Study on the Method of Increasing Rate of Power Grid Substation Production based on Big Data Statistical Analysis

Yun Huang[1a*], Weijie Zhao[1b], Kai Wang[1c], Shaoyi Wang[1d]

[a*]huyu20221@163.com , [b]173007515@qq.com, [c]250395851@qq.com, [d]981873871@qq.com

[1]Yunnan Power Grid Co.ltd., Kunming 650011, China

Abstract: As an important part of the safe and stable operation of the power grid production emergency repair project, the investment scale level has been continuously improved. However, the current project budget preparation and calculation provisions used in the pre-preparation (settlement) calculation of emergency repair projects cannot fully meet the actual billing and settlement needs of emergency repair projects. Therefore, it is urgent to deepen the adjustment of cost standards related to production emergency repair projects. In this paper on the system analysis of the power grid production current situation of emergency repair project management, through the collection of GX provincial power grid company emergency repair project cost statistics, based on the research object multiple attribute analysis, summarizes the substation maintenance engineering construction process points, and the comprehensive statistics and comparative analysis, provide basic information, data analysis and business practice comparison test reference.

Keywords: Power grid substation production emergency repair project, big data statistics, construction operation costs, rate calculation

1 INTRODUCTION

In recent years, with the rapid development of our country's social economy, the social overall electricity consumption is rising trend year by year, under the electricity demand, power grid construction rapid development, as the intermediate link connecting power generation and electricity, the security and stability of power grid to ensure the normal social power supply, to ensure the social and economic development and people's normal life order is of great significance. But as people demand for electricity rising at the same time, due to natural conditions, external damage, sudden failure cause production equipment, lines and supporting facilities damage is also increasing, driving the production emergency repair project scale gradually increase, power grid enterprise emergency repair project investment gradually increased, bring certain influence on power grid precision input control.

Literature [1] analyzes and collates the cost drivers and historical data of other infrastructure costs of H Power Company in the past three years, and studies the basis and method of standard cost preparation, budget management and standard cost formulation and

EMIS 2023, February 24-26, Chengdu, People's Republic of China

DOI 10.4108/eai.24-2-2023.2330679

implementation of other infrastructure costs of power grid engineering. Literature [2] Taking a power grid infrastructure project in a province as an example, on the basis of meeting the requirements of relevant documents, combined with the settlement data of previous years, the cost calculation standards of feasibility study in the investment stage were discussed. Literature [3] Based on the existing standard cost system of power grid companies, the standard operating cost of overhead is formulated based on the idea of operation, so that enterprises can manage costs with operation centers and corresponding responsibility centers; Secondly, the standard operation cost variance analysis is carried out on the product to find out the root cause of the cost difference and can be implemented in each operation center, that is, the production process. Literature [4] on the basis of analyzing the influencing factors of cost management of overhaul and technological transformation projects of power grid enterprises, the application strategy of the introduction of new quotas for overhaul and technological transformation of power grid enterprises for the cost management of overhaul and technological transformation projects of power grid enterprises is proposed. Literature [5] Based on the three-dimensional structure of Hall, combined with the characteristics and actual situation of power grid engineering cost management, the power grid engineering cost standard system is constructed, and the cost standard system table and system development guarantee applicable to power grid engineering are proposed.

To sum up, on the one hand, relevant scholars are mainly based on the influence of power grid engineering quota on engineering cost standard, on the other hand is more attention to power grid infrastructure engineering cost standard measurement method, and the attention to emergency repair project is relatively weak, therefore, this paper based on big data statistical analysis of power grid substation production emergency repair project construction work rate calculation method research has very good necessity.

2 BASIC PRINCIPLE OF COST REFERENCE STANDARD MEASUREMENT AND ANALYSIS METHOD

In order to determine the base selection of rate calculation scientifically, different correlation coefficient method is introduced. On the one hand, we can determine the correlation difference level of the selected cost base and determine the recommended cost base selection scheme; on the other hand, it is also an important reference and calculation basis for rate calculation.

2.1 The Pearson Correlation Coefficient Method

In statistics, the Pearson's correlation coefficient (Pearson correlation coefficient), also known as the Pearson product Pearson product-moment correlation coefficient (PPMCC or PCCs), is used to measure the correlation between two variables (linear correlation), X and Y. In natural science, the Pearson correlation coefficient is widely used to measure the degree of correlation between two variables, with values ranging between-1 and 1. It evolved from Carl Pearson from a similar but slightly different idea developed by Francis Galton in the 1880s. This correlation coefficient is also called the "Pearson's product-moment correlation coefficient".

The Pearson correlation coefficient between two variables is defined as the quotient of the covariance and standard deviation between two variables:

$$\rho_{XY} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E\left|(X-\mu x)(Y-\mu y)\right|}{\sigma_X \sigma_Y} \tag{1}$$

The above equation defines the overall correlation coefficient, often represented by the Greek lowercase letter ρ as the symbol. Estimating the covariance and standard deviation of the sample yields the Pearson correlation coefficient, which is often represented by the English lowercase letter γ :

$$\gamma = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{2}$$

γ can also be estimated from the mean of the standard fraction of the (X,Y) sample points , giving an expression equivalent to the above equation:

$$\gamma = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma_X}\right)\left(\frac{Y_i - \overline{Y}}{\sigma_Y}\right) \tag{3}$$

where $\left(\frac{X_i - \overline{X}}{\sigma_X}\right)$, $\overline{X}$, and X_i are the standard score, the sample mean, and the sample standard deviation, respectively.

The absolute values of the population and sample Pearson coefficients are less than or equal to 1. If the sample data point falls exactly on the line (if the sample Pearson coefficient is calculated), or if the bivariate distribution is exactly on the line, the correlation coefficient is equal to 1 or -1.

An important mathematical property of the Pearson's correlation coefficient is that the change in the position and scale of the two variables does not cause a change in the coefficient, that is, the invariant of the change (determined by the symbol). That is, if we move to and move Y to, where a, b, c, and d are constants, it does not change the correlation coefficients of both variables (which holds in both the population and the sample Pearson correlation coefficients). A more general linear transformation changes the correlation coefficient:

Since $\mu_x = E(X)$, $\sigma_x^2 = E\left[(X - E(X)^2\right] = E(X)^2 - E^2(X)$, Y are similar, and $E(X - E(X))(Y - E(Y)) = E(XY) - E(X)E(Y)$, the correlation coefficient can also be expressed as:

So the correlation coefficient can also be expressed as follows

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - (E(X))^2}\sqrt{E(Y^2) - (E(Y))^2}} \tag{4}$$

For the sample, the Pearson correlation coefficient:

$$\gamma_{xy}=\frac{\sum x_i y_i-n\overline{x}\overline{y}}{(n-1)s_x s_y}=\frac{n\sum x_i y_i-\sum x_i\sum y_i}{\sqrt{n\sum x_i^2-(\sum x_i)^2}\sqrt{n\sum y_i^2-(\sum y_i)^2}} \tag{5}$$

The above equation gives a simple single-process algorithm for calculating the sample Pearson correlation coefficient, but it relies on the data involved, and sometimes it can be numerically unstable.

2.2 Normal distribution analysis method

Normal distribution (Normal distribution), also known as Gaussian distribution (Gaussian distribution), is a very important probability distribution in mathematics, physics, and engineering, and has a significant influence in many aspects of statistics. In the construction cost analysis work, the rate of a large number of items to be analyzed can be regarded as the sample x. In the case of similar emergency repair engineering conditions, x should meet the normal distribution, its μ is the average rate under the normal level, and μ±σ is the reasonable rate fluctuation range. Rates greater than μ + σ or less than μ - σ are unreasonable intervals.μ±σ/ 2 is the rate fluctuation interval with large probability.

Normal distribution is the theoretical basis for many statistical methods. Various statistical methods, such as testing, ANOVA, correlation and regression analysis, all require the analyzed indicators to follow a normal distribution. Although many statistical methods do not require the analysis index to obey a normal distribution, the corresponding statistics are approximately normal distributed in large samples, so these statistical inference methods are also theoretically based on a normal distribution. Normal distribution can be used to estimate the frequency distribution. As a variable that follows a normal distribution knows, it can estimate the frequency proportion in any value range as long as it knows its mean and standard deviation according to the formula. Can also make reference value range, normal distribution method is applicable to obey normal (or approximate normal) distribution index and can obey normal distribution index, through the positive distribution analysis, can study whether the sample data meet positive distribution, and set the cost standard comparison, can test the rationality of the standard setting.

3 EMPIRICAL ANALYSIS

3.1 Sample data

This paper collects the statistics of production emergency repair projects in 14 provincial power supply companies from 2019 to June 2021. The number and cost of emergency repair projects in G cities are shown in Table 1 below.

Table 1: Statistical table of the annual fee quantity and expenses of emergency repair projects.

order number	The bureau	In 2019,		In 2020,		In 2021,		amount to	
		Quantity (individual)	amount of money (Wan Yuan)	Quantity (individual)	amount of money (Wan Yuan)	Quantity (individual)	amount of money (Wan Yuan)	Quantity (individual)	amount of money (Wan Yuan)
1	A	/	/	15	101.572	342	863.098	357	964.67
2	B	179	483.22	1459	1185.85	317	273.91	1955	1942.98
3	C	35	235.74	602	1545.04	147	866.18	784	2646.96
4	D	56	85.88	1074	3199.75	375	1042.52	1505	4328.15
6	E	2	17.67	10	74.2	19	104.8	31	196.67
7	F	93	221.21	266	604.99	63	240.19	403	1066.39
8	G	6	56.48	8	59.57	1	12.9	15	128.95
9	H	82	98.82	253	222.95	46	337.13	371	658.9
10	I	131	218.97	26	86.602	7	65.428	164	371
11	J	1	16.2	1168	1505.48	5	783.26	1174	2304.94
12	K	28	13.96	25	60.757	15	222.953	68	297.67
13	L	13	104.56	15	112.73	51	204.99	79	422.28
14	M	13	33.5	131	209.48	33	212.56	177	455.54
15	N	9	51.2	45	295.47	12	29.71	66	376.38
16	O	/	/	1	16.5	8	6.07	9	22.57
subtotal		648	1637.41	5098	9280.94	1441	5265.67	7185	16184.03
proportion		9.02%	10.12%	70.95%	57.35%	20.06%	32.54%	100%	100%

3.2 Corresponding ence of sequence of main cost of sample engineering

Through the multi-series comparison of the main costs of the sample project, the scatter plot of the comparison between the increased cost of construction operations in the emergency repair project and the six cost series of direct engineering cost, labor cost, material cost, mechanical cost, labor cost + mechanical cost is shown in Figure 1 below.

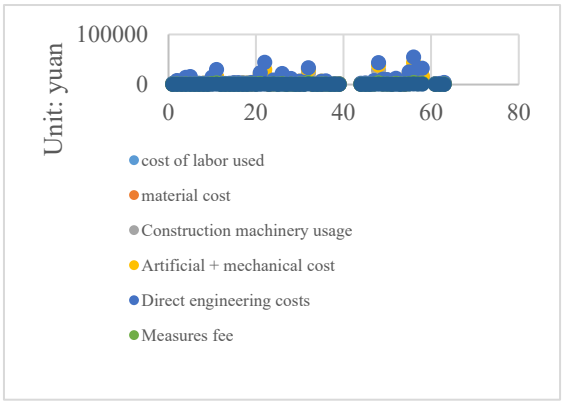

Figure 1: Corresponding sequence of the main cost of substation sample engineering.

3.3 Cost sequence correlation analysis

Based on the series correspondence shown by the scatter plot, in order to further quantify the correlation between the series of "increased costs for construction work" and the two cost series of "labor costs" and "labor + mechanical costs", this section combines the correlation coefficients of Pearson, Kendall and Spearman for econometric analysis. Through SPSS software, the Spearman correlation coefficient of the five cost series of labor costs, material costs, construction machinery usage costs, labor plus machinery costs and direct engineering costs in five types of emergency repair projects and the series of construction work increase costs is output, and the comparison chart shown in Figure 2 below is obtained.

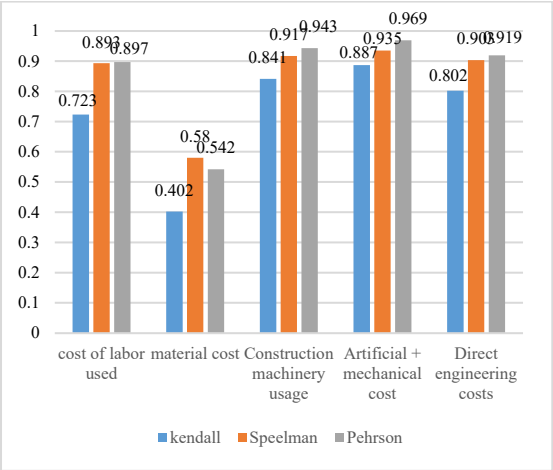

Figure 2: Comparison of the correlation coefficient of substation sample engineering cost sequence.

3.4 Measurement and analysis of fee withdrawal rate

Based on the sample statistics, the rate distribution of the substation emergency repair project is calculated from the "labor cost + machinery cost" as the base, and the rate and frequency distribution of the construction operation increase cost in the emergency repair project cost is finally obtained, as shown as follows.

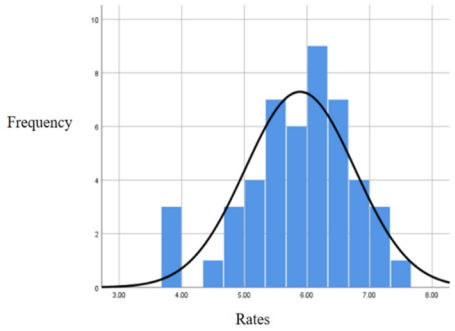

Figure 3: Histogram of additional fee rate distribution for construction operation of substation maintenance works.

From Figure 3, for substation maintenance works, the rate distribution is 5.33-6.67% based on the cost distribution chart. According to S PSS automatic statistical analysis, the distribution μ value is 5.91%, so the cost increase rate for substation maintenance works is 5.91%.

4 CONCLUSION

This paper is a basic research work based on the characteristics of power grid production emergency repair project and solving the actual problems of production emergency repair project related expenses. Based on the pre-calculation (settlement) of emergency repair projects of GX Provincial power grid Company in recent years and the statistical analysis of relevant data, In order to realize the scientific and reasonable determination of the measure cost of the emergency repair project as the goal, In-depth analysis of the cost composition and cost level of emergency repair projects of power grid projects, According to the basic theory and method of cost calculation, Introducing some guiding data processing technology and mathematical statistical method model, Combined with the characteristics of emergency repair projects of power grid engineering, Reasonable construction of power grid engineering emergency repair project measures cost calculation thinking framework, The combination of qualitative and quantitative, induction and deduction, theoretical analysis and practical guidance and application are used to summarize the typical construction technology, Research object attribute category combing, cost level standard design, Using the method of combining demand orientation and analogy analysis to determine the mode of measure items, Put forward the effective research ideas and technical methods to support it, Improve the scientific nature and applicability of the research results.

REFERENCES

[1] Chen Si. Research on other Cost Control of power grid infrastructure based on standard Cost [D]. Changsha University of Science and Technology, 2014.

[2] Fang tianrui. Discussion on the cost calculation standard of feasibility study of power grid infrastructure project [J]. Engineering and Construction, 2019,33 (01): 150-151.

[3] Lucy Wang. Research on Standard Operation Cost Management of N Power Supply Company [D]. Xi'an Petroleum University, 2020.DOI:10.27400/d.cnki.gxasc.2020.000492.

[4] Ma Ruchao. The introduction of the new quota of maintenance and technical reform on the cost management of power grid enterprises [J]. Modern Economic Information, 2010 (24): 34-35.

[5] Wang Jing, Xia Gorgeous, Ding Weiwei. Research on the Framework Construction of Power grid Engineering cost standard System [J]. Zhejiang Electric Power Company,2016,35(11):54-59.DOI:10.19585/j.zjdl.2016.11.014.