Approximation of distribution law of experimental test data to assess reliability of information-measuring and control systems

Yu T Zyryanov¹, I G Ryazanov¹, A Yu Naumova¹, D Yu Muromtsev¹, N G Chernyshov¹ and E V Trapeznikov²

¹Tambov State Technical University, 106, Sovetskaya Str., Tambov, 392000, Russia
²Omsk State Technical University, 11, Mira ave., Omsk, 644050, Russia

E-mail: crems@mail.jesby.tstu.ru

Abstract. Methods of approximation for the distribution law of experimental data find wide application in problems of reliability assessment in tests of both hardware and software of software modules of information-measuring and control systems (SM IMCS). Thus, the main task is to solve the problem of increasing the accuracy of the approximation and simplify experimental data processing. In this paper various modification types of Pearson distribution for data taking both positive and negative values, and for data taking only positive values are considered. The described modifications allow one to solve a range of problems related to the existing distributions, as well as to improve the simplicity of the approximation procedure. Pearson logarithmic distributions are proposed for experimental data that take only positive values.

1. Introduction

In modern classical mathematical statistics, the form of the distribution law is considered a priori known and the main task is to estimate the key parameters of the law using experimental observational data. But, in practice, the law of distribution is very difficult to identify, and theoretical assumptions do not provide the necessary accuracy for its unambiguous definition. The process of experimental data analysis, in turn, also does not improve the accuracy of calculation and determination of the true distribution law. In real conditions it is necessary to resort to the approximation (approximate description) of the desired law by another one, which is correlated with experimental data and which has similar parameters to the unknown true law. In solving a wide range of applied problems (for example, in assessing the reliability of technical systems [1]), the use of approximate identification of the experimental distribution laws parameters is allowed [2, 3].

The purpose of the paper is to consider the approximation features of the distribution law of experimental test data, both hardware and software of SM IMCS using modified Pearson distributions, which can take:

– positive and negative values;
– only positive values.

2. Problem statement
For experimental test data the problem of approximation is formulated in the following form: if there is an initial sample of data \((x_1, x_2, \ldots, x_n)\) in the volume of \(n\), with a random variable (RV) \(X\), it is necessary to determine the parameters of the distribution law (type and parameters), which does not contradict in the statistical sense the available initial experimental data. Existing restrictions are as follows. The sample is representative; its volume \(n \geq 10000\) is sufficient to estimate the expectation \(m_1\) and the central moments \(\mu_2, \mu_3, \mu_4\); the form of the probability distribution density (PDD) \(p(x)\) is unimodal, \(J\)-shaped or \(U\)-shaped. The solution of the problem is based on the use of modified Pearson distributions [4-6].

3. Theory

The expressions that will determine the skew coefficient and kurtosis of RV \(X\) are as follows:

\[
\beta_1 = \mu_3 / \mu_2^{1.5}; \quad \beta_2 = \mu_4 / \mu_2^2
\]  

(1)

On the basis of [4, 6] we transform the modified Pearson differential equation for the system of probability distribution densities \(p(x)\)

\[
\frac{dp(x)}{dx} = \frac{(4\beta_2 - 5)(x - m_1) - 0.5\beta_1 \sqrt{\mu_2}}{(1 - \beta_2)(x - m_1)^2 + 0.5\beta_1 \sqrt{\mu_2} (x - m_1) + (2 - \beta_2) \mu_2} p(x).
\]

(2)

The expression that describes the normalized joint skew coefficient and kurtosis \(\beta_{12}\) in (2) is presented below

\[
\beta_{12} = \frac{1.5 \beta_2^2 + 6}{\beta_2^2 + 3}.
\]

(3)

It is worth noting that the range of possible values \(\beta_{12}\) lies within the limits \(0 \leq \beta_{12} \leq 1.5\), in its upper bound \(\beta_{12}\) they correspond to the distribution as a set of two Delta functions, and in the lower bound they correspond to the distribution having less than four moments. Depending on the values of the coefficients \(\beta_1\) and \(\beta_{12}\), as well as the auxiliary coefficient

\[
\beta_{21} = \frac{\beta_1}{4(1 - \beta_{12}^2)(2 - \beta_{12})}
\]

(4)

as a solution of the differential equation (2), six modified types of distributions corresponding to twelve types of Pearson classification distributions are obtained. Table1 shows the formulas expressing the coefficients \(\beta_1, \beta_{12}, \beta_{21}\) using the shape parameters of the modified Pearson distributions; in table 2 expressions for the PDD and in table 3 expressions for the shape, scale and shift \(\mu\) parameters of these PDD are shown.

**Table 1.** Numerical characteristics of modified pearson distributions.

| Type | \(\beta_1\) | \(\beta_{12}\) | \(\beta_{21}\) |
|------|-------------|----------------|---------------|
| 1    | \(\frac{2(\nu - \alpha) \sqrt{\nu + \alpha + 1}}{\sqrt{\alpha} (\nu + \alpha + 2)}, \)\(-\infty < \beta_1 < \infty, \) \(1 < \beta_{12} < 1.5,\) \(-\infty < \beta_{21} < \infty\) | \(\frac{\nu + \alpha + 3}{\nu + \alpha + 2}, \) | \(\frac{\nu - \alpha}{2\sqrt{\alpha\nu}}, \) |
| 2a   | \(\beta_1 = -2/\sqrt{\alpha},\) | \(\beta = 1,\) | \(\beta \rightarrow -\infty\) |
| 2b   | \(\beta_1 = 2/\sqrt{\alpha},\) | \(\beta_{12} = 1, \) | \(\beta_{21} \rightarrow \infty\) |
| 3    | \(\beta_1 = 0\) | \(\beta_{12} = 1,\) | \(\beta_{21} = 0\) |
Table 2. Modified Pearson distributions.

| Type of distribution | An analytical expression for PDD $p(x)$ |
|----------------------|------------------------------------------|
| 1                    | $p(x) = \frac{(x-\mu)^{\alpha-1}}{\chi^{\nu-1}B(\alpha,\nu)}$, $\mu < x < \chi + \mu$, $\alpha > 0, v > 0$, $-\infty < \chi < \infty$, $-\infty < \mu < \infty$: $B(a,b)$ – beta function. |
| 2a                   | $p(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}(\mu-x)^{\alpha-1}\exp[-\lambda(\mu-x)]$, $-\infty < x < \mu$, $\alpha > 0, \lambda > 0$, $-\infty < \mu < \infty$: $\Gamma(z)$ - gamma function. |
| 2b                   | $p(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}(x-\mu)^{\alpha-1}\exp[-\lambda(x-\mu)]$, $\mu < x < \infty$, $\alpha > 0, \lambda > 0$, $-\infty < \mu < \infty$. |
| 3                    | $p(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$, $-\infty < x < \infty$, $\sigma > 0$, $-\infty < \mu < \infty$. |
| 4a                   | $p(x) = \frac{\lambda^\nu(\mu-x)^{\alpha-1}}{B(\alpha,\nu)(\lambda+\mu-x)^{\alpha+v}}$, $-\infty < x < \mu$, $\alpha > 0, v > 0, \lambda > 0$, $-\infty < \mu < \infty$. |
| 4b                   | $p(x) = \frac{\lambda^\nu(\mu-x)^{\alpha-1}}{B(\alpha,\nu)(\lambda+x-\mu)^{\alpha+v}}$, $\mu < x < \infty$, $\alpha > 0, v > 0, \lambda > 0$, $-\infty < \mu < \infty$. |
Table 3. Parameters of modified Pearson distributions.

| Type of distribution | Analytical expressions for the parameters of PDD \( p(x) \) |
|----------------------|---------------------------------------------------------------|
| 1                    | \( \alpha = \left( \frac{0.5}{\beta_{12} - 1} \right)^{-1} \left( 1 - \frac{\beta_{21}}{\sqrt{\beta_{21}^2 + 1}} \right) \), \( \nu = \left( \frac{0.5}{\beta_{12} - 1} \right)^{-1} \left( 1 + \frac{\beta_{21}}{\sqrt{\beta_{21}^2 + 1}} \right) \), \( \chi = 2 \sqrt{\frac{2 - \beta_{21}}{\beta_{21} - 1}} \mu_2 \left( \beta_{21}^2 + 1 \right) \), \( \mu = m_1 + \frac{2 - \beta_{21}}{\beta_{21} - 1} \mu_2 \beta_{21} \). |
| 2a                   | \( \alpha = \frac{4}{\beta_1^2} \), \( \lambda = - \frac{2}{\beta_1 \sqrt{\mu_2}} \), \( \mu = m_1 - \frac{2 \sqrt{\mu_2}}{\beta_1} \). |
| 2b                   | \( \alpha = \frac{4}{\beta_1^2} \), \( \lambda = \frac{2}{\beta_1 \sqrt{\mu_2}} \), \( \mu = m_1 - \frac{2 \sqrt{\mu_2}}{\beta_1} \). |
| 3                    | \( \sigma = \sqrt{\mu_2} \), \( \mu = m_1 \). |
| 4a                   | \( \lambda = 2 \sqrt{\left( 1 + \frac{1}{1 - \beta_{12}} \right) \left( \beta_{21}^2 - 1 \right) \mu_2} \), \( \mu = m_1 - \frac{\beta_1 \sqrt{\mu_2}}{2 (1 - \beta_{12})} \). |
| 4b                   | \( \lambda = 2 \sqrt{\left( 1 + \frac{1}{1 - \beta_{12}} \right) \left( \beta_{21}^2 - 1 \right) \mu_2} \), \( \mu = m_1 + \frac{\lambda}{2} - \frac{\beta_1 \sqrt{\mu_2}}{2 (1 - \beta_{12})} \). |
| 5a                   | \( \nu = 3 + \frac{1}{1 - \beta_{12}} \), \( \lambda = \left( 2 + \frac{1}{1 - \beta_{12}} \right) \beta_1 \sqrt{\mu_2} \), \( \mu = m_1 - \frac{\beta_1 \sqrt{\mu_2}}{4 (1 - \beta_{12})} \). |
Using the diagram in the plane of variables $\beta_1$ and $\beta_2$ shown in figure 1, it is convenient to make a topographic classification of the modified Pearson distributions.

![Diagram](image)

**Figure 1.** Diagram of the modified Pearson distributions.

### 4. Experimental results

In the case of processing experimental data that take only positive values, the approximation problem will be presented in the form where the true distribution law of RV $X$ is often skew-symmetric and it does not have positive power moments. In view of the above, for its approximation according to available data, it is possible to use logarithmic moments (estimates of mathematical expectation $m_1$ and central moments $L_2$, $L_3$, $L_4$) instead of power moments. As a result of the functional transformation $z = \exp(x)$ the distributions represented in table 2 after some transformations of parameters will take the form shown in table 4. For simplicity, we call them Pearson logarithmic distributions. The parameters of these distributions are related to the numerical characteristics $l_1$, $L_2$, $\beta_1$, $\beta_2$, and $\beta_{21}$ of
RV $Z$ with the use of the expressions presented in table 5. In this case, the skew coefficient and kurtosis ($\beta_1$ and $\beta_2$) are determined by the formulas similar to (1):

$$
\beta_1 = \frac{L_3}{L_2^3}, \quad \beta_2 = \frac{L_4}{L_2^4}.
$$

(5)

**Table 4.** Logarithmic Pearson distribution.

| Type of distribution | Analytical expression for PDD $p(z)$ |
|----------------------|--------------------------------------|
| 1                    | $p(z) = \frac{1}{B(\alpha, \nu)} \left( \ln \left( \frac{z^2}{\chi_1} \right) \right)^{1-\alpha} \left( \ln \left( \frac{z}{\chi_1} \right) \right)^{\alpha-1} \left( \ln \left( \frac{2z}{\chi} \right) \right)^{\nu-1}, \chi_1 < z < \chi_2,$$
|                      | $\alpha > 0, \nu > 0, \chi_1 > 0, \chi_2 > 0.$ |
| 2a                   | $p(z) = \frac{\nu^\alpha}{\Gamma(\alpha)z^\nu} \left( \ln \left( \frac{z}{\chi} \right) \right)^{\alpha-1}, \ 0 < z < \chi,$$
|                      | $\alpha > 0, \nu > 0, \chi > 0.$ |
| 2b                   | $p(z) = \frac{\nu^\alpha}{\Gamma(\alpha)z^\nu} \left( \ln \left( \frac{z}{\chi} \right) \right)^{\alpha-1}, \chi < z < \infty,$$
|                      | $\alpha > 0, \nu > 0, \chi > 0.$ |
| 3                    | $p(z) = \frac{\gamma^\alpha}{\sqrt{2\pi}z^\nu} \exp \left[ -\left( \ln \left( \frac{z}{\nu} \right) \right)^2 \frac{\nu^2}{2} \right], \ 0 < z < \infty,$$
|                      | $\gamma > 0, \nu > 0.$ |
| 4a                   | $p(z) = \frac{\gamma^\alpha}{B(\alpha, \nu)z^\nu} \left( \ln \left( \frac{z}{\chi} \right) \right)^{\alpha-1} \left( 1 + \gamma \cdot \ln \left( \frac{z}{\chi} \right) \right)^{-(\alpha + \nu)}, \ 0 < z < \chi,$$
|                      | $\alpha > 0, \nu > 0, \gamma > 0, \chi > 0.$ |
| 4b                   | $p(z) = \frac{\gamma^\alpha}{B(\alpha, \nu)z^\nu} \left( \ln \left( \frac{z}{\chi} \right) \right)^{\alpha-1} \left( 1 + \gamma \cdot \ln \left( \frac{z}{\chi} \right) \right)^{-(\alpha + \nu)}, \chi < z < \infty,$$
|                      | $\alpha > 0, \nu > 0, \gamma > 0, \chi > 0.$ |
| 5a                   | $p(z) = \frac{\gamma^\nu}{\Gamma(\nu)z^\nu} \left( \ln \left( \frac{z}{\chi} \right) \right)^{-(\nu + 1)} \exp \left[ -\nu \cdot \ln \left( \frac{z}{\chi} \right) \right], \ 0 < z < \chi,$$
|                      | $\nu > 0, \gamma > 0, \chi > 0.$ |
| 5b                   | $p(z) = \frac{\gamma^\nu}{\Gamma(\nu)z^\nu} \left( \ln \left( \frac{z}{\chi} \right) \right)^{-(\nu + 1)} \exp \left[ -\nu \cdot \ln \left( \frac{z}{\chi} \right) \right], \chi < z < \infty,$$
|                      | $\nu > 0, \gamma > 0, \chi > 0.$ |
| 6                    | $p(z) = \frac{\gamma}{C z} \exp \left[ \frac{b \cdot \arctg \left( \gamma \cdot \ln \left( \frac{z}{\chi} \right) \right)}{1 + \left( \gamma \cdot \ln \left( \frac{z}{\chi} \right) \right)^2} \right]^{0.5(v+1)}, \ 0 < z < \infty,$$
|                      | $\nu > 0, \gamma > 0, -\infty < b < \infty, \chi > 0.$ |
Table 5. Parameters of logarithmic Pearson distributions.

| Type of distribution | Analytical expressions for the parameters of PDD \( p(z) \) |
|----------------------|----------------------------------------------------------|
| 1                    | \( \alpha = \left( \frac{0.5}{\beta_{21} - 1} \right) \left( 1 - \frac{\beta_{21}}{\sqrt{\beta_{21}^2 + 1}} \right), \quad \nu = \left( \frac{0.5}{\beta_{21} - 1} \right) \left( 1 + \frac{\beta_{21}}{\sqrt{\beta_{21}^2 + 1}} \right) \), \quad \chi = 2 \sqrt{\frac{2 - \beta_{21}}{\beta_{21} - 1}} L_2 \left( \beta_{21}^2 + 1 \right), \quad \mu = \frac{2 - \beta_{21}}{\beta_{21} - 1} L_2 \beta_{21}, \quad \chi_1 = \exp(\chi - \mu), \quad \chi_2 = \exp(\chi + \mu) \) |
| 2a                   | \( \alpha = 4/\beta_1^2, \quad \nu = 2/\sqrt{2} \beta_1 L_2, \quad \chi = \exp \left( \nu + 1/\sqrt{2} \beta_1 L_2 \right) \) |
| 2b                   | \( \alpha = 4/\beta_1^2, \quad \nu = 2/\sqrt{2} \beta_1 L_2, \quad \chi = \exp \left( \nu + 2/\sqrt{2} \beta_1 L_2 \right) \) |
| 3                    | \( \gamma = 1/\sqrt{L_2}, \quad m = \exp \left( t_1 \right) \) |
| 4a                   | \( \alpha = \left( 1 + \frac{0.5}{\beta_{21} - 1} \right) \left( -\frac{\beta_{21}}{\sqrt{\beta_{21}^2 + 1}} - 1 \right), \quad \nu = 3 + \frac{1}{\beta_{21} - 1} \), \quad \gamma = \left( \frac{\alpha (\alpha + \nu - 1)}{(v - 1)^2 (v - 2) L_2} \right), \quad \chi = \exp \left( \nu + \frac{\nu}{v - 1} \right) \) |
| 4b                   | \( \alpha = \left( 1 + \frac{0.5}{\beta_{21} - 1} \right) \left( \frac{\beta_{21}}{\sqrt{\beta_{21}^2 + 1}} - 1 \right), \quad \nu = 3 + \frac{1}{\beta_{21} - 1} \), \quad \gamma = \left( \frac{\alpha (\alpha + \nu - 1)}{(v - 1)^2 (v - 2) L_2} \right), \quad \chi = \exp \left( \nu - \frac{\nu}{v - 1} \right) \) |
| 5a                   | \( \nu = 3 + \frac{1}{\beta_{21} - 1}, \quad \gamma = (v - 1) \sqrt{(v - 2) L_2}, \quad \chi = \exp \left( \nu + \frac{\nu}{v - 1} \right) \) |
| 5b                   | \( \nu = 3 + \frac{1}{\beta_{21} - 1}, \quad \gamma = (v - 1) \sqrt{(v - 2) L_2}, \quad \chi = \exp \left( \nu - \frac{\nu}{v - 1} \right) \) |
| 6                    | \( \nu = 3 + \frac{1}{\beta_{21} - 1}, \quad b = \left( 2 + \frac{1}{\beta_{21} - 1} \right) \frac{\beta_{21}}{1 - \beta_{21}}, \quad \gamma = \left( \frac{(v - 1)^2 + b^2}{(v - 1)^2 (v - 2) L_2} \right), \quad \chi = \exp \left( \nu + \frac{b}{v - 1} \frac{1}{\gamma} \right) \) |

5. Results discussion

Thus, the task of determining the distribution law parameters of the initial experimental data, for positive and negative values, using the modified Pearson distributions is reduced to the implementation of four main steps:

1) using the sample to receive estimates of the mathematical expectation \( m_1 \) and the central power moments \( \mu_2, \mu_3, \mu_4 \) for RV \( X \);

2) to calculate numerical characteristics \( \beta_1, \beta_2, \beta_{12}, \beta_{21} \) of RV \( X \) by formulas (1), (3) and (4);

3) using table 1 and table 2 to determine types and analytical expressions for Pearson distribution.
on the obtained values of numerical characteristics $\beta_1$, $\beta_2$, and $\beta_3$ of RV $X$;

4) to calculate the parameters of the shape, scale and shift of the PDD $p(x)$ using the formulas given in table 3 for this type of PDD.

Sometimes, for example in, the lower bound of the applicability of Pearson distributions to approximate the distribution law is indicated, as is considered in. It is a line described by the equation:

$$8\beta_2 - 15\beta_1 - 36 = 0.$$ 

This limit is the normalized joint skew coefficient and kurtosis $\beta_{12} = 0.8$.

6. Conclusion

When processing experimental test data of both hardware and software of SM IMCS, it is problematic to determine the true distribution law of experimental test data. The paper presents one of the approaches to approximating the distribution law of experimental test data on the basis of their approximation. It is proposed to use approximate identification of parameters of experimental distribution laws to solve the problems of reliability estimation of SM IMCS. The purpose of the work was achieved by considering the features of approximation of the experimental test data distribution law, these data taking positive and negative values, or only positive values using modified Pearson distributions. Modification of the Pearson method with the use of coefficients $\beta_1, \beta_{12}, \beta_{21}$ can significantly simplify the procedure for approximation of experimental distributions. Pearson logarithmic distributions for positive values of experimental test data are also presented. The proposed parameters of the modified Pearson distributions are preferably used for the preliminary selection of the experimental data distribution law for reliability tests of the IMCS.

Acknowledgments

This work was supported by the Russian Foundation for Fundamental Research, Project № 17-08-00457-a.

References

[1] Karpov I G, Zyryanov Yu T and Petrov A V 2012 Generalized distribution laws for solving problems in the theory of reliability of information-measuring and control systems Information-measuring and control systems 10 (6) 37–41

[2] Zyryanov Yu T, Chernyshov N G, Ryazanov I G, Naumova A Y, Dioumessy M F 2019 Application of Generalized Distribution Laws for the Reliability Assessment of Information and Measuring Systems of Energy Saving Control Journal of Physics Conference Series 1172(1) 012108 DOI:10.1088/1742-6596/1172/1/012108.

[3] Mikhaylov A A, Bazuyeva S A 2017 Use of probabilistic splitting distributions of the original data during the sequential synthesis problem-solving algorithm. International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM 2017 - Proceedings electronic edition 8076402

[4] Karpov I G 2003 Approximations of experimental distributions of radar signals using modified Pearson distributions Radio engineering 5 56–61

[5] Karpov I G and Zyryanov Yu T 2015 On modified Pearson distributions and their identification Automatic Control and Computer Sciences 49 (6) 366–372

[6] Karpov I G and Gribkov A N 2014 Modernization of the distributions of the Pearson approximation of the bilateral distribu-tion laws of experimental data Proc. of the Tomsk Polytechnic University vol 324 chapter 2 pp 5–10