# The limiting distributions of the multinomial distribution

by

M. FISZ (Warszawa)

Let us consider a Bernoulli schema with $r$ possible outcomes ($r \geqslant 2$) and let the probability $p_{nm}$ ($m=1,\ldots,r$) of the occurence of the $m$-th outcome depend on the number $n$ of observations. We suppose that $0<p_{nm}<1$ for all $n$ and $m$. Let $x_{nm}$ denote the number of occurrences of the $m$-th outcome in $n$ observations. Then

$$P(x_{n1}=k_1, x_{n2}=k_2, \ldots, x_{nr}=k_r) = \frac{n!}{k_1!\ldots k_r!} p_{n1}^{k_1} p_{n2}^{k_2} \ldots p_{nr}^{k_r}, \tag{1}$$

where $k_m$ are non-negative integers satisfying the equality

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

In view of the last relation the variable $(x_{n1}, x_{n2}, \ldots, x_{nr})$ can be[1]) reduced with probability 1 to an $(r-1)$-dimensional variable. The formula (1) is the distribution function of the multinomial distribution.

The aim of this paper is to find the class of all possible limiting distributions of (1) where the $p$-s may be arbitrary functions of $n$[2]). The solution of this question for $r=2$ has been given by Kozuliajev [2].

The following theorem will be proved:

THEOREM 1. *Let the sequence of distribution functions of the random variables $(A_{n1}x_{n1}+B_{n1}, A_{n2}x_{n2}+B_{n2}, \ldots, A_{nr}x_{nr}+B_{nr})$, where $A_{nm}>0$ and $B_{nm}$ $(m=1,\ldots,r)$ are some sequences of real numbers, converge as $n\to\infty$ to a distribution function of a non-singular $(r-1)$-dimensional variable. Then this variable is necessarily of the form $(\xi,\eta)$, where $\xi$ is a $j$-dimensional Poisson variable[3]) and $\eta$ is an $(r-j-1)$-dimensional normal variable $(0\leqslant j\leqslant r-1)$, $\xi$ and $\eta$ being independent.*

---

[1]) Cf. [1], § 22.5.

[2]) The problem was suggested to me by A. N. Kolmogorov during the VIII Congress of Polish Mathematicians at Warsaw.



Proof. The characteristic function of (1) is given by the formula

$$\varphi(t_1, t_2, \ldots, t_r) = (p_{n1}e^{it_1} + p_{n2}e^{it_2} + \ldots + p_{nr}e^{it_r})^n. \tag{2}$$

Let us suppose that $p_{nm_1}, p_{nm_2}, \ldots, p_{nm_j}$ satisfy the relation

$$\lim_{n\to\infty} np_{nm} = \lambda_m, \tag{3}$$

where $0<\lambda_m<\infty$. Clearly $0\leqslant j\leqslant r-1$. Without restricting the generality of our considerations we can assume $m_k=k$ $(k=1,\ldots,j)$. Let the remaining $np_{nm}$ satisfy the relations

$$\lim_{n\to\infty} p_{nm} = p_m, \qquad \lim_{n\to\infty} np_{nm} = \infty. \tag{4}$$

We shall introduce new variables

$$\begin{aligned} y_{nm} &= x_{nm} && (m=1,\ldots,j), \\ y_{nm} &= \frac{x_{nm}-np_{nm}}{\sqrt{np_{nm}}} && (m=j+1,\ldots,r). \end{aligned} \tag{5}$$

The characteristic function of the variable $(y_{n1}, y_{n2}, \ldots, y_{nr})$ is given by the formula

$$\varphi_1 = \varphi_1(t_1, t_2, \ldots, t_r)$$

$$= e^{-i\sqrt{n}\sum\limits_{m=j+1}\sqrt{p_{nm}}t_m}\left[p_{n1}e^{it_1}+\ldots+p_{nj}e^{it_j}+p_{n(j+1)}e^{\frac{it_{j+1}}{\sqrt{np_{n(j+1)}}}}+\ldots+p_{nr}e^{\frac{it_r}{\sqrt{np_{nr}}}}\right]^n.$$

Taking into account (3) and (4) we obtain

$$\log\varphi_1 = -i\sqrt{n}\sum_{m=j+1}^{r}\sqrt{p_{nm}}t_m + n\log\Big[1+p_{n_1}(e^{it_1}-1)+\ldots+p_{nj}(e^{it_j}-1)$$

$$+ i\sum_{m=j+1}^{r}\sqrt{\frac{p_{nm}}{n}}\,t_m - \sum_{m=j+1}^{r}\frac{t_m^2}{2n} + o\left(\frac{1}{n}\right)\Big]$$

$$= \sum_{m=1}^{j} np_{nm}(e^{it_m}-1) - \sum_{m=j+1}^{r}\frac{t_m^2}{2} + \frac{1}{2}\sum_{m=j+1}^{r} p_{nm}t_m^2$$

$$+ \sum_{m=j+1}^{r-1}\sum_{l=m+1}^{r}\sqrt{p_{nm}p_{nl}}\,t_m t_l + o(1),$$

---

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

$$P(y_1 = A_1k_1+B_1, y_2 = A_2k_2+B_2, \ldots, y_j = A_jk_j+B_j) = \prod_{m=1}^{j} e^{-\lambda_m}\frac{\lambda_m^{k_m}}{k_m!},$$

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

and finally

$$\lim_{n\to\infty}\log\varphi_1=\sum_{m=1}^{j}\lambda_m(e^{it_m}-1)$$

(6)
$$-\frac{1}{2}\left[\sum_{m=j+1}^{r}(1-p_m)t_m^2-2\sum_{m=j+1}^{r-1}\sum_{l=m+1}^{r}\sqrt{p_m p_l}\,t_m t_l\right].$$

The first expression on the right side of (6) is the logarithm of a characteristic function of a $j$-dimensional Poisson variable, the second one is the logarithm of a characteristic function of an $(r-j-1)$-dimensional normal variable, since the rank of the matrix[4]) of the quadratic form in the squared brackets is equal to $r-j-1$.

Let us now suppose that the assumption of the theorem is satisfied. Following the method of Cantor we can choose such a subsequence of indices $n_a$ that $p_{n_a m}\to p_m$ for $m=1,\ldots,r$ and $np_{n_a m}\to\lambda_m$, where $0\leqslant p_m\leqslant 1$ and $0\leqslant\lambda_m\leqslant\infty$. Without restricting the generality of our considerations we assume that the same holds for the sequences of indices $n$.

Let us assume that for each value of $m=1,2,\ldots,r$ relation (3) or (4) holds. Then the assumption of our theorem, formula (6) and a generalization of a theorem of Khintchin (see [3]) imply that in the limit, as $n\to\infty$, only a distribution of a variable of the form required in the assertion of the theorem can be obtained.

Let us now suppose that the assumption of the theorem is satisfied and that at least for one of the $p_{nm}$-s, say for $p_{n1}$, the following relation holds:

(7)
$$\lim_{n\to\infty} np_{n1}=0.$$

Let us introduce new variables:

$$z_{nm}=A_{nm}x_{nm}+B_{nm}\qquad (m=1,2,\ldots,r).$$

For the characteristic function $\varphi_2(t_1,\ldots,t_r)$ of the variable $(z_{n1},\ldots,z_{nr})$ the relation

(8)
$$\log\varphi_2=i\sum_{m=1}^{r}B_{nm}t_m+n\log\Big[1+\sum_{m=1}^{r}p_{nm}(e^{iA_{nm}t_m}-1)\Big]$$

holds. Further, for some $m$, let $0<p_m<1$. We have

$$\log\varphi_2(0,0,\ldots,t_m,0,\ldots,0)=iB_{nm}t_m+n\log[1+p_{nm}(e^{iA_{nm}t_m}-1)].$$

Since $\log\varphi_2$ converges with $n\to\infty$ to a characteristic function and $p_m\neq 0$, it follows from the last equality that $A_{nm}\to A_m=0$. The logarithm in formula (8) can thus be expanded in the neighbourhood of the number 1. To the same conclusion leads the assumption that for some $m$, say $m=r$, $p_r=1$. This is immediately seen if we take in (8) $t_r=0$, since for $m=1,2,\ldots,r-1$ we have $p_{nm}\to 0$. Then it follows from formula (7) that the term depending on $t_1$ will vanish. On the other hand $B_{n1}\to B_1$, where $B_1$ is finite. Then we obtain in the limit a distribution of a singular variable whose number of dimensions is at most $r-2$.

---

[4]) Loc. cit. [1]).



The theorem is thus completely proved.

Let us finally remark that the assertion of the theorem remains true if the assumption $A_{nm}>0$ is replaced by $A_{nm}\neq 0$.

### References

[1] H. Cramér, *Mathematical methods of statistics*, Princeton 1946.

[2] П. Козуляев, *Асимптотический анализ одной основной формулы теории вероятностей*, Ученые Записки Московского Университета 15 (1939), p. 179-180.

[3] M. Fisz, *A generalization of a theorem of Khintchin*, Studia Mathematica, this volume, p. 310-313.

*(Reçu par la Rédaction le 10. 10. 1953)*