



Soviet-era science, translated into English

Reports of the Academy of Sciences of the USSR

MATHEMATICS

1966

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196601.49284>

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

Abstract

Full Text

Reports of the Academy of Sciences of the USSR
1966. Volume 169, No. 4

UDC 519.272.28

MATHEMATICS

Yu. K. BELYAEV

CONFIDENCE INTERVALS FOR FUNCTIONS OF MANY UNKNOWN PARAMETERS

(Presented by Academician A. N. Kolmogorov on 18 XI 1965)

In a number of areas of practice there arises the problem of constructing a confidence interval for a function of many unknown parameters ^(1,2). In the present paper an algorithmic solution is given for the problem of constructing an upper confidence bound for a concave function

$$f(\lambda_1, \dots, \lambda_m) = \sum_{i=1}^m f_i(\lambda_i).$$

As the initial statistical data, the values d_i , $i = 1, \dots, m$, are used, of mutually independent random variables having Poisson distributions with parameters λ_i . Denote by m_d the number of i for which $d_i = d$, $d = 0, 1, \dots$; $m = \sum_{d=0}^k m_d$,

$$k = \max_{1 \leq i \leq m} d_i, \quad \delta(x) = 0, \quad x = 0, \quad \delta(x) = 1, \quad x > 0.$$

The complexity of the proposed algorithm basically increases with

$$D = \sum_{d=0}^k \delta(m_d).$$

In practically important cases, many $\lambda_i < 1$, and therefore $m_d > 0$ only for small values of d . Thus, for $D < 10$ the proposed algorithm can be used to solve the problem on computers.

The general statistical problem may be formulated as follows. There is a space $X = \{x\}$ of outcomes of trials and a parameter space $\Theta = \{\theta\}$ determining a family of probability distributions P_θ on the σ -algebra \mathfrak{B}_X of subsets of the

space X . On the product $X \times \Theta$ a function $f(x, \theta)$ is given, \mathfrak{B}_X -measurable in x for each $\theta \in \Theta$. It is required, from the observed value $x \in X$, to construct a γ -confidence interval (γ -i.) for $f(x, \theta)$, i.e., to find such \mathfrak{B}_X -measurable functions $\underline{f}(x), \bar{f}(x)$ that

$$\inf_{\theta \in \Theta} P_{\theta} \{ \underline{f}(x) \leq f(x, \theta) \leq \bar{f}(x) \} \geq \gamma.$$

The solution of the problem is carried out on the basis of a chosen system of γ -confidence sets $\{H_x\}$ (γ -s.), $H_x \subseteq \Theta, x \in X$,

$$\inf_{\theta \in \Theta} P_{\theta} \{ \theta \in H_x \} \geq \gamma$$

(3).

Theorem 1. If $\{H_x\}$ are γ -s., then the bounds of the γ -i. are equal to:

$$\underline{f}(x) = \inf_{\theta \in H_x} f(x, \theta), \quad \bar{f}(x) = \sup_{\theta \in H_x} f(x, \theta). \quad (1)$$

Corollary. If prior information on the parameter $\theta, \theta \in \Theta_0 \subseteq \Theta$, is known before the trials, then one can construct a narrower γ -i. by the formulas

$$\underline{f}'(x) = \inf_{\theta \in H_x \cap \Theta_0} f(x, \theta), \quad \bar{f}'(x) = \sup_{\theta \in H_x \cap \Theta_0} f(x, \theta). \quad (2)$$

We assume that $\underline{f}(x), \bar{f}(x), \underline{f}'(x), \bar{f}'(x)$ are \mathfrak{B}_X -measurable. In the particular problem considered below, measurability is a simple consequence of the initial assumptions.

When constructing γ -c.s., it is recommended [2] to use unbiased efficient estimates (for definitions see [3, 4]), if such exist. However, in those cases where Θ is a subset of a Euclidean space of large dimension (> 5), such a procedure is very laborious.

Let now the space $X = \{x = (d_1, \dots, d_m)\}$, where $d_i = 0, 1, \dots$, be the values of m mutually independent random variables having Poisson distributions with parameters $\lambda_1, \lambda_2, \dots, \lambda_m$. The space

$$\Theta = \{\theta\} = \{(\lambda_1, \dots, \lambda_m), \lambda_i \geq 0\}; \quad f(\theta) = \sum_{i=1}^m f_i(\lambda_i),$$

where $f_i(\lambda_i)$ are concave functions, $f_i(0) = 0$. It is required, from the observed value $x = (d_1, \dots, d_m)$, to construct an upper bound $\bar{f}(x)$ of a γ -i. for $f(\theta)$. The value $\underline{f}(x)$ is set equal to 0.

Theorem 2. *The following systems of sets are γ -c.s.: system I $\mathfrak{P}_\gamma = \{\mathfrak{P}_{\gamma, d_1, \dots, d_m}\}$; system II $\mathfrak{Q}_\gamma = \{\mathfrak{Q}_{\gamma, d_1, \dots, d_m}\}$; system III $\mathfrak{R}_\gamma = \{\mathfrak{R}_{\gamma, d_1, \dots, d_m}\}$:*

$$\mathfrak{P}_{\gamma, d_1, \dots, d_m} = \{\lambda_i : 0 \leq \lambda_i \leq \Delta_{1-\gamma_0}(d_i), i = 1, \dots, m\}; \quad \gamma_0^m = \gamma; \quad (3)$$

$$\mathfrak{Q}_{\gamma, d_1, \dots, d_m} = \left\{ \lambda_i : \sum_{i=1}^m \lambda_i \leq \Delta_{1-\gamma} \left(\sum_{i=1}^m d_i \right) \right\}; \quad (4)$$

$$\mathfrak{R}_{\gamma, d_1, \dots, d_m} = \mathfrak{P}_{\gamma_1, d_1, \dots, d_m} \cap \mathfrak{Q}_{\gamma_2, d_1, \dots, d_m}, \quad \gamma = \gamma_1 + \gamma_2 - 1. \quad (5)$$

Here $\Delta_\alpha(d)$ is the solution of the transcendental equation

$$\sum_{k=0}^d \frac{[\Delta_\alpha(d)]^k}{k!} e^{-\Delta_\alpha(d)} = \alpha,$$

tables of the values $\Delta_\alpha(d)$ are given in [2].

From Theorems 1 and 2 and the concavity property of $f(\theta)$, one can obtain the following assertion.

Theorem 3. *The upper bounds $\bar{f}(d_1, \dots, d_m)$ of the γ -i., constructed on the basis of the γ -c.s. \mathfrak{P}_γ and \mathfrak{Q}_γ , have the form*

$$\begin{aligned} \bar{f}_I(d_1, \dots, d_m) &= \sum_{i=1}^m f_i(\Delta_{1-\gamma_0}(d_i)), \quad \bar{f}_{II}(d_1, \dots, d_m) = \\ &= \max_{i=1, \dots, m} \left\{ f_i \left(\Delta_{1-\gamma} \left(\sum_{i=1}^m d_i \right) \right) \right\}. \end{aligned} \quad (6)$$

In those cases where many $d_i = 0$, the γ -c.s. \mathfrak{Q}_γ gives better results than the γ -c.s. \mathfrak{P}_γ . As many of the d_i increase, the γ -c.s. \mathfrak{Q}_γ loses its advantage. Numerical calculations show that in most cases the use of the γ -c.s. \mathfrak{R}_γ gives better results. The confidence sets \mathfrak{R}_γ are polyhedra in m -dimensional space, as follows from formulas (3)–(5).

Theorem 4. *The coordinates of the vertices $O_{\mathfrak{R}}$ of the polyhedron $\mathfrak{R}_{\gamma, d_1, \dots, d_m}$, at which the maximum of the function $f(\theta) = \sum_{i=1}^m f_i(\lambda_i)$ is attained, are specified as follows. To each such vertex there corresponds a set $S \cup i_0$*

indices i for which

$$\sum_{i \in S} \Delta_{1-\gamma_0}(d_i) \leq \Delta_{1-\gamma_1} \left(\sum_{i=1}^m d_i \right) < \sum_{i \in S} \Delta_{1-\gamma_0}(d_i) + \max_{j \in S} \Delta_{1-\gamma_0}(d_j),$$

$$\gamma_0^m = \gamma_2, \quad \gamma = \gamma_1 + \gamma_2 - 1.$$

We set the values of the coordinates of the vertex λ_i , $i \in S$, equal to $\Delta_{1-\gamma_0}(d_i)$. The remaining coordinates λ_i , $i \in S$, except for one $i_0 \in S$, are set equal to zero, and the value $\lambda_{i_0} = x$, where x is the solution of the equation

$$\sum_{i \in S} \Delta_{1-\gamma_0}(d_i) + x = \Delta_{1-\gamma_1} \left(\sum_{i=1}^m d_i \right). \quad (7)$$

The number of vertices indicated in Theorem 4 may be too large for large values of m . We shall call a set of integers $\mathcal{L} = (l_0, \dots, l_k)$ admissible if

$$0 \leq l_d \leq \min \left\{ m_d, \frac{\Delta_{1-\gamma_1}(\sum_{i=1}^m d_i)}{\Delta_{1-\gamma_0}(d)} \right\}, \quad d = 0, 1, \dots, k,$$

and if the inequalities

$$\sum_{d=0}^k l_d \Delta_{1-\gamma_0}(d) \leq \Delta_{1-\gamma_1} \left(\sum_{i=1}^m d_i \right) < \sum_{d=0}^k l_d \Delta_{1-\gamma_0}(d) + \max_{d: m_d - l_d > 0} \Delta_{1-\gamma_0}(d)$$

are satisfied.

We shall say that the sets of vertices $\Gamma_{\mathcal{L}}$ of the polyhedron $\mathcal{R}_{\gamma_{d_1, \dots, d_m}}$ correspond to the admissible set \mathcal{L} , if, for the sets S determining the coordinates of these vertices, the number of indices $i \in S \cap I_d$ is equal to l_d , $d = 0, 1, \dots, k$. Here $I_d = \{i : d_i = d\}$.

We denote the solution x of equation (7) by $x_{\mathcal{L}}$.

It follows from Theorem 4 that the vertex at which the absolute maximum $f(\theta)$ is attained belongs to one of the sets $\Gamma_{\mathcal{L}}$. The maximum of $f(\theta)$, taken over the set of vertices $\Gamma_{\mathcal{L}}$, is found as follows. Form the sets $F_d \subseteq I_d \subseteq (1, \dots, m)$, $d = 0, 1, \dots, k$. F_d contains l_d indices $i \in I_d$ for which $a_{id} = f_i(\Delta_{1-\gamma_0}(d)) \geq a_{jd}$, $j \in I_d \setminus F_d = E_d$. If for the value d , $x_{\mathcal{L}} > \Delta_{1-\gamma_0}(d)$, then the following arguments are applied to the index $d+1$. If, however, $x_{\mathcal{L}} < \Delta_{1-\gamma_0}(d)$, then the values $b_{i,\mathcal{L}} = f_i(x_{\mathcal{L}})$ are found, and the value k_d , $b_{k_d,\mathcal{L}} = \max_{i \in E_d} b_{i,\mathcal{L}}$. In this case the following cases are possible:

- 1) the inequality holds

$$(a_{i,d} + b_{k_d,\mathcal{L}}) \geq (a_{k_d,d} + b_{i,\mathcal{L}}), \quad i \in F_d; \quad (8)$$

- 2) there is an $i_d \in F_d$ such that

$$\max_{i \in F_d} [a_{k_d,d} + b_{i,\mathcal{L}} - a_{i,d} - b_{k_d,\mathcal{L}}] > 0, \quad (9)$$

the maximum in (9) being attained at $i = i_d$.

Case 2) is divided into two subcases depending on the fulfillment of one of the inequalities

$$a_{k_d,d} \geq a_{i,d}, \quad i \in E_d, \quad (10)$$

$$a_{k_d,d} < a_{j_d,d} = \max_{j \in E_d \setminus k} a_{j,d}. \quad (11)$$

When (8) is fulfilled, the value

$$\varphi_{d,\mathcal{L}}^{(1)} = \sum' a_{i,c} + b_{k_d,\mathcal{L}} \quad \text{when (9) and (10) are fulfilled,}$$

$$\varphi_{d,\mathcal{L}}^{(2)} = \sum'' a_{i,c} + b_{k_d,\mathcal{L}} \quad \text{when (9) and (11) are fulfilled;}$$

$$\varphi_{d,\mathcal{L}}^{(3)} = \sum''' a_{i,c} + b_{i_d,\mathcal{L}},$$

where \sum' is taken over values $i \in \bigcup_{c=0}^k F_c$, \sum'' over

$$i \in \left(\bigcup_{c \neq d} F_c \right) \cup (F_d \setminus i_d) \cup k_d$$

and \sum''' over

$$i \in \left(\bigcup_{c \neq d} F_c \right) \cup (F_d \setminus i_d) \cup j_d.$$

Theorem 5.

$$\max_{O_{\mathcal{X}} \in \Gamma_{\mathcal{L}}} f(\theta) = \max_{0 \leq d \leq k} \{\varphi_{d,\mathcal{L}}^{(i_d)}\},$$

where $i_d = 1$ if (8) holds, $i_d = 2$ if (9) and (10) hold, and $i_d = 3$ if (9) and (11) hold.

Thus, finding the absolute maximum of $f(\theta)$ is in fact, in complexity, equivalent to examining all admissible sets, whose number is small for small D . To compare the systems $\mathcal{P}_\gamma, \mathcal{Q}_\gamma, \mathcal{R}_\gamma$, consider the function

$$P = \prod_{i=1}^{20} (1 - (1 - p_i)^2).$$

For each p_i , binomial trials of size N_i were carried out. We assume that the numbers of “failures” d_i are small, so that Poisson approximations with $\lambda_i = N_i(1 - p_i)$ may be used.

Using Theorems 3-5 for

$$\ln P = \sum_{i=1}^{20} f_i(\lambda_i), \quad f_i(\lambda_i) = \ln(1 - (\lambda_i/N_i)^2),$$

when $m_0 = 5$, $m_1 = 15$, we find: for the system $\mathcal{P}_{0.9}$, $\underline{P} = 0.907$; for $\mathcal{Q}_{0.9}$, $\underline{P} = 0.955$; for $\mathcal{R}_{0.9}$, $\underline{P} = 0.984$. The advantages of the last method are substantial. Similar results were also obtained for other values of m_i .

Moscow State University
named after M. V. Lomonosov

Received
21 X 1965

CITED LITERATURE

- ¹ R. A. Mirnyi, A. D. Solov' ev, in: *Cybernetics in the Service of Communism*, 2, 1964, pp. 213-218.
- ² B. V. Gnedenko, Yu. K. Belyaev, A. D. Solov' ev, *Mathematical Methods of Reliability Theory*, “Nauka,” 1965.
- ³ R. Cramér, *Mathematical Methods of Statistics*, IL, 1948.
- ⁴ A. N. Kolmogorov, *Izv. Akad. Nauk SSSR, Ser. Mat.*, 14, No. 4, 303 (1950).

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

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