Estimates when parametrical hypotheses are accepted and comparison them with maximum likelihood estimates

Victor Aleksandrovich Vlasov¹, Svetlana Victorovna Vlasova², Andrey Olegovich Tolokonsky¹

¹National Research Nuclear University «MEPhI» (Moscow Engineering Physics Institute), Kashirskoye, 31, Moscow, 115409, Russian Federation

²All-Russian Research Institute for Nuclear Power Plants Operation, ¹Ferganskaya str., 25, Moscow, 109507, Russian Federation

Abstract. The estimation task of unknown parameters based on experimental data is proposed to solve taking into account the consequences of estimation errors. For this purpose the criterion of optimality which includes the weights (costs) of errors should be stated. This allows us to construct the optimum critical regions for testing competing parametrical hypotheses using variation methods. Applying the variation approach there is no more need in using confidence levels and confidence intervals as there is an opportunity to find the optimum values of estimates which are best corresponded to available experimental information. In general the optimum estimates received in this way should not coincide with maximum likelihood estimates. The examples are given below in the work.

[Vlasov V.A., Vlasova S.V., Tolokonsky A.O. Estimates when parametrical hypotheses are accepted and comparison them with maximum likelihood estimates. *Life Sci J* 2014;11(11s):210-213] (ISSN:1097-8135). http://www.lifesciencesite.com. 46

Keywords: sample values, estimate of parameter, parametrical hypothesis, criterion of optimality, variation of functional, consequences of errors, the maximum likelihood method, Pearson theorem

Estimates when parametrical hypotheses are accepted and comparison them with maximum likelihood estimates

The fact that taking into account weights of errors when parametrical hypotheses are being accepted results in difference of parameter's estimate and maximum likelihood estimate is presented here.

Finding the values of unknown parameters of distribution laws is a widespread practical task [1,2,3,4]. Let us consider for the example the estimation of exponential distribution parameter on the basis of experimental results. The maximum likelihood estimate $\hat{\lambda}$ for parameter λ will be the following:

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^{n} \xi_{i}},\tag{1}$$

where n – sample size, ξ_i – sampled values.

However there are always doubts about applying this estimate in a real practice when we have ξ_i as concrete values x_{i_i} obtained in the experiment. Therefore $\hat{\lambda}$ is the random variable and the true value of parameter is stayed unknown. More than that, when we have small sample size

considerable errors could appear in estimation of the true value of parameter.

There is another way to choose the value of unknown parameter on the basis of application of variation approach [5,6,7]. In this case we do not consider just a simple estimation of a value of the unknown parameter, but we pay main attention to making decision about the optimum value of parameter on the basis of experimental data analysis. This approach could be successfully applied for testing both simple and composite competing hypotheses about unknown parameter [7].

Let us consider a simple illustrative example, when the following optimum criterion is provided:

$$J = p_1[I_{11}(1-\alpha_1) + I_{12}\alpha_1] + p_2[I_{21}\alpha_2 + I_{22}(1-\alpha_2)]$$
, (2)

where:

$$\alpha_1 = \iint_W p_H(x_1, \dots, x_n) dx_1 \dots dx_n - \text{probability of}$$

type I error;

$$\alpha_2 = \iint_{\overline{W}} p_{\overline{H}}(x_1, ..., x_n) dx_1 ... dx_n - \text{probability}$$

of type II error;

W – critical region of hypothesis H, \overline{W} – the region opposite to critical region W;

 $p_H(x_1,...,x_n), p_{\overline{H}}(x_1,...,x_n)$ – the distribution laws of a sample in case of hypotheses H and \overline{H} ; $I_{11}, I_{12}, I_{21}, I_{22}$ – have a sense of weights (costs) of probable situations and characterize the quality of made decisions (actually they take into account the consequences of the made decisions).

As the critical region W is defined by partition of sample space, which is made by function $x_n = \varphi(x_1, ..., x_{n-1})$, the value of J is depend on this function and J is functional.

The statement (2) could be written as the following expression:

$$J = C_{0} + A\alpha_{1} + B\alpha_{2},$$

where
$$C_{0} = p_{1}I_{11} + p_{2}I_{22}, \qquad A = p_{1}(I_{12} - I_{11})$$

$$B = p_{2}(I_{21} - I_{22}), \text{ and therefore}$$

$$J = C + A \int dx_{1...} \int dx_{n-1} \int_{\varphi}^{\varphi} p_{II}(x_{1},...,x_{n}) dx_{n} + B \int dx_{1...} \int dx_{n-1} \int_{-\infty}^{\varphi} p_{II}(x_{1},...,x_{n}) dx_{n}$$

(3)

where symbol φ means function $x_n = \varphi(x_1, ..., x_{n-1})$ which is argument of the functional.

Let us calculate the variation δJ of functional J according to the famous rule:

$$\delta J = \frac{\partial}{\partial \lambda} J(\varphi + \lambda \eta)$$
, when $\lambda = 0$ (here λ – is a

real argument, symbol η means variation of functional argument depending on x_1, \dots, x_{n-1}).

Differentiation is being done on upper limit of the integral:

$$\delta J = \int [-Ap_H(x_1, \dots, x_{n-1}, \varphi) + Bp_{\overline{H}}(x_1, \dots, x_{n-1}, \varphi)] \eta(x_1, \dots, x_{n-1}) dx_1, \dots, dx_{n-1}$$

Using the main lemma of variation calculation we will obtain the extremum conditions for the functional:

$$p_{\overline{H}}(x_1,\dots,x_{n-1},\varphi) = Cp_H(x_1,\dots,x_{n-1},\varphi), \quad (4)$$

where $C = \frac{A}{B}$.

Solution of the equation (4) for variable φ will give us the expression for partition boundary of the sample space that is similar to the boundary defined according to the Pearson's theorem [1]. However, at the variation approach the value of *C* depends not only on the probabilities of type I and II errors, but also on the costs (weights) of errors $I_{11}, I_{12}, I_{21}, I_{22}$.

Note that the constant C_0 could be omitted at the construction of critical region and the quality functional of statistical testing (3) could be more simplified: $J = A\alpha_1 + B\alpha_2$. In general the expression of optimum criterion for statistical test should be chosen on the basis of practical expediency.

The constant C could be calculated from the optimum conditions of statistical test using quality characteristics $I_{11}, I_{12}, I_{21}, I_{22}$ of made decisions. It is clear that the value of the constant Ccompletely defines optimum values of probabilities of type I and II errors because the partition of sample space was constructed using the boundary (4). On the basis of Pearson's theorem it is not possible to choose the optimum couple of probabilities of type I and II errors – it only gives the chance to minimize the probability of one type of error when we preliminary have the value of other type of error. On the contrary, using the variation approach we can at all not be interested in received probabilities of type I and II errors.

Let us show on the example of exponential law that maximum likelihood estimate of unknown parameter generally speaking should not coincide with accepted parametrical hypothesis. Using the expression (4) we will consider competing of two simple hypotheses: $\lambda = \lambda_1$ (this is hypothesis H) and $\lambda = \lambda_2$ (this is hypothesis \overline{H}). For clarity we suppose that $\lambda_1 < \lambda_2$. In the simplest case, when the values $I_{11}, I_{12}, I_{21}, I_{22}$ are constant (do not depend on the errors), variation approach will leads to critical region defined by the expression:

$$\prod_{i=1}^{n} \lambda_2 \exp(-(\lambda_2 x_i)) \ge C \prod_{i=1}^{n} \lambda_1 \exp(-(\lambda_1 x_i))$$

According to the expression (4) the boundary of critical region will be described by the equation:

$$\lambda_2^n \exp(-(S\lambda_2)) = C\lambda_1^n \exp(-(S\lambda_1)),$$

where $S = \sum_{i=1}^n x_i$.

Note that $S = \sum_{i=1}^{n} \xi_i$ is sufficient statistics

and for accepting or rejecting one of the competing hypothesis the boundary value should be imposed [7]:

$$S_{g} = \frac{1}{\lambda_{2} - \lambda_{1}} \left(n \ln \frac{\lambda_{2}}{\lambda_{1}} - \ln C \right).$$
 (5)

If $S < S_g$, then the one hypothesis is accepted, if $S > S_g$ the other hypothesis is accepted.

It is clear, that accepted parametrical hypotheses depend on a constant C and the accepted value of unknown parameter in general should not coincide with the experimental maximum likelihood estimate.

Let us consider a numerical illustration. We suppose that 100 products were tested and the sum of times of their non-failure operation was equal to $S = 10^5$ hours. Numerical value of failure rate received by maximum likelihood estimate is equal to $\widehat{\lambda}_0 = 10^{-3}$ per hour. We will suppose that one of the competing hypotheses \overline{H} claims that $\lambda = \lambda_2 = 10^{-3}$ 1/hour. Let us put forward the other competing hypothesis H (alternative) that states $\lambda = \lambda_1 = 10^{-4}$ 1/hour. Then it is possible to choose such values of a constant C when the hypothesis H could be accepted. Such values are defined by the equation: $S < S_g$, which could be expressed in more detailed formula:

$$S < \frac{1}{\lambda_2 - \lambda_1} (n \ln \frac{\lambda_2}{\lambda_1} - \ln C)$$
.

Actually according to the expression (5) the

boundary value S_g depends on a constant C. We suppose that in the considering example the constant C has such a value that the value S_g exceeds the experimental value $S = 10^{-5}$ 1/hour. For this purpose according to the expression (5) it is enough to set the constant C satisfying to an equation:

$$\frac{1}{\lambda_2 - \lambda_1} (n \ln \frac{\lambda_2}{\lambda_1} - \ln C) > 10^5.$$

In this case the more preferable will be the

value $\lambda = \lambda_2 = 10^{-4}$ 1/hour that, as it was expected, does not coincide with the experimental value of maximum likelihood estimate.

Such analysis of experimental data allows us to move the hypothesis H_2 towards the hypothesis H_1 until the situation when the hypothesis H_2 could be accepted, and this could be treated as a method for receiving an estimate of unknown parameter, if hypothesis H_1 defines acceptable region for unknown parameter [8].

Important feature of variation approach is that there is no need in creating confident intervals and no problem of a choice of the value of confident probability. Particularly it simplifies the analysis of experimental data when expressions of maximum likelihood estimates are specific. For example, in case of limited testing time when products are being tested (time interval of testing is T), the maximum likelihood estimate for parameter λ is the following [4]:

$$\widehat{\lambda}_p = \frac{k}{\sum_{i=1}^k \xi_i + (n-k)T},$$

where k - the number of equipment failures that were registered during the time interval [0; T].

Testing of parametrical hypotheses in case of limited testing time could be performed according to the recommendations specified in work [9].

It is important to emphasize that the variation approach allows us to take into account the weights (costs) of errors as a function of values of these errors at accepting of the parametrical hypothesis [6, 10].

In conclusion it is important to say that in general the most correct choice of the value of unknown parameter should be made on the basis of experimental data analysis including the effects from the type I and II errors. That is exactly what the variation approach allows us to achieve with using the criterion of optimality for choice of the most rational decision.

Corresponding Author:

Dr. Vlasov Victor Aleksandrovich

National Research Nuclear University «MEPhI» (Moscow Engineering Physics Institute)

Kashirskoye, 31, Moscow, 115409, Russian Federation

References

- 1. Rao, C.R., 2002. Second Edition. Linear Statistical Inference and its Applications, New York-London-Sydney, pp:468
- Seber, G.A.F., 1977. Linear Regression Analysis, Wiley Series in Probability and Mathematical Statistics, New York-London-Sydney-Toronto. pp:456.
- Vlasov, V.A., A.O. Tolokonsky and S.V. Vlasova, 2009. Statistical Analysis of Accuracy of Algorithms for Data Processing. Proceedings of the 11th International Workshop on Computer Science and Information Technologies, Crete, Greece, October 5-8, Volume 1.
- 4. Wilks, S.S., 1961. A combinatorical tests for the problem of two samples from continuous distribution. Proc Fourth Berkeley Symp. On Math. Stat. and Prob., Vol.1. University of California Press.

- 5. Shilov, G.E., 1961. Mathematical analysis. Special course. State edition physical and mathematical literature. Second Edition. Moscow.
- 6. Èlsgolts, L.E., 2001. Differential equations and the calculus of variations. Fourth edition. Editorial.
- 7. Vlasov, V.A., 2012. Estimates, Decisions, Risks. Moscow: BINOM.
- 8. Pronichev, A.N. And E.Yu. Berdnikovich, 2011. The Analysis of Adequacy of Algorithms for Making Decisions in Automatic Medical Systems using the Technique of Checking

6/30/2014

Simple Competing Hypothesis of Parameter. Instruments and Systems: Monitoring, Control and Diagnostics, 12.

- 9. Vlasov, V.A. and S.V. Vlasova, 2013. Check of Parametrical Hypotheses about the Failure Rate in Case of Limited Testing Time. Proceedings of the 15th International Workshop on Computer Science and Information Technologies, Vienna-Budapest-Bratislava, September 15-21, 2013, Volume 1.
- Gnedenko, B.V., Y.K. Belyayev and A.D. Solovyov, 1964. Mathematical Methods in Reliability Theory. Moscow: Science.