ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

Дубна

E-2353

A.A. Tyapkin

FALLIBILITY OF USING THE χ^2 CRITERION IN THE PHASE SHIFT ANALYSIS FOR CHOOSING STATISTICALLY PERMISSIBLE AMBIGUOUS SOLUTIONS

ЛАБОРАТОРИЯ ЯДЕРНЫХ ПРОБЛЕМ

1965

E-2353

A.A. Tyapkin

FALLIBILITY OF USING THE χ^2 CRITERION IN THE PHASE SHIFT ANALYSIS FOR CHOOSING STATISTICALLY PERMISSIBLE AMBIGUOUS SOLUTIONS

In a phase shift analysis of experimental data in accordance with the general method of a regressive analysis phase shift evaluations are found from the condition of minimization of the further square functional:

$$M(y_i ; \delta_j) = \sum_{i=1}^{n} \frac{[y_i - \hat{\eta}_i(\delta_j)]^2}{\sigma_i^2} , \qquad (1)$$

where y_i is experimentally obtained values,

σ_i is root-mean-square errors of measurements of the given values, and

$\hat{\eta}_i(\delta_j)$ is theoretical values of the same quantities which depend upon the selection of the estimate of the phase shifts δ_1 , δ_2 , ..., δ_m .

If experimental data y_i is subordinated to the normal law of distribution, the minimum values of the functional M should be respectively described by the χ^2-distribution. Consequently, the mean value and the dispersion of this quantity should be, respectively, $EM = m - n$ and $DM = 2(n - m)$.

For a particular minimum value of M_0 corresponding to the given selection of experimental data we can determine basing on the χ^2-distribution the degree of statistical confidence $P(M \geq M_0)$ which defines the rate of acquiring the minimum values of $M_j > M_0$ in a series of mulitple repeated measurements of the whole assembly of experimental data y_{ij} .

The acquisition of the minimum value M_0 corresponding to a low degree of confidence can show the imperfection of the theoretical curve taken for estimating the true regression line, e.g., due to a small number of the parameters varied. The reason of a high value of M_0 can be the presence of systematic errors not taken into account in the measurements of experimental values.

It might seem that the same reliability criterion is acceptable also for clearing out the possibility of neglecting some of the solutions in such a case when the functional manifests some minima in various argument regions. Just due to the seeming validity the χ^2 - criterion has been for many years erraneously taken for clearing out the permissibility of ambiguous solutions obtained. (See, e.g., refs./1-6/). The erraneous recomendation on the use of the χ^2 - criterion in considering the problems of neglecting ambiguous solutions can be found in monographs dedicated specially to the regressive analysis (see, e.g., ref./7/).

In fact the χ^2 -distribution dominates only over the values of absolute minima obtained in the analysis of independent series of repeated experiments. Therefore, the application of the χ^2 -criterion is justifiable only for determining the statistical reliability of a solution corresponding to the deepest minimum of the functional M or the compatibility of various evaluations obtained in separate series of experiments.

The χ^2 -distribution has no bearing on the value of the relative minima of the functional M.

We have no right to consider a large value of the obtained relative minimum of the functional M to be due to the realization (in a given series of measurements) of statistically permissible, large in sums deviations of experimental values from the theoretical curve. Just for this series of experimental values we have the indication to the realization of smaller summed deviations from the curve obtained under the same theoretical assumptions but corresponding to the absolute minimum of the functional M. A selection among the general assembly has been performed, dice have been cast, so to say, the game has been made and, therefore, one cannot refer to the statistical possibility of appearing large in sums errors in repeated selections, in general, if for a particular selection the realization of considerably smaller summed errors by the deepest absolute minimum of the functional M is a fact.

The statistical reliability of solutions corresponding to the relative minimum of the functional M, can be easily determined by the maximum likelihood method. Say, apart from the solution $\hat{\eta}$ (0) corresponding to the absolute minimum M_0, there is another solution $\hat{\eta}$ (I) corresponding to the relative minimum $M_I > M_0$. The probability of obtaining the given selection (i.e., the preset experimental values $y_1, y_2, \dots, y_n$ in the intervals $dy_1, dy_2, \dots, dy_n$, respectively) from the general assembly presented by some true values $\eta_1, \eta_2, \dots \eta_n$ is expressed by the likelihood function

$$L(y_i; \eta_i) = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^{n} \sigma_i} \exp - \tfrac{1}{2} \sum_{i=1}^{n} \frac{(y_i - \eta_i)^2}{\sigma_i^2} \qquad (2)$$

multiplied by the corresponding intervals $dy_1, dy_2 \dots dy_n$.

If for true values we have hypothetically two series of values η_i(-I) and η_i(-II), then having calculated the probabilities of obtaining the given selection from different general assemblies

$$W_I = L[y_i; \eta_i(I)] \prod_{i=1}^{n} dy_i$$

and

$$W_{II} = L[y_i; \eta_i(II)] \prod_{i=1}^{n} dy_i \qquad (3)$$

we can define the statistical reliability of accepting each of the hypotheses.

Accepting one of the hypotheses, e.g., the first one to be true, we can determine the probability that the accepted hypothesis is correct

$$P_I = \frac{W_I}{W_I + W_{II}} = \frac{L[y_i; \eta_i(I)]}{L(y_i; \eta_i(I)] + L[y_i; \eta_i(II)]} = \frac{1}{1 + \exp - \tfrac{1}{2}(M_{II} - M_I)}. \qquad (4)$$

The probability of rejecting the correct hypothesis will be then

$$\alpha_I = 1 - P_I = P_{II} = \frac{W_{II}}{W_I + W_{II}}.$$

If hypotheses 1 and II are different, noncompatible with each other estimates of the general essembly determined on the basis of the given selection $y_1, y_2, \dots y_n$, the corresponding likelihood functions are of another form:

$$L[y_i; \hat{\eta}_i(I)] = \frac{1}{(2\pi)^{\frac{n-m}{2}} \left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)^{m/2} \prod_{i=1}^{n} \sigma_i} \exp - \frac{M_I}{2}$$

$$L[\bar{y}_i; \hat{\eta}_i(II)] = \frac{1}{(2\pi)^{\frac{n-m}{2}} \left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)^{m/2} \prod_{i=1}^{n} \sigma_i} \exp - \frac{M_{II}}{2}. \qquad (5)$$

It can be easily seen that this account of the dependence of the analyzed hypotheses from the experimental data selection itself does not change the expressions for the probabilities P_I and P_{II}. Thus, the problem of comparing the reliabilities of two hypotheses on the general assembly on the basis of selection data does not differ from that of comparing the reliabilities of two noncompatible estimates of the same general assembly.

The above probabilities P_I and P_{II} are related to a statistical sub-assembly consisting of the selections made according to a certain characteristic from the general assembly of possible selections. This sub-assembly consists of selections within the intervals dy_1, dy_2, ... dy_n coinciding with the given selection or with that symmetrical to the given one with respect to both the hypotheses.

We permit ourselves to explain the reason of introducing such statistical sub-assembly by very simple examples which in their essence do not differ from more complicated ones. As justification of the talk started on a statistical assembly let us note that a mess in essence in simple problems on statistics is most often due to the usage of the concept "probability" without rendering concrete a statistical assembly to which it refers. First of all, we should clear out in what sense, for instance, a small probability W_{II} of obtaining a given selection from general assembly II compared to the probability W_I of obtaining the same selection from general assembly I allows one to speak about the unreliability of hypothesis II.

We permit ourselves to dwell upon this problem in more detail apologizing to the reader beforehand for the simplicity of examples taken for illustration since a mess is committed as a result of introducing reciprocal probabilities.

Indeed, one can speak about a small probability of fulfilling hypothesis II and, respectively, about a great probability of fulfilling hypothesis I if one knows deliberatly that only one of them is carried out in nature.

All the above problems including the unlawfulness of using the χ^2-criterion in the phase shift analysis for choosing permissible ambiguous solutions, are illustrated in full measure by various versions of a more vivid example on the determination of the target position by the results of shooting. Suppose, for instance, that we are to determine the motion line $\eta(t)$ of a point target, if the coordinates of hitting $y_1(t_1)$, $y_2(t_2)$ ... $y_n(t_n)$ are known for n shots at a moving target made with the root-mean-square errors σ_i. This problem coincides in full measure with the discussed statistical problems on the phase shift analysis of experimental data on the nuclear particle interactions. This example allows one to more vividly determine simple statistical questions not related to those of special complex problems on nuclear interaction.

We shall not violate the truth if for simplicity of narration in the above example of shooting we introduce some additional simplifications. Assume that an

immovable target is fired at and the root-mean-square error σ is constant. Then instead of comparing the reliabilities of two motion lines of the target we have, respectively, the comparison of the reliabilities of two given positions η_I and η_{II} of a point target by the results of one given series from equally accurate shots.

It is clear that we can easily calculate the probabilities W_I and W_{II} of obtaining just this series of results corresponding to the target positions at the points η_I and η_{II}. These probabilities in statistical assemblies of repeated series on n-shots performed separately on the first and second targets define the corresponding rates of appearing series of results coinciding with the given one $(y_1, y_2 ... y_n)$ within the selected limits dy_1, dy_2, ... dy_n. It should be noted that here we do not mean the general coincidence of results of hitting the target with the given hittings but the coincidence of the results in strict consequence when the result of the first shot coincided just with the first result of the given series, etc.

The probabilities we are interested in

$$P_I = \frac{W_I}{W_I + W_{II}} = \frac{1}{1 + \exp -\frac{(M_{II} - M_I)}{2}} \quad \text{and} \quad P_{II} = \frac{W_{II}}{W_I + W_{II}} = \frac{1}{1 + \exp\left(\frac{M_{II} - M_I}{2}\right)} \qquad (6)$$

are easily treated in the case when the analyzed series of results is known to be a selection from the assembly in which series of results of shooting at both the targets are presented in equal measure. Indeed, in this case the value P_1 is simply a probability that the given series of results was obtained in shooting at the first target. From N cases of acquiring selections coinciding with the given one, NP_1 cases on average arose in shooting at the first target and NP_{11} cases in shooting at the second target. But despite its clarity, this application of the probabilities P_1 and P_{11} is rather limited as it requires the a priori knowledge of probabilities of carrying out each hypothesis in a general statistical body.

Thus, the above NP_I and NP_{II} rates of obtaining results coinciding with the given series are carried out in testing only in the case when a rifle-man equally often shoots on average at both the targets (the case of equality of a priory probabilities). The equality of a priori probabilities is not important since in the case of their difference one can always make a proper recalculation of the P_I and P_{II} values. But one must know the values of a priori probabilities for the above application of P_I and P_{II}. When there is no information on a priori probabilities this particular use of the P_I and P_{II} probabilities is impossible in the framework of a specially made mixed statistical body. However, the possibilities of statistical estimate on the belonging of the given selection to one of the general assemblies is not changed. The authors of monograph [8] stating that it is impossible to solve the problem on the statistical evaluation of hypotheses without the assumption on the a priori probability are absolutely not right.

In the case of interest when it is known that in nature only one hypothesis is brought about (a rifle-man shoots all the time only at one of the targets) one cannot in general compile a mixed statistical body of selections from various general assemblies. The problem of optimal statistical estimate whether a certain selection belongs to a definite assembly or not has, of course, a strict solution in this case. The degree of distinguishing between the two hypotheses depends not only uodn $\eta_I - \eta_{II}$ (the distance between the targets) and the value σ (the inaccuracy of shooting), but also upon an accidental result of the analyzed selection. Thus, in the case of hitting exactly the middle between the two targets ($W_I = W_{II}$) a shot does not carry any additional information for distinquishing between the two hypotheses since with a 50% probability all the more it is possible to simply guess the belonging of a selection to one of the general assemblies. When hitting moves off the central point the degree of possible distinguishing between the two hypotheses is increased. It is of interest that due to the square structure of the value M when the point of hitting is moved off from the positions of both the targets, the degree of possible distinguishing between the two hypotheses is increased. Thus, exact hitting the centre of one of the targets does not provide a maximum reliability for selecting this target.

All the selctions from the general assembly in the example of choosing one of the hypotheses should be classified, separated in groups on the degree of possible distinguishing between the two hypotheses given by them. In this case one and the same group, one and the same statistical sub-assembly will include cases referring symmetrically to both the hypotheses. For example, one statistical sub-assembly will contain the cases of exact hitting any of the targets.

These sub-assemblies can be made up according to the given characteristic both when only one target is shot at and when two targets are shot at in turn without any assumptions on a priori probabilities. The probabilities P_I and P_{II} obtained in accordance with relations (6) refer to such sub-assembles. Some value ξ will be accidental in this problem on the statistical determination of a correct solution by one selection. It is equal to either 1 or 0, if the chosen hypothesis turned out to be correct or false, respectively. The analogous statement of the question is possible in the case of the infinite number of hypotheses continuously turning one into another. Just in this case the artificiality and the lack of logic of introducing the so-called reciprocal probabilities become understandable [12].

A statistically grounded reliability of the second solution $P_{II} = \frac{1}{1+\exp\frac{M_{II}-M_I}{2}}$ turns out to be considerably smaller than the reliability erraneously determined by the χ^2-distribution in the cases when the value $2(n-m)$ essentially exceeds 1. Thus, solutions found in the 310 MeV phase-shift analysis of pp-scattering [1,2] have in fact the following reliabilities:

$P_1(M = 17{,}9) = 0{,}80$ $\quad$ $P_5(M = 34.2) = 0.002$

$P_{11}(M = 21.7) = 0.12$ $\quad$ $P_6(M = 34.6) = 0.002$

$P_3(M = 23.8) = 0.045$ $\quad$ $P_7(M = 41.3) \approx 2\cdot10^{-5}$

$P_4(M = 24.5) = 0.03$ $\quad$ $P_8(M = 52.3) = 3.10^{-7}$.

At the same time according to the χ^2-criterion the first four solutions entered into the 90% reliability region; the 5-th and the 6-th solutions entered into the 1% reliability region.

The incorrectness of a widely spread application of the χ^2-criterion for determining the reliability of ambiguous solutions of the phase shift analysis was cleared out by the author in discussing the paper by F.Lehar and V.V.Fyodorov [9] in which planning of experiments for the determination of a single solution is calculated. I take an opportunity to gratify N.P. Klepikov and V.V. Fyodorov for this discussion.

It is worth noting that the incorrect logic of the error found coincides with the earlier mentioned fallibility of using the χ^2 - criterion [10] for defining the permissible phase shift regions in the analysis by the method of I.M. Gelfand [11].

One has to regret and wonder that this circumstance has been left without any attention and the criticism of paper [11] has not been automatically propageted to the method accepted by everybody which determines the reliability of the ambiguous solutions of the phase shift analysis.

References

1. O.Chamberlain, E.Segre, R.D.Tripp, C.Wiegand, T.Ypsilantis. Phys.Rev., 105, 288 , 1957.
2. H.P.Stapp, T.Ypsilantis, N.Metropolis. Phys.Rev., 105 , 302, 1957.
3. P. Cziffra, M.H.MacGregor, M.J.Moravcsik, H.P.Stapp. Phys.Rev., 114, 880, 1959.
   M.H.MacGregor, M.J.Moravcsik, H.P.Stapp. Phys.Rev., 116, 1248, 1959.

4. Ю.М. Казаринов, И.Н. Силин. ЖЭТФ, 43, 692, 1962.

5. Р.Я. Зулькарнеев, И.Н. Силин. ЖЭТФ, 45, 664, 1963.

6. И. Быстрицкий, Р. Зулькарнеев. ЖЭТФ, 45, 1169, 1963.

7. Н.П. Клепиков, С.И. Соколов. Анализ и планирование экспериментов методом максимума правдоподобия. Изд. "Наука", М., 1964 г., стр. 38.

8. В.И. Гольданский, А.В. Куценко, М.И. Подгорецкий. Статистика отсчетов при регистрации ядерных частиц. Изд. физ-мат, лит., М., 1959 г., стр. 65.

9. Ф. Легар, В.В. Федоров. Ядерная физика (в печати); Препринт ОИЯИ Р-2332.

10. A.A.Tyapkin. Preprint D-642, Dubna, 1960.

11. В.А. Боровиков, И.М. Гельфанд, А.Ф. Грашин, И.Я. Померанчук. ЖЭТФ, 40, 1106, 1961.

И.М. Гельфанд, А.Ф. Грашин, Л.Н. Иванова, И.Я. Померанчук. Я.А. Смородинский. Препринт ОИЯИ Д-598 (1960).

12. Д. Орир. Препринт ОИЯИ Р-292 (1959).

Received by Publishing Department
on September 4, 1965.