Wolfgang RICHTER

1958

SovietRxiv

Abstract

Full Text

MATHEMATICS

Wolfgang RICHTER

LIMITING BEHAVIOR OF THE χ^2 DISTRIBUTION IN THE CASE OF LARGE DEVIATIONS

(Presented by Academician I. M. Vinogradov, 22 XI 1957)

1. In the present note an application is given of a certain multidimensional local theorem for large deviations (^{1}) to derive a simple case of the multidimensional integral theorem for large deviations, namely, a theorem is given on the limiting behavior of the distribution $\mathbf{P}\{\chi^2 > \tau^2\}$ as τ, together with the number of observations n, tends to infinity.

The problem of the limiting distribution of the quantity χ^2 in the following form was first posed and solved by Pearson (^{2}). Consider a sequence of independent trials on one and the same random variable. There are $s+1$ different incompatible outcomes possible, which occur with positive probabilities $p_1, \dots, p_{s+1}$, $\sum_{j=1}^{s+1} p_j = 1$. Let ν_j be the number of appearances of the j-th outcome among the first n results of the trials,

$$\sum_{j=1}^{s+1} \nu_j = n, \qquad \mathbf{E}\nu_j = np_j.$$

Following Pearson, form the sum

$$\chi^2 = \sum_{j=1}^{s+1} \frac{(\nu_j - np_j)^2}{np_j}.$$

2. **Theorem.** A. If $\tau = o(n^{1/6})$ as $n \to \infty$, then

$$\mathbf{P}\{\chi^2 > \tau^2\} = \frac{1}{2^{s/2}\Gamma(s/2)} \int_{\tau^2}^{\infty} x^{s/2-1} e^{-x/2}\, dx\, [1 + o(1)].$$

B. Let $\tau = o(\sqrt{n})$ as $n \to \infty$, $\tau > 1$; let D be a fixed sufficiently large number ($D > 4s$). Then

$$\mathbf{P}\{\chi^2 > \tau^2\} =$$

$$= \frac{1}{[2\pi]^{s/2}} \int_{\tau^2 \le \|\xi\|^2 \le D\tau^2} \cdots \int \exp\left\{ -\frac{\|\xi\|^2}{2} + n \sum_{k=3}^{\infty} Q_k\left(\frac{\xi}{\sqrt{n}}\right) \right\} d\xi \left[1 + O\left(\frac{\tau}{\sqrt{n}}\right) \right] + R,$$

where

$$R = \mathbf{P}\{\chi^2 > D\tau^2\} < 2s \exp\left\{ -\frac{D\tau^2}{4s} \right\}$$

for $\tau < \alpha\sqrt{n}$ for some $\alpha > 0$ and all n.

Here ξ denotes a row vector of s-dimensional space, $\|\xi\|$ is its length, $d\xi$ is the volume element in the same space; $Q_k(t)$ $(k = 3, 4, \dots)$ is a polylinear form of order k, whose coefficients depend

of the probabilities p_j $(j = 1, 2, \dots, s+1)$ (see (2)). The series $\sum_{k=3}^{\infty} Q_k(t)$ converges absolutely in a neighborhood of the origin

$$\|t\|^2 < \min_{1 \le j \le s+1} \{p_j\}.$$

The theorem shows that the classical χ^2 method for testing hypotheses is fully applicable for not-too-large deviations; the limit of applicability turns out to be $\tau = o(n^{1/6})$ as $n \to \infty$. For large deviations, the limiting expression necessarily involves the probabilities p_j of the distribution of the quantity ξ under special consideration.

3. Let us outline the proof for the general case B. It is easy to translate the problem into the language of vectors in $(s+1)$-dimensional space. Consider a sequence of $(s+1)$-dimensional random vectors $\vec{\mu}^{(k)}$, $k = 1, 2, \dots$, which may take $s+1$ different values

$$\mathbf{e}^{(j)} = (0, 0, \dots, 0, p_j^{-1/2}, 0, \dots, 0)$$

(only the j-th coordinate is different from zero and is equal to $p_j^{-1/2}$) with probabilities, respectively, p_j, $j = 1, 2, \dots, s+1$. The vector of mathematical expectations of the coordinates $\vec{\mu}^{(k)}$ will be

$$E\vec{\mu}^{(k)} = \mathbf{p} = (\sqrt{p_1}, \dots, \sqrt{p_{s+1}}),$$

and for the mixed second moments σ_{jl} we obtain

$$\sigma_{jl} = E(\mu_j^{(k)} - E\mu_j^{(k)})(\mu_l^{(k)} - E\mu_l^{(k)}) = \delta_{jl} - \sqrt{p_{jp}l}, \qquad j, l = 1, 2, \dots, s+1,$$

$$\Delta = \det \|\sigma_{jl}\| = 0.$$

Put

$$\bar{\mathfrak{n}} = \frac{\sum_{k=1}^{n} (\vec{\mu}^{(k)} - E\vec{\mu}^{(k)})}{\sqrt{n}}.$$

It is easy to see that

$$\chi^2 = \|\bar{\mathfrak{n}}\|^2.$$

Applying some orthogonal transformation $\mathfrak{U}$, one can arrange that the last $(s+1)$-st coordinate of all points $\mathfrak{g}^{(j)} = (\mathbf{e}^{(j)} - \mathbf{p})\mathfrak{U}$ is equal to zero. Denote

$$\vec{\rho}^{(k)} = (\vec{\mu}^{(k)} - \mathbf{p})\mathfrak{U} \quad \text{and} \quad \mathfrak{w} = \bar{\mathfrak{n}}\mathfrak{U}.$$

Then we have $E\mathfrak{w} = 0$, $E\mathfrak{w}'\mathfrak{w} = \mathfrak{E}_s$, and $\chi^2 = \|\mathfrak{w}\|^2$. We shall omit, here and in what follows, the unnecessary $(s+1)$-st coordinate in all occurring vectors. Now the vectors $\vec{\rho}^{(k)}$, $k = 1, 2, \ldots$, are independent, identically distributed, and lattice s-dimensional random vectors. The lattice is defined by the linearly independent vectors

$$\mathfrak{h}^{(j)} = \mathfrak{g}^{(j+1)} - \mathfrak{g}^{(1)}, \qquad j = 1, 2, \ldots, s.$$

All lattice points are covered by the points

$$\mathfrak{g}^{(1)} + \sum_{j=1}^{s} l_j \mathfrak{h}^{(j)},$$

where the l_j are arbitrary integers. The main characteristic of the lattice is the volume h of the parallelepiped formed by the vectors $\mathfrak{h}^{(j)}$, i.e. of the set of points

$$\sum_{j=1}^{s} \lambda_j \mathfrak{h}^{(j)}, \qquad 0 \le \lambda_j \le 1, \quad i = 1, \ldots, s.$$

Therefore the multidimensional local theorem for large deviations is applicable (^{1}). Denote

$$\mathcal{P}_n(\mathbf{l}) = \mathbf{P}\left\{ \sum_{k=1}^{n} \vec{\rho}^{(k)} = \sum_{j=1}^{s} l_j \mathfrak{h}^{(j)} + n\mathfrak{g}^{(1)} \right\},$$

$$\mathfrak{x} = \frac{1}{\sqrt{n}} \left[\sum_{j=1}^{s} l_j \mathfrak{h}^{(j)} + n\mathfrak{g}^{(1)} \right], \qquad \mathbf{l} = (l_1, \ldots, l_s),$$

where the l_j are integers.

In our particular case this theorem gives the following:

If $\|\mathfrak{z}\| = o(\sqrt{n})$ as $n \to \infty$, $\|\mathfrak{z}\| > 1$, then

$$\frac{\dfrac{n^{s/2}}{h} \mathcal{P}_n(\mathfrak{l})}{\dfrac{1}{[2\pi]^{s/2}} \exp\left\{-\frac{\|\mathfrak{z}\|^2}{2}\right\}} = \exp\left\{ n \sum_{k=3}^{\infty} Q_k\left(\frac{\mathfrak{z}}{\sqrt{n}}\right) \right\} \left[1 + O\left(\frac{\|\mathfrak{z}\|}{\sqrt{n}}\right) \right]. \tag{1}$$

Here $Q_k(t)$ is a certain multilinear form of order k; $k = 3, 4, \ldots$.

This limiting formula can also be derived directly from the expression $\mathcal{P}_n(\mathfrak{l})$ by means of Stirling' s formula. Thus, we obtain the explicit form of the multilinear forms $Q_k(t)$. Denote

$$z_j\sqrt{np_j} = l_j - np_j$$

$(j = 1, \ldots, s+1)$. If $\sum_{j=1}^{s+1}|z_j| = o(\sqrt{n})$ as $n \to \infty$, then

$$\mathcal{P}_n(\mathfrak{l}) = \frac{h}{[2\pi n]^{s/2}} \exp\left\{-\frac{\chi^2}{2} + n\sum_{k=3}^{\infty}\frac{(-1)^{k-1}}{k(k-1)}\sum_{j=1}^{s+1}p_j\left(\frac{z_j}{\sqrt{np_j}}\right)^k\right\}\left[1 + O\left(\sum_{j=1}^{s+1}|z_j|/\sqrt{n}\right)\right].$$

Applying the transformation $\mathfrak{U}$, we obtain for $Q_k(t)$:

$$Q_k(t) = \frac{(-1)^{k-1}}{k(k-1)}\sum_{j=1}^{s+1}p_j\left(t_j\sqrt{\frac{\pi_j}{p_j\pi_{j-1}}} - \sum_{i=1}^{j-1}t_i\sqrt{\frac{p_i}{\pi_i\pi_{i-1}}}\right)^j \tag{2}$$

$$\left(t_{s+1} = 0, \qquad \pi_j = 1 - \sum_{l=1}^{j}p_l, \qquad \pi_0 = 1, \qquad \pi_s = p_{s+1}\right).$$

It is easy to see that the series $\sum_{k=s}^{\infty}Q_k(t)$ converges absolutely inside the sphere

$$\|t\|^2 < \min_{1\le j\le s+1}\{p_j\}.$$

4. In order to compute $\mathbf{P}\{\chi^2 > \tau^2\}$ under the condition $\tau = o(\sqrt{n})$ as $\mathbf{P} \to \infty$, one must choose a sufficiently large number D $(D > 4s)$ and decompose $n\{\chi^2 > \tau^2\}$ into the sum

$$\mathbf{P}\{\chi^2 > \tau^2\} = \mathbf{P}\{\tau^2 < \chi^2 \le D\tau^2\} + \mathbf{P}\{\chi^2 > D\tau^2\}.$$

The second term is easily estimated with the aid of an inequality of S. N. Bernstein ((^{3}), p. 162). We have

$$\mathbf{P}\{\chi^2 > D\tau^2\} \le \sum_{j=1}^{s}\mathbf{P}\left\{|w_j| > \tau\sqrt{\frac{D}{s}}\right\} < 2s\exp\left\{-\frac{D\tau^2}{4s}\right\}$$

for all n and for all τ in the range $0 < \tau < \alpha\sqrt{n}$, for some constant $\alpha > 0$.

In computing the first term, the application of the limiting formula (1) is permitted. It can be shown that the sum thereby arising over all points of the lattice $\eta = \sqrt{n}\, r$ for which $n\tau^2 < \|\eta\|^2 \leq D\tau^2 n$ is replaced by the integral over the same region. The error allowed in doing so is of order

$$O\left(\frac{\tau}{\sqrt{n}}\right),$$

which completes the proof.

Leningrad State University
named after A. A. Zhdanov

Received
21 XI 1957

REFERENCES

[1] W. Richter, *Theory of Probability and Its Applications*, 3, no. 1 (1958).
[2] K. Pearson, *Phil. Mag.*, V, 50, 157 (1900).
[3] S. N. Bernstein, *Theory of Probability*, Moscow–Leningrad, 1946.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.