International Journal of Engineering & Technology, 7 (2.23) (2018) 292-294

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Finding interval estimates involving nuisance parameters

Nadezhda Chiganova *

A National research Moscow state university of civil engineering, 129337, Yaroslavskoye Shosse 26, Moscow, Russia
**Corresponding author E-mail: chiganovanm.mgsu@gmail.com*

Abstract

The work establishes the assertions which, in a number of cases, allow to effectively determine the possibility of applying the method suggested in the work of L.N. Bolshev and E.A. Loginov "Interval estimates in the presence of interfering parameters. Probability theory and its application" for constructing interval estimates of unknown parameters.

Keywords: *Distribution Function; Interval Estimation; Random Variables; Reliability Indices.*

1. Introduction

Suppose $X_1,...,X_n$ - are independent, identically distributed random vectors of dimension *m* with a distribution function $F(x;\theta)$ where $\theta=(\theta_1,...,\theta_k)$ is a vector of unknown parameters that takes on values in the domain θ.
[1] proposes a method for finding $X_1,...,X_n$ a confidence interval for the value of a given function $\varphi(\theta)$ from the unknown parameters θ, corresponding to the results of observations $X_1,...,X_n$ provided that the point estimate $\tilde{\varphi}=\tilde{\varphi}(X_1,...,X_n)$ of the function $\varphi(\theta)$ is known. This method is applicable in the case when the functions

$$g(y;z)=\inf_{\theta\in\phi(z)} G(y;\theta), \Im(y;z)=\sup_{\theta\in\phi(z)} G(y+0,\theta) \qquad (1)$$

are monotonic in regards to z for any fixed y. In these formulas $(z)=\{\theta\in\Theta:\varphi(\theta)=z\}$, and $G(y;\theta)$ - is the distribution function of the statistical estimate $\tilde{\varphi}$.
It is worth noting that it is not always easy to establish the monotonicity of the functions $g(y;z)$ and $\Im(y;z)$. It is shown below that the method suggested in [1], is applicable when the functions $\varphi(\theta)$ and $F(x;\theta)$ for all x are monotonous in regards to each of the components of the vector θ. It is possible to give examples of issues which are important on practical level (for this – see the Applications) in which the monotonicity of the functions $\varphi(\theta)$ and $F(x;\theta)$ is obvious.

2. Main findings

Suppose $f(\theta)$ and $\psi(\theta)$ - are some functions given in the domain Θ, and Ψ - is the possible range of the function $\psi(\theta)$. Suppose that the range Θ contains lowest θ_* (highest θ^*) point, if such $\theta_*\in\Theta(\theta^*\in\Theta)$ exists, that $\theta_*\le\theta$ $(\theta^*\ge\theta)$ for all $\theta\in\Theta$ (line $\theta'\le\theta''$ denotes the set of inequalities $\theta_1'\le\theta_1'',...,\theta_k'\le\theta_k''$). Let us introduce the following functions by analogy with (1) for the range Ψ.

2.1. Theorem 1

Suppose $\psi(\theta)$ - is a continuous non-decreasing (never-decreasing) function for all the coordinates of the vector θ, and Θ - is a convex set, containing lowest θ_* and highest θ^* points. Then, if the function $f(\theta)$ increases for all the coordinates of the vector θ, then the functions $f_*(z)$ and $f^*(z)$ on the range Ψ are non-decreasing (non-increasing); if the function $f(\theta)$ decreases for all the coordinates of the vector θ, then the functions $f_*(z)$ and $f^*(z)$ – are non-increasing (non-decreasing).
The assertion is provided only for the case when $\psi(\theta)$ is non-decreasing. Fix arbitrary $z_1,z_2\in\Psi, z_1<z_2$. Choose from the non-vacuous set $\{\theta:\psi(\theta)=z_2\}$ any θ'' element. Show that there is going to be such θ' that $\psi(\theta')=z_1,\theta'\le\theta''$.
Suppose $L=\{\theta:\theta_*+l(\theta''-\theta_*'),0\le l\le1\}$ - is an interval joining points $\theta_*\in\Theta$ and $\theta''\in\Theta$. Due to the convexity of Θ interval L is wholly contained in Θ, and on this interval $\psi(\theta)$ can be regarded as some function $\hat{\psi}(l),0\le l\le1$. Since $\psi(\theta)$ - is a continuous and non-decreasing for θ function, then the function $\hat{\psi}(l)$ exhibits the same properties, moreover, at the ends of the interval $\hat{\psi}(0)\le z_1<z_2\le\hat{\psi}(1)$. This implies that there is such $l_1(0\le l_1\le1)$, that $\hat{\psi}(l_1)=z_1$. Thus, it is shown that the point $\theta'=\theta_*+l_1(\theta''-\theta_*)$ can be chosen as a θ'.
Consequently, if $z_1<z_2$, then there are such θ' and θ'', that $\psi(\theta')=z_1,\psi(\theta'')=z_2$, moreover, $\theta'\le\theta''$, which means that $f_*(z_1)\le f_*(z_2)$, if $f(\theta)$ increases for $\theta(f_*(z_1)\le f_*(z_2))$, if $f(\theta)$

decreases for θ. The monotonicity of the function $f^*(z)$ is established in a similar way. The theorem is proven.
It can be easily seen that in order to prove the monotonicity of the function $f_*(z)(f^*(z))$, the existence of lowest $\theta_* \in \Theta$ (highest $\theta^* \in \Theta$) point suffices.
Essentially theorem 1 states that the $\psi(\theta)$ function requires continuity and monotonicity for all coordinates of the θ vector, moreover, if for the coordinate θ_i this function is non-increasing, then by changing it with $\theta_i^* = e^{\pm\theta_i}$ it can be transformed into non-decreasing.
It is of interest to establish sufficient conditions for the monotonicity of the function $G(y;\theta)$, expressed in terms of the statistic estimates $\tilde{\varphi}(X_1,...,X_n)$ and the assumed function $F(x;\theta)$.
Definition. Suppose that the set $D \subset R^n$ meets the condition (i), if in any plane sections $x_i = x_i^o$, j=1,...i-1,i+1,...,n there is a half-line $(x_1^o,...,x_{i-1}^o, x_{i+1}^o,...,x_n^o)$, $-\infty < x_i \geq x_i^o$, where $x_i^o = x_i^o(x_i^{*o})$, and $x_i^{*o} = (x_1^o,...,x_{i-1}^o,x_{i+1}^o,...,x_n^o)$.

2.2. Theorem 2

Suppose D - is a Borel set, meeting the condition (i), and $F_j(x_j;\theta_j)$ - is a non-decreasing function x_j and for all x_j decreasing (non-increasing) function $\Theta_j = (\theta_1^{(j)},...,\theta_k^{(j)})$. Then $G(\theta_1,...,\theta_k) = \int\underset{D}{...}\int \prod_{i=1}^{k} dF_i(x_i,\theta_i)$ – is a non-decreasing (non-increasing) function $\theta_j \in \Theta_j$.
Proof. Denote D_j as set D projection onto a subspace with coordinate axes $x_1,...,x_{j-1},x_{j+1},...,x_n$.
Then

$$G(\theta_1,...,\theta_k) = \int\underset{D}{...}\int \prod_{i\neq j} dF_i(x_i,\theta_i) \int_{-\infty}^{x_j(x_j^*)} dF_j(x_j,\theta_j) =$$
$$= \int\underset{D}{...}\int F_j(x_j(x_j^*);\theta_j)\prod_{i\neq j} dF_i(x_i,\theta_i)$$

This implies that the monotonicity of the function G for θ_j depends solely on the function $F_j(x_j,\theta_j)$. The proof is complete.
If $h(X_1,...,X_n)$ is a non-decreasing function $X_1,...,X_n$, defined in the space R^n, then for any $z \in h(R^n)$ the dimension $D(z)=\{(X_1,...,X_n) \in R^n : h(X_1,...,X_n) < z\}$ meets the condition (i).
Conclusion. Suppose D - is a Borel set, meeting the condition (i). If every function $F_i(x_i,\theta)$ does not decrease for the first variable x_i and for any x_i it does not decrease (does not increase) for $\theta \in \Theta$, then $G(\theta) = \int\underset{D}{...}\int \prod_{i=1}^{n} dF_i(x_i,\theta)$ does not decrease (does not increase) for θ.
The validity of this conclusion is obvious, since, according to Theorem 2, the function $G(\theta^1,\theta^2,...,\theta^n) = \int\underset{D}{...}\int \prod_{i=1}^{n} dF_i(x_i,\theta^i)$ defined on the set $\Theta \times \Theta \times ... \times \Theta = \Theta^n$ does not decrease (does not increase) for each set of variables $\theta^i \in \Theta$. It is clear that this property will be valid for $\theta^1 = ... = \theta^n = (\theta_1,...,\theta_k)$ as well.

2.3. Summary

Theorems 1-2 allow in many cases to justify the search for interval estimates by the method described in [1]. It is possible to illustrate this with two examples. It should be noted that the actual finding of the extremal points $\theta_1,...,\theta_k$ for the function $G(y,\theta)$ for the range $\Phi(z)$ is a complex computational problem which requires independent exploration.

3. Application

3.1. Example 1

Interval estimation of the reliability of the system based on the test results of its elements.
Suppose there is a system containing k of components in series. It is assumed that these units are operating independently and the failure of any element results in a system failure. Define p_i probability of failure-free operation of the i element for some fixed time t. Then the probability of failure-free operation of the system in time t will be determined by the formula $p=p_i...p_k$. Suppose that in order to estimate property p_i, the n_i of elements was tested and d_i of them did not fail for time t. The aim is to find an interval estimate for the reliability index of the system p using the information $(n_1,d_1),...,(n_k,d_k)$. A great number of publications and papers has been dedicated to that problem (see, for example [2], [3]).
Choose according to [1] statistics estimate $\tilde{p}=\tilde{p}_1...\tilde{p}_k$, where $\tilde{p}_1 = d_i / n_i$ is a point estimation. The estimate distribution function is expressed by the formula $G(y;p_1,...,p_k) = \sum\prod_{i=1}^{k} b(x_i;n_i;p_i)$, where the summation is carried out for those values x_i for which $(x_i...x_k)/(n_1...n_k) < y$ and $b(x_i;n_i;p_i) = C_{n_i}^{x_i} p_i^{x_i}(1-p_i)^{n_i-x_i}$. In order to use the method described in [1] to determine interval estimate of the characteristic property p, it is necessary to verify the monotonicity for z functions.
$g(y;z)=\inf_{p=z} G(y;p_1,...,p_k)$, $\Im(y;z)=\sup_{p=z} G(y+0;p_1,...,p_k)$.
Let us show with the help of theorems 1 and 2 that these functions decrease for z with any y. Note at the beginning that due to the [3], distribution function of the random variable d_i is expressed through formula $F(x_i;p_i) = \frac{1}{B(n_i-[x_i]+1,[x_i]+1} \int_0^{1-p_i} \xi^{n_i-[x_i]}(1-\xi)^{[x_i]} d\xi$, where $[x_i]$ – is the integral part of the number x_i, B – is a beta-function. This indicates that $F(x_i;p_i)$ – is a decreasing function p_i. Since the range x_i, which satisfies the inequality $(x_1...x_k)/(n_1...n_k) < y$, meets the condition (i), then according to the theorem 2, function $G(y;p_i,...,p_k)$ decreases for the variables $p_i,...,p_k$. Since, in addition to the above mentioned, function $p=p_i,...,p_k$ increases for all the variables, then according to the theorem 1, functions $g(y;z)$ and $\Im(y;z)$ decrease for z with any fixed y, what, as expected, coincides with the conclusion of [1].

3.2. Example 2

Interval estimation of the expected value of random variable distributed according to the Weibull law.
Suppose $X_1,...,X_n$ – are independent, identically distributed random variables, subject to Weibull distribution $F(x;\theta_1;\theta_2) = \begin{cases} 0, x \leq 0 \\ 1-\exp\left\{-\left(\frac{x}{\theta_2}\right)^{\frac{1}{\theta_1}}\right\}, x > 0 \end{cases}$, θ_1 and θ_2 - are unknown parameters). It is required to find interval estimation of the expected value $m=Mx_i$ according to these data. Choose the statis-

tics estimate $\tilde{X}=(X_1,...,X_n)/n$ as an estimate for m. Then $G(y;\theta_1,\theta_2)=\int_{\tilde{X}<y}...\int\prod_{i=1}^{n}dF(x;\theta_1,\theta_1)$.

It is clear that the integration domain $X_1+...+X_n<ny$ meets the condition (i), and function $F(x;\theta_1,\theta_1)$ increases for all the variables. Conclusion of the theorem 2 implies that the distribution $G(y;\theta_1,\theta_2)$ also exhibits this property.

It is known that the value m is connected with the parameters θ_1 and θ_2 by the formula $m=\theta_2\Gamma(\theta_1+1)$. This implies that m is a monotonically increasing function θ_1 and θ_2. According to Theorem 1 interval estimate m can be found by the method described in [1], because $g(y;z)=\inf_{\theta_2\Gamma(\theta_1+1)=z}G(y;\theta_1,\theta_2)$, $\Im(y;z)=\sup_{p=z}G(y+0;p_1,...,p_k)$ - is a monotonically increasing functions for z.

4. Conclusions and discussion

The above-mentioned examples show that the estimation of the statistical characteristics of parameters, processes, and any measured experimental information of the technical objects is an integral procedure in determining their accuracy and reliability. In engineering practice, the need to solve the estimation problems involving nuisance parameters often arises.

In terms of traditional mathematics, there are no models by which it is possible to accurately reflect the fuzziness of the original data. This fuzziness is usually associated with the interval uncertainty of describing a particular technical parameter.

The proposed mathematical apparatus allows to approach the classical problem of estimating such numerical characteristics as expected value and variance in a flexible way and, on their basis, to carry out a complex evaluation of the influence of the given parameters on the object under study.

As a result, it is possible to indicate confidence limits for the average and variance from one observation and drastically reduce the amount of worked out calculations.

References

[1] Bolshev LN & Loginov EA (1966), Interval estimates in the presence of interfering parameters, *Probability theory and its application* XI, 94-107.

[2] Shiryaev AN, Ehrlich IG & Oskov PA (2013), *Probability in theorems and problems (with proofs and solutions)*, MCCME, 648.

[3] Chiganova NM (2015), Logarithmic convexity with respect to the parameter of some distributions, *Journal of Natural and Technical Sciences* 6, 49-52.

[4] Bolshev LN & Smirnov EV (2012), *Tables of Mathematical Statistics*, Nauka, 416.

[5] Gnedenko BV, Belyaev YuK, Soloviev AD & Kashtanov VA (2013), *Mathematical Methods in Reliability Theory*, Book House "LIBROKOM", 550.

[6] Ushakov IA (2007), Reliability: Theory & Applications, *Journal of International Group on Reliability* 1 (1), 6–19

[7] Chiganova NM (2016), Reliability theory application for building structures reliability determination, *MATEC Web of Conferences* 86, 02009, https://doi.org/10.1051/matecconf/20168602009

[8] Gnedenko BV (1983), *Aspects of Mathematical Theory of Reliability (Russian)*, Radio I Svyaz, 376.

[9] Medvedev VV & Chiganova NM (2015), Evaluation of the reliability of products based on the results of software tests, *Scientific Review (Russia)* 14, 232-236.