A limit theorem for a modified Bernoulli scheme

by

M. FISZ (Warszawa)

1. An essential role in the considerations of this paper is played by the known theorem of Kolmogorov [2], stating that the probability function P in the space $(x_1, x_2, x_3, \ldots)$ is uniquely determined by the set of all finite dimensional marginal probability functions $P_{l_1, l_2, \ldots, l_s}$ of P defined for all Borel sets in the corresponding s-dimensional space $(x_{l_1}, x_{l_2}, \ldots, x_{l_s})$ for $s = 1, 2, 3, \ldots$ and for arbitrary integers $l_1, l_2, \ldots, l_s$.

DEFINITION 1. We shall say that the denumerably dimensional probability function P, given in the space $(x_1, x_2, x_3, \ldots)$, *is not singular* if for $s = 1, 2, 3, \ldots$ and for arbitrary integers $l_1, l_2, \ldots, l_s$ the marginal probability function $P_{l_1, l_2, \ldots, l_s}$ is not singular[1]) in the usual sense.

DEFINITION 2. We shall say that the sequence P_n of denumerably dimensional probability functions in the space $(x_1, x_2, x_3, \ldots)$ *converges, as* $n \to \infty$, *to a probability function* P if for $s = 1, 2, \ldots$ and arbitrary integers $l_1, l_2, \ldots, l_s$ the sequence $P_{n(l_1, l_2, \ldots, l_s)}$ of s-dimensional marginal probability functions of P_n converges, as $n \to \infty$, to the corresponding marginal probability function $P_{l_1, l_2, \ldots, l_s}$ of the limiting probability function P.

DEFINITION 3. The non-singular probability function P in the space $(x_1, x_2, \ldots, x_s)$ is of the *Poisson-normal type* if it is a probability function of a variable (ξ, η) where ξ is a j-dimensional Poisson[2]) variable, η an $(s-j)$-dimensional normal variable $(0 \leqslant j \leqslant s)$, ξ and η being independent.

DEFINITION 3a. We shall say that the non-singular denumerably dimensional probability function P in the space $(x_1, x_2, x_3, \ldots)$ is of the *Poisson-normal type* if each finite dimensional marginal probability function of P is of the Poisson-normal type.

[1]) The probability function P, defined in the space $(x_1, x_2, \ldots, x_s)$, is called *singular* if the whole mass of probability lies in a z-dimensional hyperplane, where $z < s$.

[2]) A j-dimensional random variable $(y_1, y_2, \ldots, y_j)$ is called a *Poisson variable* if its probability function is given by the formula

$$P(y_1 = A_1 k_1 + B_1, y_2 = A_2 k_2 + B_2, \ldots, y_j = A_j k_j + B_j) = \prod_{m=1}^{j} e^{-\lambda_m} \lambda_m^{k_m} / k_m!,$$

where $k_m = 0, 1, 2, \ldots$ and $\lambda_m > 0$, $A_m \neq 0$ and B_m are real constants.

Let us now consider the multinomial distribution given by the formula

$$(1) \qquad P_n(x_1 = k_1, x_2 = k_2, \ldots, x_r = k_r) = \frac{n!}{k_1! k_2! \ldots k_r!} p_{n1}^{k_1} p_{n2}^{k_2} \ldots p_{nr}^{k_r}$$

where $0 < p_{nm} < 1$, p_{nm} are arbitrary functions of n and the k_m $(m = 1, 2, \ldots, r)$ are non-negative integers satisfying the equality

$$\sum_{m=1}^{r} k_m = n.$$

The variable $(x_1, x_2, \ldots, x_r)$ can be reduced with probability 1 — in view of the last equality — to an $(r-1)$-dimensional variable.

The following theorem has been proved by the author [1]:

THEOREM 1. *Let the random variable* $(x_1, x_2, \ldots, x_r)$ *be distributed according to* (1) *and let the sequence* G_n *of probability functions of the random variables*

$$(A_{n1} x_1 + B_{n1}, A_{n2} x_2 + B_{n2}, \ldots, A_{nr} x_r + B_{nr}),$$

where $A_{nm} \neq 0$ *and* B_{nm} $(n = 1, 2, 3, \ldots;\ m = 1, 2, \ldots, r)$ *are real numbers, converge, as* $n \to \infty$, *to a non-singular* $(r-1)$*-dimensional probability function* G. *Then* G *is necessarily of the Poisson normal type.*

Let us now modify the multinomial distribution given by (1). Namely, let us suppose that the number r in (1) — which we shall denote by r_n — is a non-decreasing function of n, increasing to infinity, as $n \to \infty$. The following theorem answers a question put to the author by G. Hajos:

THEOREM 1a. *Let the sequence* G_n *of probability functions of the random variables*

$$(A_{n1} x_1 + B_{n1}, A_{n2} x_2 + B_{n2}, \ldots, A_{nr_n} x_{r_n} + B_{nr_n}),$$

where $A_{nm} \neq 0$ *and* B_{nm} $(n = 1, 2, 3, \ldots;\ m = 1, 2, \ldots, r_n)$ *are real constants, converge, as* $n \to \infty$ *and* $r_n \to \infty$, *to a non-singular denumerably dimensional probability function* G. *Then* G *is necessarily of the Poisson-normal type.*

Proof. Let the assumptions of theorem 1a be satisfied. Thus the sequences of arbitrary s-dimensional $(s = 1, 2, 3, \ldots)$ marginal probability functions of G_n converge, as $n \to \infty$, to the corresponding marginal probability function of the limiting probability function G. However, from theorem 1 it follows that arbitrary marginal probability functions of G are of the Poisson-normal type. Thus, taking into account definition 3a, we obtain the assertion of theorem 1a.

Let us now observe that, the assumption of theorem 1a being satisfied, one can assume[3]) that the set $M=\{1,2,3,\ldots\}$ of indices m can be divided into two subsets $M_1=\{m_{1_1},m_{1_2},\ldots\}$ and $M_2=\{m_{2_1},m_{2_2},\ldots\}$ in such a way that for $m\in M_1$ the relation

$$\lim_{n\to\infty} np_{nm}=\lambda_m, \tag{2}$$

where $0<\lambda_m<\infty$, holds, and for $m\in M_2$ the relations

$$\lim_{n\to\infty} p_{nm}=p_m, \quad \lim_{n\to\infty} np_{nm}=\infty \tag{3}$$

hold. Let us introduce new variables

$$\begin{gathered} y_m=x_m \qquad (m\in M_1), \\ y_m=(x_m-np_{nm})/\sqrt{np_{nm}} \qquad (m\in M_2). \end{gathered} \tag{4}$$

Let

$$\varphi_{ns}=\varphi_{ns}(t_{m_1},\ldots,t_{m_j},t_{m_{j+1}},\ldots,t_{m_s})$$

denote the characteristic function of the random variable $(y_{m_1},\ldots,y_{m_j},y_{m_{j+1}},\ldots,y_{m_s})$, where s is an arbitrary constant integer and $m_v\in M_1$ for $v=1,2,3,\ldots,j$ and $m_v\in M_2$ for $v=j+1,\ldots,s$. As has been shown in the cited paper [1], it follows from relations (2) and (3) the following one:

$$\lim_{n\to\infty}\log\varphi_{ns}=\sum_{v=1}^{j}\lambda_{m_v}(e^{it m_v}-1)-\\ -0{,}5\Big[\sum_{v=j+1}^{s}(1-p_{m_v})t_{m_v}^2-2\sum_{v=j+1}^{s-1}\sum_{w=v+1}^{s}\sqrt{p_{m_v}p_{m_w}}\,t_{m_v}t_{m_w}\Big]. \tag{5}$$

On the right side of (5) we have the logarithm of the characteristic function of a non-singular $(s-1)$-dimensional probability function of the Poisson-normal type. It is easily seen that the whole set of finite dimensional distribution functions of the Poisson-normal type given on the right side of (5) satisfies Kolmogorov's [2] consistency conditions, thus it determines a denumerably dimensional probability function of the Poisson-normal type. We have obtained the following conclusion:

If for some constants A_{nm} and B_{nm} the assumption of theorem 1a *holds then it holds also for the norming of* $(x_1,x_2,x_3,\ldots)$ *given by* (4).

[3]) Following a method of Cantor, we can choose a subsequence n_α for which the relations (2) and (3) hold. As one can consider only this subsequence, the assumptions (2) and (3) do not restrict the generality of our considerations. The case $\lambda_m=0$ is — as has been shown in [1] — excluded by the assumed non-singularity of the limiting probability function.

Let us now consider the special case with equal probabilities p_{nm}, *i. e.* for each value of n let

$$p_{nm}=\frac{1}{r_n} \qquad (m=1,2,\ldots,r_n).$$

Then we conclude from (2) and (3) that if we use the normation (4), the limiting probability distribution will be of the Poisson type if $r_n=O(n)$ and of the normal type if $r_n=o(n)$. The conclusions obtained are in accordance with intuition.

Thus if we use the χ^2-test in a very large sample and divide the sample into classes with equal theoretical probabilities[4]), the number of classes should be of order $o(n)$.

The theorem 1a can have the following physical interpretation. Let n denote the number of particles of gas contained in a vessel V divided into r_n parts $V_1,V_2,\ldots,V_{r_n}$. Let us imagine a machine pressing the gas into the vessel V and at the same time dividing this vessel into more and more parts so that one can assume that n and r_n converge to infinity. Let the joint probabilities of finding k_m $(m=1,2,\ldots,r_n)$ particles of gas in V_m be given by (1). Then if a limiting probability distribution assumed in theorem 1a exists then 1° this limiting distribution must be of the Poisson-normal type, 2° the rapidity of the convergence of r_n to infinity is restricted by the rapidity of the convergence of n to infinity.

[4]) Comp. the paper of Mann and Wald [3].

References

[1] M. Fisz, *The limiting distributions of the multinomial distribution*, Studia Mathematica 14 (1954), p. 272-275.

[2] A. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933.

[3] B. Mann and A. Wald, *On the choice of the number of class intervals in the application of the chi-square test*, Annals of Mathematical Statistics 13 (1942), p. 306-317

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK
MATHEMATICAL INSTITUTE OF THE POLISH ACADEMY OF SCIENCES

Reçu par la Rédaction le 15. 9. 1954