сообщения
объединенного
института
ядерных
исследований
дубна

E5-87-698

N.I.Chernov, G.A.Ososkov, V.S.Kurbatov

THE ESTIMATION OF PARAMETERS AND THE GOODNESS-OF-FIT TEST FOR MULTIVARIATE DISTRIBUTIONS FOR DATA PRESENTED AS ONE-DIMENSIONAL HISTOGRAMS

1987

1. <u>Introduction</u>. Some years ago the problem of pion-nucleon interactions with the production of an additional pion near 1 GeV was intensively investigated. It was typical of these processes that the final state interaction, particularly production of the resonance $\Delta(1236)$ strongly affected the final particle spectra. These reactions (see, for example [1]) were analysed in the framework of the so-called isobar models which permitted finding the contribution of different quantum states. One of the main states near the production threshold of the $\Delta(1236)$ resonance is D_{13} (angular momentum of the $\Delta(1236)$ resonance relative to the additional pion in the final state is 2, isotopic spin is 1/2, total angular momentum J=3/2).

In this energy region the reaction $\pi^- p \rightarrow \pi^+ \pi^- n$ is one of the main ones. During computer simulation we used a simplified model where we assumed that the main characteristics of the reaction are described by isobar production in the D_{13} state and the contribution of other states is taken into account as a uniform background. Besides, we neglected the interference between the separate channels of isobar production, i.e. the following expression was used for the joint probability distribution function of the energies E_1, E_2 of the secondary pions [2]:

$$\frac{\partial^2 F}{\partial E_1 \partial E_2} \sim |aR_1|^2 + |bR_2|^2 + a_1 \quad , \qquad (1)$$

where $a = -(\sqrt{3})^{-1} a_2\, e^{i\Phi}$, $b = -(3\sqrt{3})^{-1} a_2\, e^{i\Phi}$,

$$R_1 = \left(\frac{\Gamma_1}{2\pi p_1'}\right)^{1/2} \frac{1}{\omega_0 - \omega_{13} - 1/2\, i\, \Gamma_1} \quad ,$$

$$R_2 = \left(\frac{\Gamma_1}{2\pi p_2'}\right)^{1/2} \frac{1}{\omega_0 - \omega_{23} - 1/2\, i\, \Gamma_1} \quad ,$$

Γ_1 is the width of the $\Delta(1236)$ resonance, ω_0 is the mass of $\Delta(1236)$, ω_{13} and ω_{23} are the masses of $\pi^- n$ and $\pi^+ n$ systems in the final state respectively, p_1' and p_2' are the momenta of $\Delta(1236)$ in the centre-of-mass reaction (see also [2]), a_1 and a_2 are the parameters to be estimated.

Analysis of these reactions required the maximum of available information, i.e. the information from other experiments which, as a rule, was not full, for example, it was available as spectra over a single variable though a three-particle reaction is described by four variables. Sometimes the experiment itself did not permit the full information (for example, in the reaction $\pi^- p \to \pi^0 \pi^0 n$ only the neutron was measured in the experiments of that time). So there was a problem: how to get correct statistical conclusions about the joint distribution of final particle parameters without full experimental information.

From the point of view of mathematical statistics this problem is reduced to the construction of a test permitting estimation of multivariate distribution parameters using data in the form of histograms over single variables. Let us consider this problem in more detail for the two-variable case (generalization of our conclusions to a multivariate case is simple, all the necessary changes are noted below in the text). We take the energies E_1 and E_2 of the secondary pions in (1) as the two variables in question.

Let (E_1', E_1'') and (E_2', E_2'') be the intervals of the range of these variables divided into m_1 and m_2 subintervals respectively. It creates a grid with m_1 m_2 cells on the plane. Using hypothetical distribution function (1) we can calculate the probabilities p_{kl} $(k=1,\dots,m_1,\ l=1,\dots,m_2)$ for all cells. The experimental sample of N events can be grouped over cells as a histogram $\{N_{kl}\}$, $k=1,\dots,m_1$, $l=1,\dots,m_2$, $\sum N_{kl} = N$. But we have not the full data, but the data as histograms over single variables:

$$N_{k\cdot} = \sum_l N_{kl}, \qquad N_{\cdot l} = \sum_k N_{kl} .$$

In the analysis of such experiments the following statistic was used:

$$T = \sum_k \frac{(N_{k\cdot} - N p_{k\cdot})^2}{N p_{k\cdot}} + \sum_l \frac{(N_{\cdot l} - N p_{\cdot l})^2}{N p_{\cdot l}} \qquad (2)$$

(here $p_{k\cdot} = \sum_l p_{kl}$, $p_{\cdot l} = \sum_k p_{kl}$) that was just the sum of χ^2-statistics for single-variable histograms. When the hypothesis on the distribution (1) was tested, the value of T was compared with the table for the χ^2-distribution with (m_1+m_2-2) degrees of freedom. For the estimation of unknown parameters of (1) T statistic was minimized over them. Both methods are not substantiated and can yield significant errors if the distribution parameters are not independent (see the numerical example in the Appendix).

Here we investigate the correct distribution of T-statistic (2) (Section 2), construct modified (generalized) T_m statistic for testing the hypothesis on the distribution $F(E_1,E_2)$ (Section 3) and investigate the estimates of unknown parameters of $F(E_1,E_2)$ by minimizing T_m statistic (Section 4).

2. <u>Distribution of T statistic</u>. Let $m=m_1+m_2$, $N_k=N_{k\cdot}$ and $p_k=p_{k\cdot}$ for $k=1,\dots,m_1$ and $N_{m_1+l}=N_{\cdot l}$, $p_{m_1+l}=p_{\cdot l}$ for $l=1,\dots,m_2$ ("through" numeration). Let also $x_k = (N_k - Np_k)/\sqrt{Np_k}$ for $k=1,\dots,m$ and $x=(x_1,x_2,\dots,x_m)$ be a column vector. For this notation $T = x^T x$ (upper index T means transposition).

<u>Lemma 1</u>. The vector x is asymptotically normal with mean zero and covariance matrix

$$V = \left[\begin{array}{c|c} V_1 & V_{12} \\ \hline V_{12}^T & V_2 \end{array}\right]$$

consisting of blocks V_1 $(m_1 \times m_1)$, V_2 $(m_2 \times m_2)$, V_{12} $(m_1 \times m_2)$. These blocks have the form $V_1 = I_{m_1} - u_1 u_1^T$, $V_2 = I_{m_2} - u_2 u_2^T$, here and on I_k is for the $k \times k$ unit matrix, u_1 and u_2 are vector columns

$$u_1 = (\sqrt{p_{1\cdot}}, \sqrt{p_{2\cdot}}, \dots, \sqrt{p_{m_1\cdot}}), \quad u_2 = (\sqrt{p_{\cdot 1}}, \sqrt{p_{\cdot 2}}, \dots, \sqrt{p_{\cdot m_2}});$$

the (k,l)-th element v_{kl} of the block V_{12} has the form

$$v_{kl} = \frac{p_{kl} - p_{k\cdot} p_{\cdot l}}{\sqrt{p_{k\cdot} p_{\cdot l}}} .$$

<u>Proof</u>. We follow the idea of H.Cramer (/3/, p.p. 418-419). The joint characteristic function of the quantities N_{kl} $(k=1,\dots,m_1, l=1,\dots,m_2)$ is

$$\varphi_1(t_{11}, t_{12}, \dots, t_{m_1 m_2}) = \left(\sum_{k=1}^{m_1} \sum_{l=1}^{m_2} p_{kl}\, e^{i t_{kl}} \right)^N .$$

Therefore the joint characteristic function of the quantities N_k $(k=1,\dots,m)$ is

$$\varphi_2(u_1, u_2, \dots, u_{m_1}, v_1, \dots, v_{m_2}) = \left(\sum_{k=1}^{m_1} \sum_{l=1}^{m_2} p_{kl}\, e^{i(u_k + v_l)} \right)^N .$$

Then we obtain the joint characteristic function of the variables x_k $(k=1,\dots,m)$

$$\varphi_3(u_1,u_2,\dots,u_{m_1},v_1,\dots,v_{m_2}) = \exp\left[-\sqrt{N}\left(\sum_{k=1}^{m_1} u_k\sqrt{p_{k\cdot}} + \sum_{l=1}^{m_2} v_l\sqrt{p_{\cdot l}}\right)\right]\times$$

$$\times\left[\sum_{k=1}^{m_1}\sum_{l=1}^{m_2} p_{kl}\exp\left(\frac{iu_k}{\sqrt{Np_{k\cdot}}} + \frac{iv_l}{\sqrt{Np_{\cdot l}}}\right)\right]^N ,$$

and, using the MacLaurin expansion of its logarithm, we deduce by some easy calculation

$$\ln\varphi_3(u_1,u_2,\dots,u_{m_1},v_1,\dots,v_{m_2}) = -\frac{1}{2}U^TVU + O\left(\frac{1}{\sqrt{N}}\right),$$

where U denotes the vector column $(u_1,u_2,\dots,u_{m_1},v_1,\dots,v_{m_2})$. The function φ_3 tends to the characteristic function of the multivariate normal distribution with mean zero and covariance matrix V as $N\to\infty$. Using the continuity theorem (/3/, p.96) we complete the proof.

In the case of the d-dimensional distribution F $(d>2)$ the matrix V consists of $d\times d$ blocks of the same form.

Theorem 1. The characteristic function of the statistic T has the limit

$$\lim_{N\to\infty}\varphi_T(t) = \prod_{k=1}^{m}\left(1-2i\sqrt{\lambda_k}\,t\right)^{-\frac{1}{2}}, \qquad (3)$$

where $\lambda_1\le\lambda_2\le\dots\le\lambda_m$ are the eigenvalues of the matrix V. The density of the asymptotic distribution of T statistic is

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\prod_{k=1}^{m}\left(1-2i\sqrt{\lambda_k}\,t\right)^{-\frac{1}{2}}dt .$$

Proof. Let C be an arbitrary orthogonal matrix, then $e=C^Tx$ is a random vector $e=(e_1,e_2,\dots,e_m)$ also having an asymptotic normal distribution with mean zero and covariance matrix

$$\lim_{N\to\infty}\langle e\,e^T\rangle = \lim_{N\to\infty}\langle C^Txx^TC\rangle = C^TVC$$

(here and further $\langle\cdot\rangle$ denotes the expectation of a random value). Besides, $e_1^2+\dots+e_m^2 = e^Te = x^TCC^Tx = x^Tx = T$. One can find such a matrix C that C^TVC is the diagonal matrix with diagonal elements $\lambda_1,\dots,\lambda_m$. Then the statistic T is the sum of squares of asymptotically independent normal random values $e_1,\dots,e_m$, so

$$\lim_{N\to\infty}\varphi_T(t) = \prod_{k=1}^{m}\lim_{N\to\infty}\varphi_{e_k^2}(t).$$

We have $\lim_{N\to\infty}\varphi_{e_k^2}(t) = \left(1-2i\sqrt{\lambda_k}\,t\right)^{-\frac{1}{2}}$ (see /3/, p.233), so the theorem is proved.

Function (3) coincides with the characteristic function of a χ^2-distribution if and only if $\lambda_k = 1$ for $k=3,\dots,m$. We failed to obtain an analytic expression for the density function f(x), but in any way the distribution of T statistic depends on $F(E_1,E_2)$ and cannot be tabulated (contrary to χ^2-distribution).

Corollary. If the random quantities E_1 and E_2 are independent, i.e. $p_{kl}=p_{k\cdot}\,p_{\cdot l}$, then the statistic T has a limiting χ^2-distribution with $m-2$ degrees of freedom.

Proof. Since $p_{kl}=p_{k\cdot}\,p_{\cdot l}$, the matrix V_{12} consists of zeroes alone. Hence the eigenvalues of V matrix coincide with those of V_1 and V_2 matrices. These matrices have one zero eigenvalue each, and the rest of their eigenvalues are ones (see /3/, p.419). This proves the Corollary.

3. Generalization of T statistic to the case of dependent variables E_1, E_2. Let v_1 and v_2 be vector columns $\left(\sqrt{p_{1\cdot}},\sqrt{p_{2\cdot}},\dots,\sqrt{p_{m_1\cdot}},0,0,\dots,0\right)$ and $\left(0,0,\dots,0,\sqrt{p_{\cdot 1}},\sqrt{p_{\cdot 2}},\dots,\sqrt{p_{\cdot m_2}}\right)$ respectively. One can easy find that $Vv_1 = Vv_2 = 0$, so v_1 and v_2 are eigenvectors of V with the eigenvalue 0. Then the rank of the V matrix does not exceed $m-2$. The ranks of the V_1 and V_2 matrices are m_1-1 and m_2-1 respectively (/3/, p.419), therefore rank $V=m-2$ in general case.

We shall suppose that the distribution $F(E_1,E_2)$ has a general form, i.e. rank $V=m-2$ (or $\lambda_3>0$ - see Theorem 1). In this case one can find a normal random vector $e=(e_1,\dots,e_m)$ with mean zero and the unit covariance matrix such that

$$x = Ae + o_p(1), \qquad (4)$$

where $o_p(1)$ denotes a random value tending to 0 in probability as $N\to\infty$ and A is a $m\times m$ matrix with zeroes in the m-th and m_1-th columns. The matrices A and V are linked by the following relationship:

$$V = \lim_{N\to\infty}\langle xx^T\rangle = \langle Aee^TA^T\rangle = A\langle ee^T\rangle A^T = AA^T. \qquad (5)$$

One can choose the A matrix as a low-triangular one, then its elements can be calculated directly from (5).

We propose a generalization T_m of the statistic T as a quadratic form

$$T_m = \sum_{k,l=1}^{m} x_k x_l q_{kl} = x^T Q x , \qquad (6)$$

where $Q = (q_{kl})$ is the symmetric $m \times m$ matrix which will be specified below. From (4), (6) we obtain $T_m = e^T(A^TQA)e + o_p(1)$. Let the Q matrix be chosen so that

$$A^TQA = I_m - c_{m_1}c_{m_1}^T - c_m c_m^T , \qquad (7)$$

where c_k denotes the vector column of the length m with one at the k-th position and zeroes at the others. If (7) holds then we obtain $T_m = e_1^2+e_2^2+\ldots+e_{m_1-1}^2+e_{m_1+1}^2+\ldots+e_{m-1}^2+o_p(1)$. Therefore we have proved the following theorem:

Theorem 2. Suppose that rank $V = m-2$ and the Q matrix satisfies the condition (7). Then the asymptotic distribution of T_m statistic is that of χ^2 with $m-2$ degrees of freedom.

In the case of a d-dimensional distribution F $(d>2)$ the condition (7) has the form $A^TQA = I_{m_1+\ldots+m_d} - c_{m_1}c_{m_1}^T - c_{m_1+m_2}c_{m_1+m_2}^T - \ldots - c_{m_1+\ldots+m_d}c_{m_1+\ldots+m_d}^T$, where m_p is the number of cells in the p-th axis of the distribution $F(E_1,\ldots,E_d)$, $p=1,\ldots,d$.

Condition (7) defines the Q matrix correctly, but not uniquely. We can define the Q matrix uniquely adding a natural condition: the m_1-th and m-th columns and rows of Q consist of zeroes alone. After that the elements of Q can be consequently evaluated from (7) from right to left and from bottom up (see the Appendix for the FORTRAN-IV program for calculating the matrices A and Q from the given matrix V).

Note. The ambiguity of the definition of the Q matrix in (7) does not influence the value of T_m statistic, i.e. any choice of the matrix Q satisfying (7) does not involve any loss of information. Actually, let L be the subspace in the m-dimensional vector space which is orthogonal to the vectors v_1 and v_2. Then L is a characteristic space of the matrix V. The direct analysis of the geometrical interpretation of formulae (5) and (7) shows that the Q matrix is uniquely defined in L, namely $(Q|_L) = (V|_L)^{-1}$, where $\cdot|_L$ denotes the restriction of an operator to the subspace L. The vectors Qv_1 and Qv_2 may take arbitrary values. One can easily

check it up that the vector x is orthogonal to v_1 and v_2, i.e. $x \in L$. Then the vector Qx and the value $T_m = x^TQx$ are defined by (7) uniquely. In our further considerations we shall suppose that the operator Q is identical at v_1 and v_2: $Qv_1 = v_1$ and $Qv_2 = v_2$. In that case the matrix Q is strictly positive: $Q > 0$.

4. Estimating unknown parameters on the basis of T_m statistic. It is known (/3/, p.p. 426-427) that minimizing the χ^2-sum leads to a test statistic having an asymptotical χ^2-distribution with (m-s-1) degrees of freedom, where m is the number of cells in the data histogram and s is the number of estimated parameters. We obtain an analogous result for T_m statistic in the case of the multivariate distribution with incomplete data.

Let $\alpha = (\alpha_1,\ldots,\alpha_s)$ be the vector of unknown parameters defined in a domain $I \subset \mathbb{R}^s$, and $\alpha_0 = (\alpha_1^0,\ldots,\alpha_s^0)$ be its "true" value $(s<m-1)$. We suppose that

(i) functions $p_{kl}(\alpha)$ have continuous second derivatives;

(ii) α_0 is an inner point of the domain I.

Under these conditions the estimate minimizing T_m is a solution of the equation $\partial T_m/\partial\alpha_r$, $r=1,\ldots,s$, or

$$0 = \frac{1}{\sqrt{N}} \sum_{k,l=1}^{m} \left(2\frac{\partial x_k}{\partial\alpha_r} x_l q_{kl} + x_k x_l \frac{\partial q_{kl}}{\partial\alpha_r} \right) =$$

$$= -2\sum_{k,l=1}^{m} \frac{\partial p_k}{\partial\alpha_r}\left(\frac{1}{\sqrt{p_k}} + \frac{N_k - Np_k}{2Np_k\sqrt{p_k}}\right)x_l q_{kl} + \frac{1}{\sqrt{N}}\sum_{k,l=1}^{m} x_k x_l \frac{\partial q_{kl}}{\partial\alpha_r} .$$

As in /3/, p.426, we shall consider the modified estimate minimizing T_m which is obtained by excluding all terms of the order $O_p(1/\sqrt{N})$ from the last expression and keeping only the terms of the order $O_p(1)$. Now we have simpler equations

$$\sum_{k,l=1}^{m} \frac{\partial p_k}{\partial\alpha_r}\frac{1}{\sqrt{p_k}} x_l q_{kl} = 0 . \qquad (8)$$

Let us suppose that the following additional conditions are valid:

(iii) $p_k(\alpha) > c^2$ for some $c>0$ and all k,α;

(iv) matrix $D = (\partial p_k/\partial\alpha_r)$, $k=1,\ldots,m$, $r=1,\ldots,s$ is of rank s for all α;

(v) rank $V = m-2$ and, moreover, $\lambda_3(\alpha) > c$ for all α (see notation in Theorem 1).

The last condition is an extra one compared with the classic theorem (/3/, p.p. 426-427).

<u>Theorem 3</u>. Under conditions (i)-(v) system (8) has one and only one solution $\hat{\alpha} = (\hat{\alpha}_1,\dots,\hat{\alpha}_s)$ such that $\hat{\alpha}$ tends to α_o in probability as $N\to\infty$. The statistic T_m at the point $\hat{\alpha}$ has an asymptotic χ^2-distribution with $(m-s-2)$ degrees of freedom.

<u>Proof</u>. The proof of the theorem is very much similar to that of the classic theorem (/3/, p.p. 427-434). Therefore we shortly describe the common steps preserving the notation of H.Cramer /3/ and discuss new ones in detail.

Let us denote $p_k^o = p_k(\alpha_o)$, $q_{kl}^o = q_{kl}(\alpha_o)$ and $\partial p_k^o/\partial\alpha_r = \partial p_k/\partial\alpha_r(\alpha_o)$. System (8) is equivalent to

$$\sum_{t=1}^{s}(\alpha_t-\alpha_t^o)\sum_{k,l=1}^{m}\frac{\partial p_k^o}{\partial\alpha_t}\frac{\partial p_l^o}{\partial\alpha_r}\frac{q_{kl}^o}{\sqrt{p_k^o p_l^o}} = \sum_{k,l=1}^{m}\frac{N_k-Np_k^o}{N}\frac{\partial p_l^o}{\partial\alpha_r}\frac{q_{kl}^o}{\sqrt{p_k^o p_l^o}}+\omega_r\,, \qquad (8^a)$$

where

$$\omega_r = \sum_{k,l=1}^{m}\frac{N_k-Np_k^o}{N}\left[\frac{\partial p_l}{\partial\alpha_r}\frac{q_{kl}}{\sqrt{p_k p_l}} - \frac{\partial p_l^o}{\partial\alpha_r}\frac{q_{kl}^o}{\sqrt{p_k^o p_l^o}}\right] - \sum_{k,l=1}^{m}(p_k-p_k^o)\frac{\partial p_l}{\partial\alpha_r}\frac{q_{kl}}{\sqrt{p_k p_l}} - \frac{\partial p_l^o}{\partial\alpha_r}\frac{q_{kl}^o}{\sqrt{p_k^o p_l^o}} - \sum_{k,l=1}^{m}\frac{\partial p_l^o}{\partial\alpha_r}\frac{q_{kl}^o}{\sqrt{p_k^o p_l^o}}\left[(p_k-p_k^o) - \sum_{t=1}^{s}\frac{\partial p_k^o}{\partial\alpha_t}(\alpha_t-\alpha_t^o)\right]. \qquad (8^b)$$

Let us introduce a matrix $B = (b_{kl})$, $b_{kl} = \dfrac{1}{\sqrt{p_k^o}}\dfrac{\partial p_k^o}{\partial\alpha_l}$, $k=1,\dots,m$, $l=1,\dots,s$ and a vector column $\omega = (\omega_1,\omega_2,\dots,\omega_s)$. Then one can rewrite (8^a) in the matrix form

$$B^TQB(\alpha-\alpha_o) = N^{-1/2}B^TQx+\omega\,, \qquad (9)$$

where $Q = (q_{kl}^o)$ and $x = x(\alpha_o)$. Note that the matrix B can be represented as $B = P_oD_o$, where $D_o = D(\alpha_o)$ and P_o is the diagonal matrix formed by the diagonal elements $1/\sqrt{p_1^o}, 1/\sqrt{p_2^o},\dots,1/\sqrt{p_m^o}$. Hence, by virtue of assumption (iv), the matrix B is of rank s. Using the note at the end of Section 3, we deduce that the matrix B^TQB is not a singular one, so equation (9) can be rewritten as

$$\alpha = \alpha_o + N^{-1/2}(B^TQB)^{-1}B^TQx + (B^TQB)^{-1}\omega\,. \qquad (10)$$

According to Chebyshev's inequality, $P\{|N_k-Np_k^o| \geq \delta\sqrt{N}\} \leq p_k^o/\delta^2$ for any $\delta>0$. Hence

$$P\{|N_k-Np_k^o| \geq \delta\sqrt{N} \text{ for at least one } k=1,\dots,m\} \leq \delta^{-2}\sum_{k=1}^{m}p_k^o = 2\delta^{-2}\,.$$

Taking $\delta = N^{1/8}$ and using (iii) we obtain $|x_k| < \delta/c$ for every $k=1,\dots,m$ with the probability $P_N = 1-2N^{-1/4}$ ($P_N\to 1$ as $N\to\infty$).

By virtue of (i),(iii),(iv) and the note at the end of Section 3 the quantities q_{kl} and ω_r are smooth function of α. Therefore, as in /3/, for any α' and α'' with the probability P_N the following inequality holds:

$$|\omega_k(\alpha') - \omega_k(\alpha'')| \leq K|\alpha'-\alpha''|\,(|\alpha'-\alpha_o| + |\alpha''-\alpha_o| + \delta/\sqrt{N}) \qquad (11)$$

where $K>0$ is a constant.

Let us introduce a sequence of vectors $\{\alpha_\nu\}$, $\nu=1,2,\dots$:

$$\alpha_\nu = \alpha_o + N^{-1/2}(B^TQB)^{-1}B^TQx + (B^TQB)^{-1}\omega(\alpha_{\nu-1})$$

with

$$\alpha_1 - \alpha_o = N^{-1/2}(B^TQB)^{-1}B^TQx\,, \qquad (12)$$

i.e. we have $|\alpha_1-\alpha_o| < c_1\delta/\sqrt{N}$ with the probability P_N, where $c_1>0$ is a constant. By virtue of (11) for some constant $c_2>0$ we have with the probability P_N

$$|\alpha_{\nu+1}-\alpha_\nu| = |(B^TQB)^{-1}[\omega(\alpha_\nu) - \omega(\alpha_{\nu-1})]| \leq$$
$$\leq c_2|\alpha_\nu-\alpha_{\nu-1}|\,(|\alpha_\nu-\alpha_o| + |\alpha_{\nu-1}-\alpha_o| + \delta/\sqrt{N})\,.$$

Using induction, one can easily show that for sufficiently large N with the probability P_N

$$|\alpha_{\nu+1}-\alpha_\nu| \leq c_1c_3^{\nu}(\delta/\sqrt{N})^{\nu+1}\,, \qquad (13)$$

where $c_3 = (4c_1+1)c_2$. Therefore, with the probability P_N, the sequence $\{\alpha_\nu\}$ has a limit - a vector $\hat{\alpha}$ which is a solution of (10). By virtue of (13) we have $|\hat{\alpha}-\alpha_o| \leq 2c_1\delta/\sqrt{N}$ with the probability P_N, i.e. $\hat{\alpha}$ tends to α_o in probability as $N\to\infty$. If $\tilde{\alpha}$ is another solution of (10) tending to α_o in probability, then by virtue of (11), with the probability P_N,

$$|\hat{\alpha}-\tilde{\alpha}| = |(B^TQB)^{-1}[\omega(\tilde{\alpha}) - \omega(\hat{\alpha})]| \leq$$
$$\leq \text{const}\;|\tilde{\alpha}-\hat{\alpha}|\,(|\tilde{\alpha}-\alpha_o| + |\hat{\alpha}-\alpha_o| + \delta/\sqrt{N})\,.$$

But the expression within parenthesis tends to 0 in probability as

$N \to \infty$, so a contradiction arises. Therefore the first part of the Theorem has been proved.

Note. So far we did not use the special form of the matrix Q defined by (7). Hence the first part of the Theorem is valid for any matrix $Q = Q(\alpha)$ which is a smooth function of α and has eigenvalues $\lambda_1, \ldots, \lambda_m$ satisfying the restriction $0 < c' < |\lambda_k| < c'' < \infty$ for all k and α for some constants c', c''.

Now we continue the proof of the Theorem. According to (9) and (11) we have $(B^TQB)^{-1}\omega(\hat{\alpha}) = \hat{\alpha} - \alpha_1 = (\alpha_2 - \alpha_1) + (\alpha_3 - \alpha_2) + \ldots$ and by virtue of (13) we have

$$|(B^TQB)^{-1}\omega(\hat{\alpha})| \leq \text{const}\,\frac{\delta^2}{N} \tag{14}$$

with the probability P_N. Hence

$$\hat{\alpha} - \alpha_o = N^{-1/2}\,(B^TQB)^{-1}B^TQx + o_p\left(\frac{\delta^2}{N}\right) , \tag{15}$$

(here and further $o_p(N^{\delta})$ denotes such a random value ξ that $\sup\limits_N P\{N^{-\delta}\xi > c\} \to 0$ as $c \to \infty$).

Let us consider the quantities $\hat{p}_k = p_k(\hat{\alpha})$ and

$$y_k = \frac{N_k - N\hat{p}_k}{\sqrt{N\hat{p}_k}} = x_k - x_k\frac{\hat{p}_k - p_k^o}{\sqrt{\hat{p}_k p_k^o}(\sqrt{\hat{p}_k}+\sqrt{p_k^o})} + \frac{(\hat{p}_k - p_k^o)^2}{\sqrt{N}\sqrt{\hat{p}_k p_k^o}(\sqrt{\hat{p}_k}+\sqrt{p_k^o})} -$$

$$-\sqrt{N}\,\frac{\hat{p}_k - p_k^o}{\sqrt{p_k^o}} = x_k - \sqrt{\frac{N}{p_k^o}}\sum_{l=1}^{s}\frac{\partial p_k^o}{\partial \alpha_l}(\hat{\alpha}_l - \alpha_l^o) + o_p\left(\frac{\delta^2}{N}\right) .$$

These relations can be rewritten as

$$y = x - \sqrt{N}\cdot B\,(\hat{\alpha} - \alpha_o) + o_p\left(\frac{\delta^2}{N}\right) ,$$

where y is the vector column $(y_1, \ldots, y_m)$. Applying (14) and (15) we obtain

$$y = x - B(B^TQB)^{-1}B^TQx - \sqrt{N}\cdot B(B^TQB)^{-1}\omega(\hat{\alpha}) + o_p\left(\frac{\delta^2}{N}\right) =$$

$$= Rx + o_p\left(\frac{\delta^2}{N}\right) ,$$

where $R = I_m - B(B^TQB)^{-1}B^TQ$. By virtue of Lemma 1 the vector y is asymptotically normal with mean zero and covariance matrix $V_1 = RVR^T$.

The value of the statistic T_m at the point $\alpha = \hat{\alpha}$ is

$$T_m(\hat{\alpha}) = y^T\hat{Q}y = y^TQy + y^T(\hat{Q}-Q)y ,$$

where $\hat{Q} = Q(\hat{\alpha})$. Since the matrix Q is a smooth function of α, $\|\hat{Q} - Q\| = o_p(\delta/\sqrt{N})$, i.e. $T_m(\hat{\alpha}) = y^TQy + o_p(1)$.

Let us return to the vectors v_1, v_2 and the subspace L (see Section 3). It is easy to check it up that $B^Tv_1 = B^Tv_2 = 0$, i.e. the image of the s-dimensional vector space $\mathbb{R}^s$ lies inside the space L due to B operation. Let $L_o = QB\mathbb{R}^s \subset L$ and let L_1 be the orthogonal complement to L_o in L. Note that $\dim L_o = s$, $\dim L_1 = m-s-2$. Let us represent the vector y as the sum $y = y_1 + y_o$, $y_1 \in L_1$, $y_o \perp L_1$. Remember that $x \in L$ (Section 3), hence

$$(v_k, y) = (v_k, Rx) + o_p(1) = o_p(1) \quad \text{for } k=1,2 .$$

Then for any s-dimensional vector w we have

$$(QBw, y) = (w, B^TQRx) + o_p(1) = o_p(1) .$$

Therefore $\|y_o\| = o_p(1)$ and $T_m(\hat{\alpha}) = y_1^TQy_1 + o_p(1)$. The vector y_1, like y, is asymptotically normal with mean zero and covariance matrix V_1, i.e. $\operatorname{rank} V_1 = m-s-2$.

One can find such a normal vector e with mean zero and the unit covariance matrix that $y_1 = Ae + o_p(1)$, A is a matrix of the order $m \times m$ which has zeroes in its $(s+2)$ last columns. Then $V_1 = \langle y_1 y_1^T \rangle = \langle Aee^TA^T \rangle = AA^T$. Since $T_m(\hat{\alpha}) = e^T(A^TQA)e + o_p(1)$, the statistic $T_m(\hat{\alpha})$ has a limiting χ^2-distribution with $m-s-2$ degrees of freedom if and only if the matrix $V_2 = A^TQA$ is a diagonal one, its first $(m-s-2)$ diagonal elements are ones and the rest of its elements are zeroes. Since $\operatorname{rank} A = m-s-2$, this condition is equivalent to $AA^TQAA^T = AA^T$, i.e. $V_1QV_1 = V_1$. The last condition can be proved directly if one uses the substitution $V_1 = RVR^T$ and the formula $QV = VQ = I_m - v_1^Tv_1 - v_2^Tv_2$ (see Section 3). Theorem 3 has been completely proved.

In the case of the d-dimensional distribution F $(d>2)$ the statistic $T_m(\hat{\alpha})$ has the asymptotic distribution of χ^2 with $m-s-d$ degrees of freedom. The proof of the Theorem does not change in this case.

Note that the asymptotic covariance matrix of $\hat{\alpha}$ estimate is $N^{-1/2}\,(B^TQB)^{-1}$.

Appendix. For numerical comparison of the discussed distributions we have made a computer experiment. We constructed distribution (1) for $a_1=90.74$ and $a_2=3.71$. The ranges of the variables E_1, E_2 were subdivided into 15 equal intervals each $(m_1=m_2=15)$, and 500 events were being generated with the distribution according to law

(1), i.e. N=500. It was done 12.000 times and the sampling distribution functions of T and T_m statistics were calculated at four points - quantiles of the χ^2-distribution with 28 degrees of freedom ($m_1+m_2-2=28$). The results are presented in the table

x	41.3	45.4	48.3	56.9
$P\{\chi^2_{28} > x\}$	5%	2%	1%	0.1%
$P\{T_m > x\}$	4.7%	1.75%	1.12%	0.15%
$P\{T > x\}$	5.5%	2.67%	1.80%	0.30%

It is seen that the T statistic distribution considerably deviates from the χ^2-distribution n the "tail", i.e. in the most important region of big values of the χ^2-variable: the error of the first kind can be 2-3 times larger than the tabular one.

In conclusion we give the text of the FORTRAN-IV program for calculating the A and Q matrices from the given V matrix. This algorithm requires $m^4/2 + 5/3\ m^3 + 7/2\ m^2 + 2\ m$ floating-point operation and m square root subroutine calls.

```
      SUBROUTINE  QMATR (Q,A,V,M,M1)
C  IT CALCULATES THE MATRICES  Q(M,M)  AND  A(M,M)  FROM THE GIVEN
C  MATRIX  V(M,M). SEE THE NOTATION IN THE TEXT. ONLY FOR D=2.
C
      DIMENSION Q(M,M),A(M,M),V(M,M)
C --------- FIRST STEP: CALCULATE THE MATRIX A.
      DO 4 I=1,M
         DO 3 J=1,M
            A(I,J)=Ø.
            IF (J.GT.I) GO TO 3
            SUM=Ø.
            IF (J.EQ.1) GO TO 2
            J1=J-1
            DO 1 L=1,J1
    1       SUM=SUM+A(I,L)*A(J,L)
    2       IF (J.LT.I) A(I,J)=(V(I,J)-SUM)/A(I,J)
            IF (J.EQ.I) A(I,J)=SQRT(ABS(V(I,J)-SUM))
    3    CONTINUE
    4 CONTINUE
C -------------- SECOND STEP: CALCULATE THE MATRIX Q.
      DO 9 II=1,M
      DO 8 JJ=1,M
      I=M+1-II
      J=M+1-JJ
      Q(I,J)=0.
      IF (I.EQ.M1.OR.I.EQ.M) GO TO 8
      IF (J.EQ.M1.OR.J.EQ.M) GO TO 8
      I1=I+1
      J1=J+1
      SUM=Ø.
      DO 5 K=I1,M
    5 SUM=SUM+A(K,I)*Q(K,J)
```



```
      SUM=SUM*A(J,J)
      DO 7 L=J1,M
      D=Ø.
      DO 6 K=I,M
    6 D=D+A(K,I)*Q(K,L)
      SUM=SUM+D*A(L,J)
    7 CONTINUE
      S=Ø.
      IF (I.EQ.J) S=1.
      Q(I,J)=(S-SUM)/(A(I,I)*A(J,J))
    8 Q(J,I)=Q(I,J)
    9 CONTINUE
      RETURN
      END
```

We acknowledge the valuable help of N.V.Guseva and A.M.Rozhdestvensky in computer simulation.

REFERENCES

1. Bunyatov S.A., Kurbatov V.S., Likhoded A.K. Yad. Fiz. 1972, 16(6), p. 1279-1285.
2. Olsson M.G., Yodh G.B. Phys. Rev., 1966, 145, p. 1309.
3. Cramer H. Mathematical methods of statistics, Princeton Univ. Press, Princeton, 1946.

Received by Publishing Department
on September 17, 1987.

Чернов Н.И., Ососков Г.А., Курбатов В.С. E5-87-698

Оценка параметров и критерий согласия для многомерных распределений в случае данных, представленных в виде одномерных гистограмм

Исследованы задачи оценки параметров многомерных вероятностных распределений и проверки статистических гипотез в физических экспериментах, в которых данные поступают в виде одномерных гистограмм отдельно по каждой переменной. Исследован применяющийся обычно метод "суммы χ^2-статистик" по всем отдельным переменным, показана его некорректность в случае зависимых переменных. Предложено корректное обобщение этого метода, для которого доказаны аналоги классических теорем о предельном распределении χ^2-статистик в случае проверки гипотез и в случае оценки неизвестных параметров. Предложенный метод реализован в виде программы и проиллюстрирован на численном примере для одной модели пион-нуклонных взаимодействий.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Сообщение Объединенного института ядерных исследований. Дубна 1987

Chernov N.I., Ososkov G.A., Kurbatov V.S. E5-87-698

The Estimation of Parameters and the Goodness-of-Fit Test for Multivariate Distributions for Data Presented as One-Dimensional Histograms

The problems of estimating parameters of multivariate probability distributions and testing hypothesis are studied in the framework of physical experiments where data are presented as one-dimensional histograms for every single variable separately. A commonly used method of "sum of χ^2-statistics" over all single variables is investigated. It is shown that this method is incorrect for dependent variables. A correct generalisation of this method is proposed for which analogues of classical theorems concerning asymptotic distributions of χ^2-statistics are proved for testing hypothesis and estimating unknown parameters. The proposed method is implemented as a FORTRAN-IV program and illustrated by a numerical example for some model of pion-nucleon interactions.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Communication of the Joint Institute for Nuclear Research. Dubna 1987