Research on the Determination of Reasonable Cost Range of Power Grid Production Technological Transformation Project

Xin Zhao¹, Tongyun Ni², Xujuan Miao¹ and Lijuan Guo*¹

¹Economy and Technology Research Institute, State grid Xin Jiang Electric Power Corporation, Urumqi, Xinjiang, 830011, China
²State grid Xin Jiang Electric Power Corporation, Urumqi, Xinjiang, 830011, China.
*Guo-1182206039@ncepu.edu.cn
*Corresponding author’s e-mail: guolijuanxinj@163.com

Abstract. This paper proposed a method to determine the reasonable cost interval of power grid production technological transformation project based on Kernel Density Estimation (KDE). The general idea of the study of reasonable cost interval is to select sample data, determine the sample estimation method, remove the outlier data, and finally calculate the reasonable unit cost interval of power grid. The general idea of the study of reasonable cost interval is to select sample data, determine the sample estimation method, remove the outlier data, and finally calculate the reasonable unit cost interval of power grid technical transformation project. Taking the actual data of the disconnector replacement project as an example, this paper made statistics of the reasonable cost range of the power grid production technological transformation project, which provides a new idea and method for the reasonable determination of the cost range in the case of large amount of data.

1. Introduction

With the continuous development of economy and society, the rapid development of China's power grid construction, the implementation of power grid production technological transformation project has a new era background. With the development and application of new technology, new equipment, new materials and new technology, the actual cost of technological transformation project deviates from the corresponding standards to a certain extent [1]. Because of the increasing demand of electric power, the investment of technical transformation of power grid equipment is increasing. As an important link of technological transformation project construction management, cost management directly affects the economic and social benefits of enterprises. Therefore, effective control of the cost of technological transformation projects and determination of a reasonable cost range will not only affect the economic interests of the project management unit, but will also affect the future development of power engineering construction. This article uses scientific statistical analysis methods as a means to determine the reasonable cost range of technological transformation projects, and is committed to achieving the deep unification of the cost standards of technological transformation projects, which has certain practical significance in promoting the controllability of project costs.
2. Nonparametric density estimation theory

In statistics, the probability density function of the sample can be estimated by parametric or non-parametric methods. The parameter density estimation method assumes that the distribution of the sample is known, and is a probability density function \( f(x|\theta) \) including the parameter \( \theta \), where \( \theta \) is unknown. After obtaining the optimal estimate \( \hat{\theta} \) of the parameter \( \theta \), the whole density function is estimated by \( f(x|\hat{\theta}) \). The advantage of parameter density estimation is that the distribution of samples is described in an accurate way through fewer parameters. The parameter expression of density provides an effective solution to the calculation of density. The biggest defect of parameter density estimation is that the distribution form of samples is often unknown, and parameter estimation needs to assume the distribution of samples first. If the assumed distribution of samples does not conform to the actual density distribution, it will cause a large estimation error.

Non-parameter estimation can make up for the above defects. It does not need to assume a specific distribution in advance, but only takes the data on the sample itself as the basis of probability density function estimation. Starting from a large amount of data itself, study the distribution characteristics of the data, so as to simulate the true probability density curve, so it has higher application value. Kernel Density Estimation is a very effective non-parametric estimation method \([3]\), and its general expression is as follow:

\[
 f(x) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{x - X_i}{h} \right)
\]

Where \( n \) is the total number of samples; \( h \) is the bandwidth or smoothing parameter; \( X_i \) is the given sample; \( K(g) \) is the kernel function. The selection of kernel function should meet the following conditions:

\[
\begin{align*}
 K(t) & \geq 0 \\
 \int_{-\infty}^{+\infty} K(t)dt & = 1 \\
 \sup_{t} K(t) & < +\infty \\
 \int_{-\infty}^{+\infty} K^2(t)dt & < +\infty \\
 \lim_{x \to +\infty} K(t) & = 0
\end{align*}
\]

After the kernel function is determined, the variance of the kernel function is determined by the bandwidth \( h \). \( h \) reflects the smoothness of the kernel density estimation curve. The smaller \( h \), the smaller the weight of the observed sample points in the final kernel density estimation curve. The steeper the nuclear density estimation curve; conversely, the greater the \( h \), the smoother the nuclear density estimation curve \([4]\).

3. Determination of reasonable interval of technical reform project cost based on KDE

In this paper, the Gaussian kernel function is selected as the kernel function for kernel density estimation. At this time, the expression of the probability density function \( f(x) \) is as follow:

\[
 f(x) = \frac{1}{\sqrt{2\pi} nh} \sum_{i=1}^{n} \exp\left[ -\frac{1}{2} \left( \frac{x - X_i^2}{h^2} \right) \right]
\]

If \( h \) is too small, the density estimation is too limited to the vicinity of the observed data, and many erroneous peaks are generated, which is difficult to interpret the data. If \( h \) is too large, some important feature information (such as multimodality) is lost, and unbiased estimation cannot be
guaranteed. At this point it is estimated to be meaningless. The kernel function estimation results corresponding to different $h$ are very different. In this paper, the calculation is based on the optimal bandwidth given by MATLAB.

The reasonable cost interval of the technical transformation project needs to be determined in conjunction with the given confidence level. After obtaining the probability density function $f(\varepsilon)$ of the unit cost of the technical transformation project based on the non-parametric kernel density estimation, the cumulative probability distribution function $F(\varepsilon)$ is obtained by the integral shown in formula (4).

$$F(\varepsilon) = \int_{-\infty}^{\varepsilon} f(\varepsilon) d\varepsilon$$

According to the cumulative probability distribution function, the confidence interval of the confidence level $1 - \alpha$ can be obtained.

4. Example analysis

Taking the disconnector replacement project in the power grid production technological transformation project in Xinjiang from 2016 to 2018 as an example, the actual completion data was used to verify the role of the non-parametric kernel density estimation method in determining the reasonable interval of the power grid production technical renovation project.

In this analysis, a total of 167 disconnector replacement projects were collected. Among them, 46 disconnector replacement projects were completed and put into production in 2018, and a total of 431 disconnectors were transformed, with a static investment of 9.533 million yuan; 72 disconnector replacement projects were completed and put into operation in 2017, and a total of 693 disconnectors were transformed, with a static investment of 15.9347 million yuan; 49 disconnector replacement projects were completed and put into operation in 2016, and a total of 462 disconnectors were transformed, with a static investment of 11.0986 million yuan. The statistical summary of the unit cost of replacement isolation switch engineering from 2016 to 2018 is shown in Figure 1.

![Figure 1. Statistics of unit cost of disconnector replacement projects.](image-url)
It can be seen from the above figure that the unit cost of disconnector replacement projects in 2016-2018 is up to 99,700 yuan/set and the lowest is 1100 yuan/set. This cost range cannot meet the needs of cost management accuracy.

The kernel density estimation can be used to obtain the probability density of the unit cost of the disconnector replacement projects, and then the cubic spline interpolation can be used to fit the cumulative probability distribution curve of the unit cost of the disconnector replacement projects. The estimated probability density curve of the unit cost of disconnector replacement projects is shown in Figure 2.

![Figure 2. Unit cost probability density curve.](image)

It can be seen from the above figure that the kernel density curve and the frequency histogram cover more area in the sample concentration area, and the kernel density curve is basically consistent with the normal distribution trend at the first and last ends of the sample, proving that the probability density interval retains its internal characteristics. At this time, the window width is 1.1169. The cubic spline interpolation theory is used to curve fit the probability distribution of the unit cost of the disconnector replacement project, and the confidence intervals when the confidence levels are 95%, 90%, 85%, and 80% are calculated, as shown in Figure 3.

![Figure 3. Unit cost interval at various confidence levels.](image)
In this paper, Interval Coverage Probability (ICP) is used to evaluate the reliability of the cost interval. The calculation formula is:

\[ ICP = \frac{1}{N} \sum_{i=1}^{N} \rho_i \]  

(3)

Among them, \( N \) is the total number of samples; \( \rho_i \) is a Boolean quantity, if the sample falls within the interval, then \( \rho_i = 1 \), otherwise \( \rho_i = 0 \). When \( ICP = 1 \), it means that all samples are within the cost interval, and the reliability is the highest at this time, but at the same time, the cost interval loses its practical significance. It is generally believed that ICP should be as close to the preset confidence level as possible. The interval at this time is more reasonable. The coverage of the cost interval at each confidence level is shown in Table 1.

| confidence level | Number of projects covered | ICP       |
|------------------|---------------------------|-----------|
| 80%              | 130                       | 77.84%    |
| 85%              | 143                       | 85.63%    |
| 90%              | 147                       | 88.02%    |
| 95%              | 156                       | 93.41%    |

It can be seen from the above table that the coverage of the cost interval at the 85% confidence level is closest to the confidence level, indicating that the cost interval at the 85% confidence level has high reliability, so the cost interval at the 85% confidence level is the most guaranteed. At this time, the cost interval is \([0.69, 8.97]\), which is more convergent and provides theoretical support for the cost interval to play a role in actual engineering decision-making and engineering review.

5. Conclusion

Based on the kernel density estimation, this paper constructs a reasonable interval research model for the cost of power grid production technological transformation projects. And taking the disconnector replacement project as an example to calculate the reasonable cost interval, the results show that the cost interval determination model has a high reliability under 85% confidence level. In the process of engineering review, the use of reasonable cost intervals as an index reference can reduce the degree of dependence on expert experience in the work. When the sample data is large enough, this method can calculate more representative cost interval values, which can provide a more reliable analysis basis for investment control of regional power grid technological transformation projects, and has great practical significance.

Acknowledgments

The authors would like to record their thanks to the guidance and help of Professor Liu Jinpeng of North China Electric Power University.

References

[1] Qi, X., Yang, Y.J., Ye, H.N. (2020) Cost deviation analysis of power grid production and technical transformation project. China Power Enterprise Management, (03): 82–83.

[2] Sun, J.B., Wu, X.S., Zhang B.H. (2013) Wind power interval prediction based on nonparametric kernel density estimation. Water Resources and Power, 31(09): 233-235.
[3] Li, Z.W., He, P. (2018) Data-Based Optimal Bandwidth for Kernel Density Estimation of Statistical Samples. Commun Theor Phys, 70(6): 728-734.

[4] Zhao, X.H., Miao, X.J., Zhang Z.G., Zheng H. (2019) Research on Prediction Method of Reasonable Cost Level of Transmission Line Project Based on PCA-LSSVM-KDE. MATHEMATICAL PROBLEMS IN ENGINEERING.