

Error Analysis of the Cable Force on Skew Normal Distribution

SHEN Lin^a, LIU Lai-jun^b, WU Fang-wen^c, LI Yu^d

Key Laboratory for Bridge and Tunnel of Shaanxi Province, Chang'an University, Xi'an 710064, Shaanxi, China;

^ashenlin1618@163.com, ^bliulj@chd.edu.cn, ^c527723433@qq.com, ^dshennergou123@163.com

Keywords: bridge engineering; cable-stayed bridge; cable force; skewness coefficient

Abstract. In order to investigate the the mathematical model considering the skewness-parameter, which covers the shortage of long-span bridge in sensitive parameter analyzing. Random variable ξ along with its statistic probability model is taken. The application of influence coefficient helps to build up a determining index that abnormal values could affect on the precision. A mathematical model of skew normal distribution and cable tension, bases on skewness coefficient and variation coefficient, is set up to analyze the state of cable tension counting on engineering. Results show that this method is reliable. It not only provides a relatively high recognition precision, but also shows advantages on forecast precision comparing with the engineering practice which is suitable to engineering practice.

1 Introduction

The complexity for the construction control of long-span bridge is that it's difficult to predictive control the output value of input information. The processing and analysis of data is often based on the assumed or some approximation of the premise, and the error which was caused by the assumption has become necessary to be solved. Foreign scholars has made some achievements in related fields, Hosking and Wallis [1] presented normal deflection coefficient with generalized, Klima M S et al [2] proposed to describe the deviation by two auxiliary index curves. The research in the field of domestic have few research in bridge. Therefore, this paper carried out a study on the statistical probability model of cable force, which is suitable to engineering practice.

2 Theory analysis of disturbance on random error

The random error is made of numerous and tiny change factors, which is transitional from relative intact (RI) and fully adjusted (FA) states. It caused running state point departure with the design state point. Parameters of stochastic error include: 1) correlation; 2) normality; 3) heteroscedasticity. The linear regression model is

$$x^t = x^m + \varepsilon \quad (1)$$

In which x^t is theoretical value, and x^m denotes the observed value, ε is error term. At the assumption of ε is probability density function probability, then the probability of error terms is

$$p(x) = \int \cdots \int f(x) dx_1 \cdots dx_n, n=1,2,\cdots,k \quad (2)$$

Using the method of maximum likelihood, then

$$L(\hat{\theta}) = \max L(\theta) = p(X_1 = x_1, X_2 = x_2 \cdots X_n = x_n) = \prod_{n=1}^N \int f(x_n; \theta) dx_n \quad (3)$$

In which, θ is regards as an unknown parameter. As is known that the error distribution is the real embodiment between theoretical model and the project in appropriate structural topology, geometry and material characteristics. The curve of error value may bended down in a large numerical region, or bend up by the end of small value, both of them possible exist at the same time. And it means that it is not entirely accurate to describe the distribution curve by normal distribution, which can possible describe part of the area. In order to solve this problem, it is necessary to achieve a more accurate description of the asymmetric nature of the distribution curve by introducing the concept of coefficient of variation,

2.1 effect of the outliers for the accuracy on statistical value

In all of the characteristic that could reflect population distribution, actual residual is regard as one that could descript the variation of the observed values, which including the degree of migration and bias parameters extent and variation extent

1) Skewness coefficient. Skewness coefficient is a measure of migration.

$$\partial = \frac{m_3}{\sigma_3} = \frac{\sum (x - \bar{x})^3 f}{\left(\frac{\sum (x - \bar{x})^2 f}{\sum f} \right)^{3/2}}$$

When the skewness coefficient $\partial > 0$; means the distribution is right skewed, and shows left skewed when $\partial < 0$. However, the value of outlier $G(x_i)$ will plays a stimulating role on m_3 . The greater the departure degree is, the more influence m_3 is. The reliability of skewness coefficient will reduce without considering the outlier, or just remove it. Therefore, it is inaccurate with the use of the assumption of normal.

2) Coefficient of kurtosis. Coefficient of kurtosis is to describe the distribution of steepness, which is a characterization that shows the probability of density distribution curve at the peak level of the average in the number of features.

$$K = \frac{\sum_{i=1}^k (x_i - \bar{x})^4 f_i}{(n-1) s^4}$$

Coefficient of kurtosis was K , x_i is patient samples, \bar{x} is the average of patient samples, and s is the sample standard deviation. Kurtosis of the normal distribution is three. When $k < 3$ means the kurtosis of distribution is insufficient, density curve performance chunky and flatter than normal; $k > 3$ means density curve has excess kurtosis and the performance is sharp peak.

As the feature point in analysis the skewness and kurtosis, outlier is not only the substandard point, but also the catastrophe point of the trend, which will strongly influence the distribution of data in the form hypothesis testing, and even the choosing of the null hypothesis. To determine the affection of outlier on standard deviation, then decision on basis of removing observed values. The ratio of generalized variance on observation generalized and all the data, namely the impact factor^[1]

$$cov\ ratio(x_i) = \frac{MSE_i^{m+1}}{MSE^{m+1}} \left(\frac{1}{1 - h_i} \right) \quad (4)$$

Take the maximum value as the value to judgment

$$\sigma' = cov\ ratio(x_i)_{max} = \left| \frac{MSE_i^{m+1}}{MSE^{m+1}} \left(\frac{1}{1 - h_i} \right) \right| \quad (5)$$

in which, MSE_i^{m+1} means square error after deleting the data x_i , h_i is expressed as a diagonal elements of hat matrix. When $cov\ ratio(x_i)$ is more than 1, it improves the accuracy of parameter estimates; and the accuracy of the observations reduce when $cov\ ratio(x_i)$ is less than 1, which will be the evidence to cancel the sample or not.

2.2 mathematical model Based on the multi-parameter skew normal distribution

1) normally distributed variables is equivalent in probability distribution function of non-normal distribution mathematical model treatment of normality skew normal distribution, according to the normal condition, set normal, non-normal distribution.

① normally distributed variables is equivalent in probability distribution function of non-normal distribution (first part), namely $F(x) = G(x)$; ;

② normally distributed variables is equivalent in probability density function of non-normal distribution, namely $f(x_i) = g(x_i)$. When the coefficient of skewness is

When $\partial > 0$,

$$G(x) = F(x) = \int_{-\infty}^{x_i} f(x)dx = \int_{-\infty}^{x_i} g(x)dx = \Phi\left(\frac{x_i - \mu'}{\sigma'}\right) \quad (6)$$

$$\mu' = x_i - \Phi^{-1}(G(x))\sigma' \quad (7)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (8)$$

if $Y = a + bX$, then

$$f_Y(y) = \frac{1}{b} f_x\left(\frac{y-a}{b}\right) \quad (9)$$

$$f_Y(y) = \sigma f_x(\sigma y + \mu)$$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma'} e^{-\frac{(x-\mu')^2}{2\sigma'^2}} = f\left(\frac{x-\mu'}{\sigma'}\right) \frac{1}{\sigma'} \quad (10)$$

$$\sigma' = \frac{f[\Phi^{-1}(G(x))]}{f(x)} \quad (11)$$

if $\partial < 0$, then

$$G(x) = F(x) = \int_{x_i}^{+\infty} f(x)dx = \int_{-\infty}^{x_i} g(x)dx = 1 - \Phi\left(\frac{x_i - \mu'}{\sigma'}\right) \quad (12)$$

3 Generalized normal distribution

According to Gram-Charlie expansions, it represent generalized normal distribution, make up pf normal distribution and polynomial, which including skewness coefficient and coefficient of kurtosis. [2], the distribution function is

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} dx + \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \left[\frac{\partial}{6} (u^3 - 3u) + \frac{K}{24} (u^4 - 6u^2 + 3) + \frac{1}{120} \frac{\beta^5}{\sigma^5} (u^5 - 10u^3 + 15u) \right] + \quad (13)$$

It took up to first two terms only, so that the series formula became

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} dx + \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \left[\frac{\partial}{6} (u^3 - 3u) + \frac{K}{24} (u^4 - 6u^2 + 3) \right] \quad (14)$$

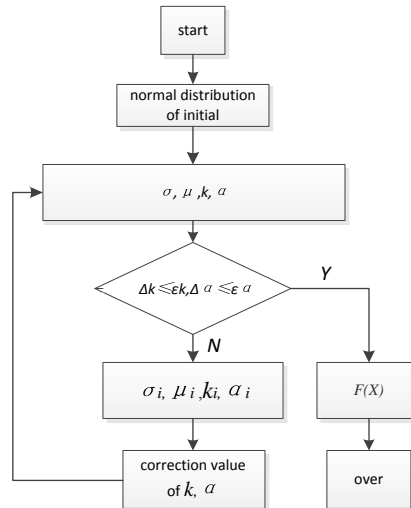


Fig. 1 process of mathematical model

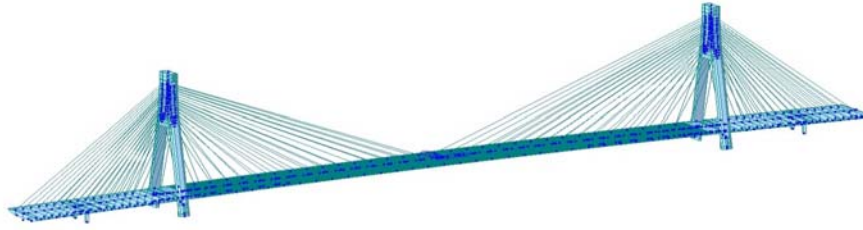


Fig. 2 Finite element model of Cable-stayed bridge

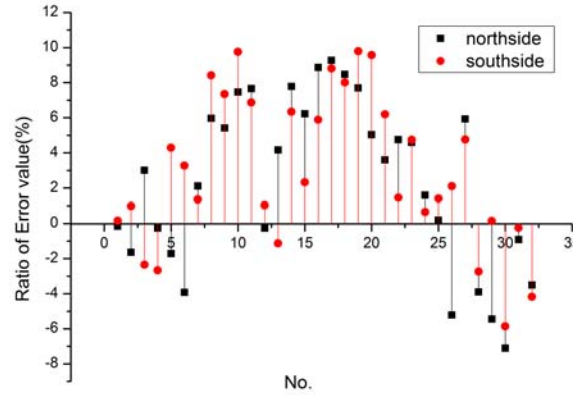


Fig. 3 Error distribution of cable force

The results in Figure 3 shows that the cable tension had error between 5%, and 8% accounted for 24% in total, 17% in total had the error between 8% to 10%, and the maximum is 9.87%. But the results showed that error asymmetric capital and no constant changes. Among them, the error of S1 ~ S5 is large and M6 had the maximum of upstream, S3 in downstream had the peak error level of cable tension, S13 is minimized. The distribution curve of cable force showing the error skewed is normal distribution.

3.1 Affect the accuracy of determining the value

In order to improve the accuracy of the system control large point observations were took and calculate the impact coefficient in formulae 4.

Tab 1 Error distribution of cable force

No.	$MSE_{(i)}$	MSE	h_i	$covration(x_i)$
B4	29946.63	30890.35	0.048	1.018
Z6	30704.13	30890.35	0.022	1.016
Z1'	30120.18	30890.35	0.009	0.984
B3	32579.52	33262.78	0.037	1.017
B4'	32442.82	33262.78	0.042	1.018
Z7	32877.97	33262.78	0.028	1.017
B1'	33094.74	33262.78	0.021	1.016

The datas above showed that when the impact factor value is more than 1, the sample points are indeed influential point of the distribution, and was thought to improve the accuracy of parameter estimates, the value should be retained. In the samples, the coefficient of MI' is less than 1, the square error of ,excluding the observed value, an increase in the mean will improve the accuracy of parameter estimation worth.

3.2 method of nonlinear fitting

Through the proposed generalized normal distribution based on correction model g normal after the generalized normal distribution, and the fitting equation of cable tension is shown formulae 5.

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-5.759)^2}{142.2}} dx + \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-5.759)^2}{142.2}} \left\{ \frac{-0.014}{6} \left[\frac{(x-5.759)^6}{8.43^3} - 3 \frac{(x-5.759)^2}{8.43} \right] + \frac{1.77}{24} \left[\frac{(x-5.759)^8}{8.43^4} - 6 \frac{(x-5.759)^4}{8.43^2} + 3 \right] \right\} \quad (5)$$

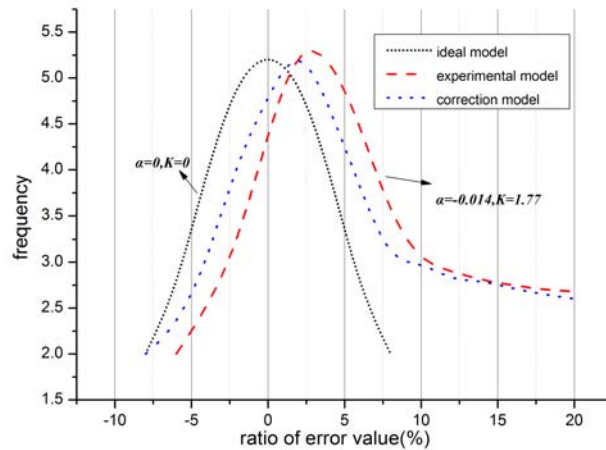


Fig. 3 distribution curve shape considering $\hat{\alpha}$ and K

- (1) Skew normal distribution as post correction skewness coefficient $\hat{\alpha} = -0.014$, $k = 1.77$, and the measured error is coinciding a little with actually measured value .
- (2) the skewness coefficient of cable force error is negative, model curve skewed present right deviation, the density probability curve in the low end of the upturned.
- (3) regression accuracy of skew distribution was higher than normal error characteristic equation.

4 Summary

- 1) In the evaluation of impacting on outliers, impact factor value can better fulfill its influence coefficient of the distribution which helps to improve the accuracy of the relative error value
- 2) The result is more practical by introducing skew normal distribution coefficient of skewness and kurtosis coefficients
- 3) Normal distribution skewed probability iterative mathematical model is contented the engineering requirements by reusing iterative information times, which is higher than the normal distribution with zero mean and the assumption of random error.
- 4) The method that mention in this paper could applies not only in the construction of Cable control and optimization, but also the state of cable tension.

If you follow the “checklist” your paper will conform to the requirements of the publisher and facilitate a problem-free publication process.

Acknowledgements

This work was financially supported by the National Natural Science Foundation (51408040, 51408042), Program of Yunnan Provincial Transport Department of Science and Technology (2013(A)022).

References

- [1] Hosking J R.M, Waills J Regional frequency analysis: an approach based on L-moments[M].New Yourk:Cambridge Universitu Press,1997.
- [2] Klima M S,Luckie P T.Evaluating coal separation processes[J].Coal Preparation,1988,6:91-101
- [3] WANG GUO-hong, SHENG Hong-yu. Finite difference method to solve transverse vibration of the suspender and the application in cable tension testing.[J].Journal of Hefei university of technology, 2013, 36 (8): 978-982

- [4] DONG Lian-ping, FAN Min-qiang, YANG Hong-li,et. Mathematical model for partition curve based on generalized normal skew distribution [J]. Journal of china coal society, 2013, 38 (10): 1856-1861.
- [5] YAN Pei-pei a, ZHANG Xing-fang a, FAN Min-qiang. Particle Size Distribution Functions Containing a Skew Coefficient [J]. China powder science and technology. 2013, 19 (3): 58-61
- [6] ZANG Zhi-jun, LING Neng-xiang. Strong convergence of parametric estimation in semi-functional partial linear regression model with AR(1) error [J].Journal of Hefei university of technology, 2010, 33 (4): 616-620
- [7]Jong-Tae Kim, and Norris Stubbs , Model-uncertainty impact and damage detection accuracy in plate girder, Journal of Structural Engineering,ACSE,1995,121(10),1409~1417
- [8]D.Janjic,M.Pircher,etc.Optimization of Cable Tensioning in Cable Stayed Bridge[J],Bridge of Engineerings,2003:131-137
- [9]Fleishman A.I.A method for simulatane non-normal distributions [J] .Psuchometrika, 1978, 43(4):521-532