

## ON TYPES OF LIMIT DISTRIBUTIONS FOR SOME CLASS OF STATISTICS

R. Mukhamedkhanova

For statistical applications it is necessary to investigate the limit distributions of the functions of statistics based on a random sample from a general population (cf. [1], [2], [3]). All of these authors consider this problem under the condition that the distribution function of the population random variable does not depend on the sample size  $N$ .

For the first time Zubkov ([6]) has given sufficient conditions for the convergence to the normal and  $\chi^2$  distributions, in the case of a concrete function of the parameters (statistical estimation of entropy in the multinomial scheme), under the assumption that the sampling probabilities depend on  $N$ .

We shall now determine the class of all the limit distributions of the distribution of a function  $H(\cdot)$  depending on the frequencies of the occurrences of outcomes in a multinomial scheme when the probabilities of the outcomes may depend on  $N$ .

Let  $\hat{p}_i$ ,  $i = 1, \dots, s$ , be the frequencies of the occurrences of the outcomes in a multinomial scheme,  $p_i = p_i(N)$  the true probabilities of the outcomes satisfying the conditions: the limits in (1) and (2) exist and have the following properties

$$(1) \quad \lim_{N \rightarrow \infty} p_i(N) = p_i^0, \quad p_1^0 \geq \dots \geq p_s^0, \quad i = 1, \dots, s,$$

$$(2) \quad \lim_{N \rightarrow \infty} N p_i(N), \quad \lim_{N \rightarrow \infty} N p_2 \geq \dots \geq \lim_{N \rightarrow \infty} N p_s, \quad i = 1, \dots, s.$$

Let us consider a function  $H(\cdot)$  on an  $s$ -dimensional simplex  $S$  and let us define the function  $\tilde{H}(\cdot)$  by

$$(3) \quad \begin{cases} \tilde{H}(\bar{y}) = H(\bar{x}) & \text{for } \bar{x} = (x_1, \dots, x_s), \\ x_1 = 1 - \sum_{i=2}^s x_i, & \bar{y} = (x_2, \dots, x_s). \end{cases}$$

Let us assume that  $\tilde{H}(\cdot)$  satisfies the following conditions:

$$(4) \quad \begin{cases} \text{the derivatives } \frac{\partial \tilde{H}(\bar{y})}{\partial x_{i+1}} = H_i(\bar{y}), \quad \frac{\partial^2 \tilde{H}(\bar{y})}{\partial x_{i+1} \partial x_{j+1}} = H_{ij}(\bar{y}), \\ i, j = 1, \dots, s-1, \text{ exist and are continuous,} \end{cases}$$

$$(5) \quad \begin{cases} \text{if for some } k = 1, \dots, s-1, & H_i(\bar{y}) = 0, \quad i = 1, \dots, k, \\ \text{then } H_{ii}(\bar{y}) \neq 0, \quad i = 1, \dots, k. \end{cases}$$

**Definition 1.** We shall say that a function  $H(\cdot)$  belongs to the class  $\mathcal{K}_s$  if the function  $\tilde{H}(\cdot)$  satisfies the conditions (4) and (5).

Below we shall consider the following statistics:  $\hat{H} = H(\hat{p}_1, \dots, \hat{p}_s)$ , where  $\hat{p} = (\hat{p}_1, \dots, \hat{p}_s)$  is the vector of the observed frequencies in the multinomial scheme under the conditions (1), (2) and  $H(\cdot) \in \mathcal{K}_s$ .

Let us denote

$$(6) \quad \begin{aligned} \hat{H} &= H(\bar{p}), \quad H_i = H_i(\bar{q}), \quad i = 1, \dots, s-1, \quad \bar{q} = (p_2, \dots, p_s) \\ H_i^0 &= H_i(\bar{q}^0), \quad \bar{q}^0 = (p_2^0, \dots, p_s^0). \end{aligned}$$

Let  $\Phi_\sigma(\cdot)$  be the distribution function of the normal law with the parameters  $(0, \sigma^2)$ .

Let

$$\tilde{\chi}_n^2(x; a, \alpha) = \begin{cases} \chi_n^2(a^{-1}x; \alpha), & a > 0, \\ 1 - \chi_n^2(a^{-1}x; \alpha), & a < 0, \end{cases}$$

where  $\chi_n^2(\cdot; \alpha)$  is the  $\chi^2$ -distribution function with  $n$  degrees of freedom (for  $\alpha = 0$  it is central, for  $\alpha \neq 0$  it is non-central),

$$\tilde{\Pi}(x; b, \lambda) = \begin{cases} \Pi_\lambda(b^{-1}x), & \text{if } b > 0, \\ 1 - \Pi_\lambda(b^{-1}x + 0), & \text{if } b < 0, \end{cases}$$

where  $\Pi_\lambda(\cdot)$  is the distribution function of a Poisson random variable with parameter  $\lambda$ , and

$$\tilde{\Pi}^2(x; m, \mathbf{B}, \bar{\mu}) = P\{\bar{\eta}\mathbf{B}\bar{\eta}' < x\}, \quad m \geq 1,$$

where  $\eta = (\eta_1, \dots, \eta_m)$  is a vector of independent Poisson distributed random variables with the parameters  $\mu_i$ ,  $i = 1, \dots, m$ , and  $\mathbf{B}$  is a symmetric  $m \times m$  matrix. Finally, we denote

$$\Phi_0 \equiv E(x), \quad \varphi(z) = \begin{cases} 1, & z \neq 0, \\ 0, & z = 0, \end{cases}$$

i.e.,  $\Phi_0$  is the degenerate distribution function at the origin.

**Definition 2.** Let us define  $\mathcal{F}_s$  as the class of following distribution functions:

$$\begin{aligned} &\Phi_\sigma(x+c), \quad \Phi_{\sigma_1} * \tilde{\chi}_{n_1}^2(\cdot; a_1, \alpha_1) * \dots * \tilde{\chi}_{n_{k_1}}^2(x+c; a_{k_1}, \alpha_{k_1}), \\ &\Phi_{\sigma_2} * \tilde{\Pi}(\cdot; b_1, \mu_1) * \dots * \tilde{\Pi}(x+c; b_{l_1}, \mu_{l_1}), \\ &\Phi_{\sigma_3} * \tilde{\chi}_{n_1}^2(\cdot; a_1, \alpha_1) * \dots * \tilde{\chi}_{n_{k_2}}^2(\cdot; a_{k_2}, \alpha_{k_2}) * \tilde{\Pi}(\cdot; b_1, \mu_1) * \dots * \tilde{\Pi}(x+c; b_{l_2}, \mu_{l_2}), \\ &\tilde{\Pi}^2(x+c; m, \mathbf{B}, \bar{\mu}), \quad \sigma > 0, \quad \sigma_1 \geq 0, \quad k_i \geq 1, \quad l_i \geq 1, \end{aligned}$$

under the following conditions

$$\begin{aligned} \varphi(\sigma_1) + n_1 + \dots + n_{k_1} &\leq s - 1, \quad \varphi(\sigma_2) + l_1 \leq s - 1, \\ \varphi(\sigma_3) + n_1 + \dots + n_{k_2} + l_2 &\leq s - 1. \end{aligned}$$

**Theorem 1.**  $\mathcal{F}_s$  is the class of all non-degenerate limit distribution functions for  $C_N(\hat{H} - d_N)$ , where  $\hat{H} = H(\hat{p}_1, \dots, \hat{p}_s)$ ,  $H(\cdot) \in \mathcal{K}_s$ ,  $C_N$  and  $d_N$  are normalizing constants.

**Remark 1.** For specific functions in  $\mathcal{F}_s$  necessary and sufficient conditions for the convergence may be obtained in terms of the following characteristics:

$$\begin{aligned} u &= \max\{i = 0, \dots, s - 1 : p_{1+i}^0 > 0\}, \\ v &= \max\{i = 0, \dots, s - 1 : Np_{1+i} \rightarrow \infty\}, \\ w &= \max\{i = 0, \dots, s - v - 1 : \lim Np_{1+v+i} > 0\}, \\ t &= \begin{cases} \min\{i = 0, \dots, s - 1 : H_i^0 \neq 0\}, & \text{if } (H_1^0)^2 + \dots + (H_{s-1}^0)^2 \neq 0, \\ s, & \text{if } H_1^0 = \dots = H_{s-1}^0 = 0. \end{cases} \end{aligned}$$

As an example let us formulate the following theorem about the convergence to the normal law.

**Theorem 2.** For  $\hat{H} = H(\hat{p}_1, \dots, \hat{p}_s)$ , with  $H(\cdot) \in \mathcal{K}_s$ , the conditions

- 1)  $v > 0$ , and
- 2)  $\sqrt{N}\sigma_N \rightarrow \infty$  when  $t \leq v + w$ ,

$$\sqrt{N}\sigma_N/\rho_N \rightarrow \infty \quad \text{when } t > v + w,$$

are necessary and sufficient for the convergence

$$P\{C_N(\hat{H} - d_N) < x\} \implies \Phi_1(x+c).$$

Here  $C_N \sim \sqrt{N}/\sigma_N$ ,  $d_N = H - \beta_N + [c + o(1)]\sigma_N/\sqrt{N}$ ,

$$\sigma_N^2 = \begin{cases} ND \left[ \sum_{i=1}^v H_i (\hat{p}_{i+1} - p_{i+1}) \right], & v > 0 \\ 0, & \text{if } v = 0 \end{cases}$$

$$\beta_N = \begin{cases} \sum_{i=v+w+1}^s H_i p_{i+1}, & \text{if } v + w < s - 1 \\ 0, & \text{if } v + w = s - 1, \end{cases}$$

$$\rho_N = p_2 + \dots + p_s.$$

These results have been extended also to the case of a non-homogeneous multinomial scheme, i.e., when we consider a sequence of independent experiments in a "triangular array" scheme ( $N$  experiments in the series number  $N$ ) with  $s$  outcomes in each, and the probability of outcomes depends not only on  $N$  but also on the number of the experiment in each series.

To derive these results for schemes with dependent events let us consider the case when the events are connected to a Markov chain. For example when the Markov chain has two states we can describe the class of all limit distributions for functions of the statistics of its parameters.

#### References

- [1] BHATTACHARYA, R.N., and J.K. GHOSH: On the validity of the formal Edgeworth expansion. - *Ann. Statist.* 6, 1978, 434-451.
- [2] CHANDRA T.K., and J.K. GHOSH: Valid asymptotic expansions for the likelihood ratio statistics and other perturbed chi-square variables. - *Sankhyā Ser. A* 41, 1979, 22-47.
- [3] CRAMÉR, H.: *Mathematical methods of statistics.* - Princeton University Press, Princeton, 1946. (First printing.)
- [4] MUKHAMEDKHANOVA, R.: A class of limit distributions for functions of the vector of the frequency of appearance of events in multinomial scheme. - *Izv. Akad. Nauk UzSSR, Ser. Fiz.-Mat. Nauk* 1984, 4, 30-36 (Russian).
- [5] MUKHAMEDKHANOVA, R.: A class of limit theorems for functions of statistical estimates of parameters of a Markov chain with two states. - *Izv. Akad. Nauk UzSSR, Ser. Fiz.-Mat. Nauk* 1985, 1, 29-35 (Russian).
- [6] ZUBKOV, A.M.: Limit distributions for a statistical estimate of the entropy. - *Theory Probab. Appl.* 18, 1973, 611-618 (English translation of *Teor. Veroyatnost. i Primenen.* 18, 1973, 643-650).

Academy of Sciences Uzbek SSR  
 Institute of Mathematics  
 F. Hodjaeva 29  
 700143 Tashkent  
 USSR