$  et l'auteur de cette communication le cas où  $\frac{\partial^2 \log M}{\partial x \partial y} \neq 0$  et  $\frac{\partial C}{\partial y} = 0$   $\frac{\partial D}{\partial x} = 0$  dans son article Sur l'anamorphose, Věstník Královské české společnosti nauk, IV (1949).

## O DŁUGOŚCI LINIJ KRZYWYCH EMPIRYCZNYCH.

HUGO STEINHAUS, Wrocław.

Paradoks długości bierze się stąd, że długość jest funkcjonalem nieograniczonym w sąsiedztwie każdego łuku. Chcąc go usunąć opieramy się na takim sposobie mierzenia długości: Na płycie przezroczystej narysowana jest gromada linii równoległych o odstępach  $d$ ; tę płytkę kładziemy na łuku  $S$  i liczymy liczbę  $n_1$  przecięć, potem obracamy o  $1/k$  kąta półpełnego i znowu liczymy przecięcia; tak otrzymujemy liczby  $n_1, n_2, \dots, n_k$  i  $N = \sum_{i=1}^k n_i$ ; przybliżona długość łuku  $S$  jest

$$L = \frac{Nd\pi}{2k}, \quad (1)$$

a dokładność zależy od  $d$  i  $k$ . Wzór (1) wcale nie usuwa sam paradoksu długości, ale można go zmodyfikować biorąc zamiast liczb  $n_1, n_2, \dots, n_k$  liczby  $n'_1, n'_2, \dots$  określone przez

$$n'_i = n_i \text{ dla } n_i \leq p, \quad n'_i = p \text{ dla } n_i > p. \quad (2)$$

Tak dochodzimy przy użyciu (1) do *długości rzędu  $p$* , który to funkcjonal oznaczamy przez  $L_p(S)$ . Ten funkcjonal jest ograniczony. Można obliczyć wartość  $L_p(S)$  z dowolną dokładnością. Można porównywać długości granic dwóch obszarów ze sobą opierając się o wybrany rząd  $p$ , na przykład  $p = 10$ . Gdy dane są dwie mapy o różnych generalizacjach i chcemy wykorzystać zawarty w nich materiał informacyjny, n. p. porównać długość rzeki, która jest tylko na jednej mapie z długością innej, która jest tylko na drugiej mapie, musimy przystosować rząd  $p$  do mapy o grubszej generalizacji. Nawet gdy mamy do czynienia z obiektami, które mają w przyrodzie długość nieskończoną, możemy określić ich stosunek długości  $s$  przez

$$s = \lim_{p \rightarrow \infty} L_p(A)/L_p(B) \text{ dla } p \rightarrow \infty, \quad (3)$$

gdzie  $A, B$  są to owe objekty, n. p. lewy i prawy brzeg Wełtawy. Tam, gdzie nie chcemy użyć tego sposobu, lepiej w geografii unikać pojęcia długości.



**On the length of empirical curves.**

HUGO STEINHAUS, Wrocław.

A paradox in the concept of length arises from the fact that length is an unbounded functional in the neighbourhood of every arc. If we wish to eliminate this paradox, we can use the following method of measuring length: On transparent paper we draw a system of equidistant, parallel lines; distance of the lines is  $d$ . We place this paper over the arc  $S$  and count the number of intersections of the system with the curve  $S$ , then we turn the paper by the angle  $\pi/k$  and again we count the number of intersections; in this way we obtain the numbers  $n_1, n_2, \dots, n_k$  and  $N = \sum_{i=1}^k n_i$ ; the approximate length of the arc  $S$  is then

$$L = Nd \frac{\pi}{2k} \tag{1}$$

and the accuracy depends on  $d$  and  $k$ . Equation (1) alone does not eliminate the paradox of length, but can be modified by taking, in stead of the numbers  $n_1, n_2, \dots, n_k$ , the numbers  $n'_1, n'_2, \dots, n'_k$ , defined by the equations

$$n'_i = n_i \text{ for } n_i \leq p, \quad n'_i = p \text{ for } n_i > p. \tag{2}$$

Using (1) we obtain a *length of  $p^{\text{th}}$  order*, which functional we denote by  $L_p(S)$ . This functional is bounded. We can calculate the value  $L_p(S)$  with any chosen degree of accuracy. We can compare the lengths of the boundaries of two regions using any chosen order  $p$ , e. g.  $p = 10$ . If we are given two maps of different accuracy and if we wish to use the information contained in them, e. g. to compare the lengths of two rivers, one of which appears on the one map and the other on the other map, we must adjust the order  $p$  to the map of rougher accuracy. Even if we are investigating objects which in fact have infinite length, we can define the relation of their lengths  $s$  by the expression

$$s = \lim_{p \rightarrow \infty} L_p(A)/L_p(B), \tag{3}$$

where  $A, B$  are these objects, e. g. the left and the right bank of the River Vltava. In the cases where we do not wish to use this method, it is better to avoid the concept of length.

## STAVBA A PLÁNOVÁNÍ MATEMATICKÝCH STROJŮ.

ANTONÍN SVOBODA, Praha.

V oboru *matematických strojů analogických* dokončuje se v závodech Zbrojovky Brno n. p. prototyp *lineárního analyzátoru*, jehož hlavním úkolem má být řešení soustav lineárních rovnic.

V oboru *číslicových matematických strojů* pracuje se pokusně na *elektronkové násobičce* s elektrostatickou pamětí, sestrojen byl model *stolní elektromechanické kalkulačky* a vypracován byl plán *samočinného počítače* s magnetickou pamětí.

Lineární analyzátor je analogický matematický stroj, který řeší soustavu rovnic

$$\sum_k a_{ik}x_k + y_i = 0, \quad i, k = 1, 2, \dots, 10$$

tím, že vytvoří síť měničů proudu, kde každé rovnici odpovídá jeden měnič proudu, každé neznámé pak jeden samostatný obvod, do kterého zapojena jsou vinutí o počtu závitů úměrném součinitelům u téže neznámé. Konečně jest tu samostatný obvod, vytvořený vinutími o počtu závitů úměrném součinitelům  $y_i$ . Vinutí, odpovídající součiniteli  $a_{ik}$ , zapojeno je do  $k$ -tého obvodu a je umístěno na  $i$ -tém měniči proudu; vinutí, odpovídající součiniteli  $y_i$  jest na  $i$ -tém měniči.

Zavedeme-li do samostatného obvodu, vytvořeného vinutími součinitelů  $y_i$  střídavý proud  $I_0$ , vznikne v  $k$ -tém obvodu proud  $I_k$  a platí dostatečně přesně

$$\sum_k n_{ik}I_k + m_iI_0 = 0, \quad i, k = 1, 2, \dots, 10,$$

kde  $n_{ik} = c \cdot a_{ik}$ ,  $m_i = cy_i$ . Změříme-li proudy  $I_0, I_1, \dots, I_{10}$ , najdeme ihned

$$x_k = \frac{I_k}{I_0}.$$

Kompensací dosáhneme plné platnosti rovnice o nulovém součtu ampéřzávitů. Opakovaným použitím stroje vypočteme výsledky na libovolný počet platných číslic.

Elektronková násobička je typu seriového, kde číslo v binární formě zobrazeno je rytmem elektrických pulsů v jediném vodiči. Násobení provádí se postupným sečítáním. V přednášce vysvětlen byl podrobněji postup při sčítání. Čtyři rotační kapacitní paměti pracují při tomto pochodu. Jsou to v podstatě kolektory, jejichž lamely jsou přemostěny kondensátory. Otáčením kolektorů zapínají se do vnějšího obvodu postupně kondensátory v téže frekvenci, ve které přichází z vnějšího obvodu rytmus napěťových pulsů, zobrazující nějaké číslo. Časově rozvrstvený obraz čísla změní se tím v prostorový, neboť kondensátory nesou obraz dyadické číslice po celou otočku kolektoru. Kondensátory

vřazují se postupně do několika elektrických obvodů tak, aby součet jejich napětí zapůsobil na vstupní svorky elektronového čtyřpólu (model zhotoven byl Fysikálním výzkumem KOVO). Na výstupních svorkách čtyřpólu vznikne napětí, které je funkčně vázáno na vstupní napětí čtyřpólu a které nabíjí jiné kondensátory kolektorů. Výsledkem pracovního postupu, jak bylo podrobně vyloženo na přednášce, jest rytmus pulsů napětí, odpovídající dyadicky vyjádřenému součtu obou vstupujících sčítanců.

Stolní elektromechanická kalkulačka ukázána byla na exkursi účastníků sjezdu do závodu Zbrojovky Brno ve Vokovicích. Bylo na ní provedeno několik aritmetických operací. Současně předveden byl model rotační kapacitní paměti a nedokončený ještě model lineárního analyzátoru.

Samočinný počítač s magnetickou pamětí jest dosud pouze v plánu.

\*

Summary. -- Výtah.

### **Construction and planning of mathematical machines.**

ANTONÍN SVOBODA, Praha.

Several types of mathematical machines are under construction: a Linear Equation Solver (Linear Analyser) for systems of 10 equations, an Electronic Multiplier with electrostatic memory and an Electromechanical Calculator (table-type).

The *linear analyser* is an analogue mathematical machine formed by a network of transformers of currents. The well known physical relationship expressing that the sum of Ampereturns is zero at each instant and for each transformer of currents obtains the form of the given set of linear equations.

The *electronic multiplier* is a digital mathematical machine with a single digital channel including a system of electronic adders with electrostatic memories. The adders themselves consist of several memories operated by electronic quadripols.

The *memory* is a circular system of condensers connected to two collectors and rotated at uniform speed. The brushes of the collectors connect the condensers of the memory consecutively with external electric networks.

The *electronic quadripol* is a unit possessing a pair of input terminals and a pair of output terminals. The internal structure including electronic tubes establishes a definite relationship between the input and output voltages of the quadripol. The output voltage is zero if the input voltage remains below a certain well defined limit. Above this limit

the outputvoltage has a well stabilized (from inputvoltage independent) value.

To add two binary numbers stored in their rotating memories the voltages of two condensers representing the digits of the same order ( $n$ ) are added with the voltage of a third (carry) condenser and the resulting voltage acts as the input of a quadripol. The critical inputvoltage of the quadripol is set so that the outputvoltage corresponds to the carry of the order next higher ( $n + 1$ ); it is stored then in a (carry) condenser preceding the one which was used at the order  $n$ .

A laboratory model of an *electromechanical calculator* was presented during the excursion to the factory of Zbrojovka Brno at Vokovice.

## SAMPLING PLANS FOR PER CENT DEFECTIVE, WHICH MINIMIZE THE MAXIMUM OF A GIVEN RISK FUNCTION.

ANTONÍN ŠPAČEK, Praha.

Given a lot of size  $N$  of manufactured product which contains an *unknown* fraction defective  $p$ . The acceptance or rejection of this lot is to be decided on the basis of a single random sample of size  $n$  with acceptance number  $k$ . The decision procedure is determined as follows: *If the random sample of size  $n$  from this lot contains  $k$  or less defective items, the lot will be accepted, and otherwise rejected.* Evidently, according to this rule, the two integers  $n \leq N$  and  $k \leq n$  completely determine the decision procedure and with each  $N$  there is associated the set  $S_N$  of all decision procedures of this type, i. e. the set of all pairs of integers  $(n, k)$ , where  $0 \leq k \leq n \leq N$ . The choice of an element of  $S_N$  corresponds to the choice of a decision procedure. The probability of acceptance is

$$P(p, n, k) = \frac{1}{\binom{N}{n}} \sum_{j=0}^k \binom{Np}{j} \binom{N(1-p)}{n-j}. \quad (1)$$

If  $n/N$  is sufficiently small (for example  $\leq 1/10$ ), then (1) may be approximated by

$$L(p, n, k) = \sum_{j=0}^k \binom{n}{j} p^j (1-p)^{n-j}. \quad (2)$$

With each  $(n, k) \in S_N$  and each fraction defective  $p$  there is associated a risk  $R(p, n, k)$ , which expresses the economic outlay involved by the given decision procedure. If rejected lots are returned by the consumer to the producer, it seems natural to establish the risk-function as follows:  $cN + a(N - n)pL(p, n, k)$ , where  $c$  is the cost inspection of one item and  $a$  is the loss incurred by accepting one defective. Since  $cN$  is the cost

of making 100% inspection, we may write

$$R(p, n, k) = \frac{n}{N} + \alpha \left(1 - \frac{n}{N}\right) pL(p, n, k), \quad (3)$$

where  $\alpha = a/c$ .

A question of fundamental importance is that of the proper choice of sampling plans. Following an idea of A. Wald we shall now establish a principle, on the basis of which the sampling plan will be chosen.

*The decision procedure  $(v, \kappa) \in S_N$  is optimum if and only if simultaneously*

$$L(p_0, v, \kappa) \geq 1 - \varepsilon, \quad 0 < \varepsilon < 1, \quad (4)$$

(5) the maximum of (3) with respect to  $p$  is minimum for  $n = v, k = \kappa$ .

The condition (4) excludes the unreasonable consumers specifications, which cannot be met by the producer, i. e. there is given a  $p = p_0$  and a value  $0 < \varepsilon < 1$ , such that (4) holds. It is easy to show, that  $\min_n \max_p R(p, n, k)$  always exists and that

$$\min_n \max_p R(p, n, k) = \max_p \min_n R(p, n, k), \quad (6)$$

for each  $k = 0, 1, 2, \dots$ . Clearly, (6) corresponds to the condition of strict determinateness in the sense of von Neumann's theory of zero-sum two-person games.

For each  $k = 0, 1, 2, \dots$  let  $n_k$  denote the value of  $n$  which minimizes the maximum with respect to  $p$  of (3). We have  $n_0 \approx \sqrt{Nde^{-1}}$  and  $n_k$  for  $k = 1, 2, 3, \dots$  may be easily computed by successive approximations. Since  $R(p, n, k) \leq R(p, n, k + 1)$  for each  $p, n, k$ , we have

$$\min_n \max_p R(p, n, k) \leq \min_n \max_p R(p, n, k + 1) \quad \text{for each } k = 0, 1, 2, \dots,$$

and the optimum decision procedure  $(v, \kappa)$  can be found as follows:

For  $k = 0, 1, 2, \dots$  we compute  $n_0, n_1, n_2, \dots$  as long as

$$L(p_0, n_k, k) < 1 - \varepsilon.$$

Then

$$L(p_0, n_{k+1}, k + 1) \geq 1 - \varepsilon$$

and  $(v, \kappa) \in S_N$ , where  $v = n_{k+1}, \kappa = k + 1$  is the optimum decision procedure. If  $v/N$  is not sufficiently small, then (2) must be replaced by (1) and the computations are obviously much more laborious. Generalisation of the problem may be obtained by replacing  $n$  in (3) by the average sample number. It is further possible to consider double multiple or sequential sampling schemes.

Tesla Electronic National Corporation, Prague.

**Výběrové plány pro procento vadných výrobků, pro které maximum dané risikové funkce je minimální.**

ANTONÍN ŠPAČEK, Praha.

Jedná se o nalezení takového výběrového plánu pro procento vadných výrobků, který splňuje technické možnosti výrobce a je s ohledem na risiko spotřebitele hospodářsky nejvýhodnější.

**O OBLICZANIU PÓL OBSZARÓW PŁASKICH ZA POMOCĄ SIATEK RÓWNOLEGŁOBCZNYCH.**

MIECZYŚLAW WARMUS, Wrocław.

Dana jest na płaszczyźnie siatka  $W$ , której oczka są kwadratami o boku  $l$ , oraz obszar  $A$  o polu  $a$  i średnicy  $b$ , ograniczony krzywą pojedynczą rektyfikowalną i zamkniętą o długości  $l$ . W obszarze  $A$  leży dokładnie  $w$  węzłów siatki  $W$ . Określona jest ponadto liczba  $\varphi$  równaniem

$$\varphi + \sin\varphi \cos\varphi - 2\sin^2\varphi = 0 \text{ z warunkiem } 0 < \varphi < \frac{1}{2}\pi. \quad (1)$$

Położmy  $c = \frac{1}{2 \sin\varphi}$ . Wartości przybliżone liczb  $\varphi$  i  $c$  są następujące:

$$\varphi = 1,0311579 \dots, \quad c = 0,5828222 \dots$$

**Twierdzenie 1.** —  $c \cdot l - 1 < a - w \leq c \cdot l - 1$ , gdy  $b \geq 1$ .

Niech  $W^*$  będzie obrazem afinicznym siatki  $W$ , a więc siatką płaską, której oczka są równoległobokami. Niech pole oczka siatki  $W^*$  wynosi 1, a odległość między dwoma najbliższymi węzłami tej siatki  $m$ . Niech  $\varphi^*$  będzie liczbą określoną równaniem

$$m^2(\varphi^* + \sin\varphi^* \cos\varphi^*) - 2 \sin^2\varphi^* = 0 \text{ z warunkiem } 0 < \varphi^* < \frac{1}{2}\pi. \quad (2)$$

Położmy

$$c^* = \frac{1}{\sqrt{2\varphi^* + \sin 2\varphi^*}} \quad (3)$$

i oznaczymy przez  $w$  ilość węzłów siatki  $W^*$  leżących wewnątrz tego samego obszaru  $A$ .

**Twierdzenie 2.** —  $c^* \cdot l - 1 < a - w^* \leq c^* \cdot l - 1$ , gdy  $b \geq m$ .

**Twierdzenie 3.** Funkcja  $c^* = c^*[\varphi^*(m)]$ , określona wzorami (2) i (3) osiąga minimum, gdy oczka siatki  $W^*$  są rombami powstałymi z połączenia dwóch przystających trójkątów równobocznych.

Jest wtedy  $c^*_{\min} = 0,568575873599 \dots$

Taka siatka minimalna daje przy oszacowaniu pola  $a$  najmniejszy błąd maksymalny (dla zbioru wszystkich możliwych obszarów  $A$ ) ze wszystkich siatek równoległobocznych.

**Twierdzenie 4. Nierówność** —  $c_1 \cdot l + c_2 \leq a - w^* \leq c_3 \cdot l + c_4$ ,  
 gdzie  $c_1, c_2, c_3$  i  $c_4$  są liczbami zależnymi jedynie od kształtu oczek siatki  $W^*$ ,  
 pociąga za sobą nierówności  $c_1 \geq c^*$  i  $c_3 \geq c^*$ .

(Praca ukaże się w Pracach Wrocławskiego Towarzystwa Naukowego, seria B).

\*

Résumé. — Streszczenie.

## Sur l'évaluation des aires des régions planes à l'aide des réseaux de parallélogrammes.

MIECZYŚLAW WARMUS, Wrocław.

Soit  $W$  un réseau plan aux mailles quadratiques de côté 1. Soit, sur le même plan,  $A$  une région d'aire  $a$  et de diamètre  $b$ , bornée par une courbe simple fermée rectifiable de longueur  $l$ . La région  $A$  contient exactement  $w$  noeuds du réseau  $W$ . En outre, un nombre  $\varphi$  est défini par l'équation

$$\varphi + \sin\varphi \cos\varphi - 2 \sin^2\varphi = 0 \quad \text{où } 0 < \varphi < \frac{1}{2}\pi. \quad (1)$$

Posons  $c = \frac{1}{2 \sin\varphi}$ . Les valeurs approchées de  $\varphi$  et  $c$  sont les suivantes:  
 $\varphi = 1,0311579 \dots$ ,  $c = 0,5828222 \dots$

**Théorème 1.** —  $c \cdot l - 1 < a - w \leq c \cdot l - 1$  pour  $b \geq 1$ .

Soit  $W^*$  une image affine du réseau  $W$ , à savoir un réseau plan aux mailles en forme de parallélogrammes d'aire 1, et dont la distance entre deux noeuds les plus proches l'un de l'autre est  $m$ . Soit  $\varphi^*$  le nombre défini par l'équation

$$m^2(\varphi^* + \sin\varphi^* \cos\varphi^*) - 2 \sin^2\varphi^* = 0, \quad \text{où } 0 < \varphi^* < \frac{1}{2}\pi. \quad (2)$$

Posons

$$c^* = \frac{1}{\sqrt{2\varphi^* + \sin 2\varphi^*}} \quad (3)$$

et désignons par  $w^*$  le nombre de noeuds du réseau  $W^*$  situés dans la même région  $A$ .

**Théorème 2.** —  $c^* \cdot l - 1 < a - w^* \leq c^* \cdot l - 1$  pour  $b \geq m$ .

**Théorème 3.** La fonction  $c^* = c^*[\varphi^*(m)]$ , définie par les formules (2) et (3), atteint son minimum lorsque les mailles du réseau  $W^*$  sont des rhombes formés de couples de triangles équilatéraux.

On a alors  $c^*_{\min} = 0,568575873599 \dots$

Ce réseau minimal donne, pour l'évaluation de l'aire  $a$ , la plus petite erreur maximum (en tenant compte de toutes les régions  $A$  possibles) parmi tous les réseaux  $W^*$ .

**Théorème 4.** L'inégalité —  $c_1 \cdot l + c_2 \leq a - w^* \leq c_3 \cdot l + c_4$ , où  $c_1, c_2, c_3$  et  $c_4$  sont des nombres dépendant uniquement de la forme des mailles du réseau  $W^*$ , entraîne les inégalités  $c_1 \geq c^*$  et  $c_3 \geq c^*$ .