12th International Research/Expert Conference
”Trends in the Development of Machinery and Associated Technology”
TMT 2008, Istanbul, Turkey, 26-30 August, 2008

ASSESSMENT OF WORKPIECE DIMENSIONS BY USE OF BAYESIAN CONFIDENCE INTERVAL

Dževad Zečić,
University of Zenica,
Faculty of Mechanical Engineering,
Bosnia & Herzegovina

Almir Huskanović,
University of Zenica,
Faculty of Mechanical Engineering,
Bosnia & Herzegovina

Hermina Alajbegović,
University of Zenica,
Faculty of Mechanical Engineering,
Bosnia & Herzegovina

ABSTRACT
The Bayesian confidence interval method provides a wide spectrum of application in different fields, including mechanical engineering. With aforegiven number $\gamma, 0<\gamma<1$ *called reliability, an interval is obtained* (a,b) *for assessment of machine operation accuracy. The interval is defined on the basis of aposteriory probability distribution of an unknown parameter. In this calculation some standard statistical methods such as beta distribution and Chebyshev inequality are used. The workpiece quality assessment method will be shown in an example taken from industrial production.*
Key words: confidence interval, beta distribution, reliability, Bayesian method

1. INTRODUCTION
One of the most commonly used distributions in statistics is Beta distribution, which is determined by its probability density function

$$f(x)=\begin{cases} 0, & x\le 1 \text{ i } x\ge 1 \\ \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & 0<x<1 \end{cases}, \qquad (1)$$

where $$\Gamma(t)=\int_0^{\infty} x^{t-1}e^{-x}dx \ \ (t>0) \qquad (2)$$

is Gamma function.

Let us assume that larger product series is manufactured. We wish to establish which percentage of manufactured products is within desirable dimensions, by taking sample of *n* products, where *k* is a number of products with desirable dimensions. It is clear that $k\in\{0,1,...,n\}$.

2. THE BAYESIAN CONFIDENCE INTERVAL

Assuming that we have measured data $x_1, x_2, ..., x_n$ of a random variable *X*, that its probability distribution is known and we want to estimate value of a parameter *t* by these data. In the problem described above, $t = \frac{k}{n}$ is a percent (if *t* is multiplied by 100) of precisely manufactured products.

Π distribution, which describes behavior of the parameter *t,* is also a priori known. That distribution will be presented by random variable *T*. On the other side, we have random variable $Y = \varphi(x_1, x_2, ..., x_n)$. In that case probability distribution of a random vector (T, Y) is determined by Π distribution and by conditioned probability distributions of *Y*. We are interested in a conditioned probability distribution of *T,* where $Y = y$. Let us specify random variable which presents that probability distribution as a T_y.

If $\gamma, 0 < \gamma < 1,$ is in advance given number, we look for such numbers *a* and *b*, *a* < *b*, with

$$P\left(a < T_y < b\right) = \gamma \tag{3}$$

We say that (*a*, *b*) is the Bayesian Confidence Interval of an unknown parameter *t* with γ confidence. We also say that the interval (*a*, *b*) is defined on the basis of aposteriory probability distribution of an unknown parameter *t*.

Probability distribution of random vector (T, Y) is defined by its density function

$$f(t,k) = f_1(t)P(Y = k / T = t). \tag{4}$$

At the same time f_1 is density probability function of the uniform distribution *U* (0,1), therefore

$$f_1(t) = \begin{cases} 1, \ t \in (0,1) \\ 0, \ t \notin (0,1) \end{cases} \tag{5}$$

while, $P(Y = k / T = t) = \binom{n}{k} t^k (1-t)^{n-k}, \ k = 0,1,2,...,n$ (6)

Consequently, $f(t,k) = \begin{cases} \binom{n}{k} t^k (1-t)^{n-k}, \ 0 < t < 1, \ k = 0,1,2,...,n \\ 0, \ t \notin (0,1) \end{cases}$ (7)

then $P(Y = k) = \binom{n}{k} \int_0^1 t^k (1-t)^{n-k} dt = \frac{1}{n+1}, \ (k = 0,1,...,n)$ (8)

Distribution of the random variable T_y is defined by the density function

$$\overline{f}(t) = \frac{f(t,k)}{P(Y = k)} = \begin{cases} (n+1)\binom{n}{k} t^k (1-t)^{n-k}, \ 0 < t < 1 \\ 0, \qquad t \notin (0,1) \end{cases} \tag{9}$$

Comparing this with the density function of Beta distribution (1), some congruency may occur if $\alpha = k+1, \beta = n-k+1.$ Therefore, unknown limits *a* and *b* of the Bayesian confidence interval could be defined from integral equations

$$\int_0^b \overline{f}(t)dt = \int_a^1 \overline{f}(t)dt = \frac{1-\gamma}{2} \tag{10}$$

for given confidence coefficient γ. Nevertheless, it is pretty complicated, so *a* and *b* are to be calculated approximately by using of the Chebyshev inequality.

In fact, if *X* is random variable with $E(X)=\mu$ expectation, and with $Var(X)=\sigma^2$ variance, then for approximate ε > 0 following inequality

$$P(\mu-\varepsilon\sigma < X < \mu+\varepsilon\sigma) \geq 1-\frac{1}{\varepsilon^2}. \tag{11}$$

is valid.

For the density function (1) follows $\mu=\frac{\alpha}{\alpha+\beta}, \sigma=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$, (12)

or when $\alpha=k+1, \beta=n-k+1,$ is incorporated, then for random variable T_y we get that expectation and variance is

$$\mu_y=\frac{k+1}{n+2}; \ \sigma_y=\frac{(k+1)(n-k+1)}{(n+2)^2(n+3)}. \tag{13}$$

If in (2) we take $\varepsilon=\frac{1}{\sqrt{1-\gamma}}$, then follows $1-\frac{1}{\varepsilon^2}=\gamma$, so requested limits of the Bayesian confidence interval are

$$a=\mu_y-\frac{\sigma_y}{\sqrt{1-\lambda}}, b=\mu_y+\frac{\sigma_y}{\sqrt{1-\lambda}}. \tag{14}$$

This result can be improved when *n* is a relatively large number and $k\approx\frac{n}{2}$, since then Beta distribution can be substituted by normal one $N(\mu_y,\sigma_y{}^2)$, and following formulas are obtained

$$a=\mu_y-z_\gamma\sigma_y, b=\mu_y+z_\gamma\sigma_y, \tag{15}$$

where, z_γ is the result of the equation $\Phi(z_\gamma)=\frac{\gamma}{2}$, and Φ is a Laplace's function. Number z_γ can be found from the Laplace's function tables.

3. APPLICATION OF CONFIDENCE INTERVAL

Example: The series of 1000 products is manufactured. To verify if the products are manufactured in desired dimensions, sample of $n=100$ is taken, where $k=70$ products are manufactured in desired dimensions. If one wishes to get the Bayesian interval with 95% of confidence, from Laplace's function tables $\Phi(1{,}96)=\frac{0{,}95}{2}$ is found, therefore $z_\gamma=1{,}96.$ Hence, according to the formula (13):

$\mu_y=\frac{71}{102}=0{,}696$ and $\sigma_y=\frac{71\cdot 31}{102^2\cdot 103}=0{,}002$,

and formula (15): $a=0{,}696-1{,}96\cdot 0{,}002=0{,}69208$ and $b=0{,}696+1{,}96\cdot 0{,}002=0{,}69992$.

Consequently, with the 95% confidence, it could be argued that in observed series of products, between 69,21% and 69,99% of products are manufactured within desired dimensions.

4. CONCLUSION

The Bayesian Probability Theory, named after Thomas Bayes (1702. – 1761.) is one of the theories which are still being developed and further explored based on the original ideas of the founder. The Bayesian Confidence Interval is just an example of one of these ideas. Unlike the frequentist confidence intervals, this one is based on of aposteriory probability, while the common confidence intervals are mainly based on known (measured) parameters. The Bayesian confidence interval method is sometimes used in elections voting results assessment, as well as in other situations. Accordingly, application of this method is also possible in problems concerning production, which is illustrated in this article.

5. REFERENCES

[1] A. Gelman, John B. Carlin, Hal S. Stern, Donald B. Rubin; Bayesian Data Analysis, Chapman & Hall / CRC, Washington, 2004.

[2] S. M. Lynch; Introduction to Applied Bayesian Statistics and Estimation for Social Scientists; Springer, New York, 2007.

[3] William M. Bolstad; Introduction to Bayesian Statistic, Wiley Interscience, New Jersey, 2004.

[4] Ž. Pauše; Uvod u matematičku statistiku, Školska knjiga, Zagreb, 1993.

[5] S. K. Kopparapu, U.B. Desai; Bayesian Approach to Image Interpretation, Kluwer Academic Publishers, New York / Boston, 2002.