Rock and Soil Mechanics

Volume 41 | Issue 4 Article 4

9-27-2020

Methods of determining the minimum number of samples for statistical measurement of rock joint roughness

Qin-kuan HOU

Rui YONG
yongrui_usx@hotmail.com

Shi-gui DU

Min-na XU

See next page for additional authors

Custom Citation

HOU Qin-kuan, YONG Rui, DU Shi-gui, XU Min-na, CAO Ze-min. Methods of determining the minimum number of samples for statistical measurement of rock joint roughness[J]. Rock and Soil Mechanics, 2020, 41(4): 1259-1269.

Rock and Soil Mechanics 2020 41(4): 1259−1269
https: //doi.org/10.16285/j.rsm.2019.5788
ISSN 1000−7598
rocksoilmech.researchcommons.org/journal

Methods of determining the minimum number of samples for statistical measurement of rock joint roughness

HOU Qin-kuan, YONG Rui, DU Shi-gui, XU Min-na, CAO Ze-min
Shaoxing University, Shaoxing, Zhejiang 312000, China

Abstract: The rock joint roughness has many characteristics like heterogeneity, anisotropy, nonuniformity and scale effect. In engineering practice, different statistical methods are utilized for analyzing the rock joint roughness due to its uncertainty. However, previous studies often neglected the impact of insufficient samples on statistical results. To solve the problem that reasonable number of samples cannot be determined during the statistical measurement of joint roughness, the methods based on the coefficient of class ratio analysis and the simple random sampling principle are proposed for determining the minimum number of samples (MNS), respectively. In the case study, the MNS of statistical measurements is determined based on the proposed methods. The results of rock joint samples are compared and analyzed with different sample sizes. The results indicate that the coefficient of variation (COV) value of the small samples is significantly larger than that of large ones, and the COV value decreases with increasing size of samples. The COV values of the joint samples with the sizes of 10−50 cm basically are in a range of 0.31−0.47, and the values for those of 60−100 cm samples are between 0.21−0.31. The relationship between MNS and sample size basically satisfies the power function relationship, and the MNS decreases with the sample size. The MNS determined by the former method with an allowable error of ±2% is consistent with that calculated by the latter with a maximum allowable error of 10% and a confidence level of 95%. The similarity of the results based on these two methods is greater than 0.997. This study can provide basis for quantitatively obtaining the MNS in rock joint roughness statistical measurement, and can ensure the accuracy of JRC statistical analysis. It is of great significance to accurately obtain mechanical parameters of rock joints in rock mass stability evaluation.
Keywords: joint roughness coefficient (JRC); class ratio analysis; simple random sampling principle; minimum number of samples (MNS)

1 Introduction

Rock joint is one of the main factors controlling the stability of rock engineering such as the slope, underground chamber, etc.[1−3]The surface morphology of the joint is composed of three elements: macro-geometrical contour, surface undulation morphology and micro-roughness. Among them, the surface undulation morphology constitutes the peak-valley undulation of the common scale rock joint, which is the main factor affecting the joint roughness[4]. In 1973, Barton[5] firstly proposed joint roughness coefficient (JRC), which has been used to quantitatively describe the roughness of the rock joint and to construct a model for estimating the shear strength of the joint. It is now widely used in rock engineering practice. Du et al. [6−9] firstly proposed that the JRC has characteristics of heterogeneity, anisotropy, non-uniformity and size effect. Specifically, the heterogeneity refers to the difference in surface morphology of rock joint along the same measuring direction due to different compositions of rocks on the same rock joint; the anisotropy refers to the difference in surface morphology along different measuring directions under the same sampling length of the same joint developed in the same wall rock; the nonuniformity refers to that the same rock joint with the same rock composition, the surface morphology of different parts is different; and the size effect refers to the phenomenon that the mechanical parameters of rock mass decrease with increasing sample size. These characteristics will lead to the uncertainty of JRC acquisition. Field engineering surveys also found that sampling deviation and geometric measurement accuracy will cause changes in the mechanical properties of the rock joint due to the difference of technicians' experience and cognitive level. Among them, the complex characteristics of the rock joint are the major cause of the uncertainty of the mechanical properties, and the sampling deviation and sampling size are the direct causes of this result. Due to the complexity, ambiguity and uncertainty of the mechanical properties of rock mass joints and the limitations of JRC measurement methods, it is difficult to obtain accurate rock joint mechanical parameters.

Received: 30 April 2019 Revised: 17 July 2019
This work was supported by the National Natural Science Foundation of China (41502300, 41427802) and the Natural Science Foundation of Zhejiang Province (LQ16D020001).
First author; HOU Qin-Kuan, male, born in 1995, Master degree candidate, research interests include the mechanical properties of rock mass joints and slope engineering stability evaluation. E-mail: 771575689@qq.com
Corresponding author: YONG Rui, male, born in 1987, Post-doctor, associate Professor, research interests include the mechanical properties of rock mass joints and slope engineering stability evaluation.
E-mail: yongrui_usx@hotmail.com

In view of the above difficulties, statistical methods are particularly useful for analyzing the JRC characteristics. Du [10] calculated the JRC value of 12 contour curves with a length of 10 cm, and converted the overall JRC value of the rock joint by using the JRC size effect fractal model, which provides an effective means for quantitative statistical analysis of the JRC size effect. Ji [11] carried out a quantitative determination of the JRC of hard joints, which provided sufficient conditions for rapid prediction of joint parameters by the Barton's theoretical formula. Wu et al.[12] used random sampling with truncated normal distribution to obtain rough joints with a certain trace length. The JRC parameters were counted, and the regression relationship between each statistical parameter and JRC was obtained. In order to overcome the limitations of existing JRC fractal evaluation methods, Ma et al. [13] proposed an optimal sampling interval based on the fractal dimensions of the contour curve of rock joint under different sampling intervals. Yong et al. [14-15] proposed a method for statistically determining the maximum sampling interval based on Fourier series and a sampling method for determining series scale JRC. The above-mentioned studies all analyzed the JRC through statistical methods, but none of them proposed methods for determining the number of statistical measurement samples, or considered the relationship between sample size and sample number. Reasonable sampling size and the minimum number of statistical measurement samples play a key role in accurately obtaining the JRC value. For the study of the size effect of the joint roughness, determining a reasonable minimum number of samples (MNS) under different sample sizes is also an urgent problem to be solved.

The sample size is directly related to the JRC estimation accuracy[16], and there is always a balance between them. Effectively reducing the statistical uncertainty is the key to improving sampling accuracy. In recent years, many researchers have investigated the possible effect of the statistical measurement samples. Among them, Gao [17]proposed a method for determining the parameters of rock and soil, and concluded that at 95% confidence level, the relative error of the sample mean can be less than 3% when the number of samples is greater than or equal to 6, which may meet the requirement for engineering practice. Dong et al. [18] proposed a reasonable number of sub-samples for the actual variability coefficients of different physical indexes in different soil layers using *t* distribution theory. Li[19] claimed that 6 samples is only applicable to the specifically recommended formula, and cannot be used as the generic standard for controlling the sample size. Zhao et al.[20] discussed the reasonable number of samples of rock and soil. The above-mentioned methods can obtain a certain number of sample statistical values, but they cannot accurately calculate the relationship between the minimum number of samples and the allowable error, nor can they put forward a quantitative method to determine the reasonable number of samples for statistical measurement of joint roughness, which is not universal.

This paper proposes methods of determining MNS based on the analysis of the coefficient of class ratio and the principle of simple random sampling. Through the statistical analysis of JRC, the variation rule and the allowable error range of MNS for statistical measurement under series sizes are obtained. This method has some improvements over the traditional method only specifying the sample number based on empirical standards, and can solve the problem that MNS cannot be determined with samples collected in the past. In this regard, the MNS obtained by the proposed statistical method for JRC measurement will be more reasonable, which is of great significance for the accurate extraction of rock joint mechanical parameters in related projects.

2 Method of determining MNS for statistical measurement of JRC

2.1 Method for determining MNS based on class ratio analysis of the coefficient of variation

In rock mass engineering, the analysis and value of JRC is the basis for the study of rock mass mechanical parameters. Coefficient of variation (COV) is a dimensionless parameter introduced to evaluate the variational characteristics of geotechnical parameters. *Code for investigation of geotechnical engineering*（GB 50021－94）[21] gives the evaluation criteria for the variability of this parameter, in Table 1. The dispersion of the data can also be expressed by the standard deviation, but because this parameter is a dimensional index, it is not suitable for the comparative study of discrete analysis due to the difference in the mean value of the parameters. In engineering practice, the dimensionless COV is usually used for the evaluation of geotechnical parameters.

The class ratio analysis is usually used to characterize the smoothness of the data series, and to improve the approximation accuracy of the exponential model in the grey theory. Grey theory plays an important role in determining the accuracy of prediction models [22–23]. In this paper, the stability value of COV is determined by analyzing the class ratio relationship of the COV, so as to quantitatively obtain the MNS of the statistical measurement of the joint roughness under the series size. When the COV's of the adjacent samples are same, the class ratio is equal to unity; when the difference between the two samples is large, the class ratio will be far from unity; when the COV corresponding to the number of different samples tends to be stable, it means that the number of samples has been sufficient to meet the requirements of calculation accuracy. Statistically, the error range represents the random sampling error in the survey results. The probability of the result falling into the error range is usually 98%. Therefore, the allowable error range $\varepsilon = \pm 2\%$ used in this paper is sufficient for determining the reasonable MNS. All subsequent data points are strictly controlled within ±2%, that is, when the class ratio coefficient is controlled between 0.98 and 1.02, it can be considered that the number of samples is sufficient and meets the requirements for calculation accuracy, and the number of samples can be determined as the MNS of the rock joint roughness statistical measurement.

Table 1 Classification of parameter variability

COV	COV<0.1	0.1≤COV<0.2	0.2≤COV<0.3	0.3≤COV<0.4	COV≥0.4
Variation	Very low	Low	Moderate	High	Very high

This data processing method is mainly used to quantitatively determine the MNS for statistical measurement of rock joint roughness. A method for determining the MNS based on grade ratio analysis of the COV values is proposed. The main steps are as follows:

(1) Select proper rock joints, and use a contour curve meter to extract *k* contour curves of length *l* on the surface of the joint, where the value of *k* should be sufficiently large.

(2) Calculate JRC_i of the rock joint corresponding to the *i*-th measurement segment, where $i \leqslant k$.

(3) *k* samples are grouped, the first group is *s* samples, the second group is *s*+*c*, the third group is *s*+2*c*,……, where *s* is any number of samples, and *c* is any number of samples added by each group, a total of *m* groups (*m* is a natural number), calculate and list the corresponding average μ_i and standard deviation σ_i of the rock joint JRC of each group, where

$$\mu_i = \frac{1}{k}\sum_{i=1}^{k}\mathrm{JRC}_i \tag{1}$$

$$\sigma_i = \sqrt{\frac{1}{k-1}\sum_{i=1}^{k}(\mathrm{JRC}_i - \mu)^2} \tag{2}$$

(4) Calculate and record the coefficient of variation CV_i of each group using the obtained average and standard deviation, where:

$$\mathrm{CV}_i = \frac{\sigma_i}{\mu_i} \tag{3}$$

(5) Obtain the class ratio coefficient y_i of each group by the following formula, and record the results in order to obtain the gradation sequence.

$$y_i = \frac{\mathrm{CV}_i}{\mathrm{CV}_{i+1}} \tag{4}$$

(6) Draw a scatterplot with the calculated sequence of class ratios. When all points after a certain point fall within the interval $1+\varepsilon(\varepsilon = \pm 2\%)$, it means that the number of samples represented by this point is the MNS required when the sample length is *l*. Note that the MNS determined based on the analysis of the COV of the series size is n_c.

2.2 Method for determining MNS based on simple random sampling principle

Sampling survey is based on scientific probability theory. It can calculate statistical errors of statistical data and control the reliability of statistical data. It has many advantages such as saving survey cost and timeliness[24]. Simple random sampling, also known as pure random sampling, is the most basic and simple sampling organization form in mathematical statistics. During the statistical measurement of the joint roughness in the project, the sampling samples at different measurement positions are different. The process of each sampling is essentially the same as the principle of simple random sampling. The number of random sampling is the number of samples of the statistical measurement of the roughness of the joint.

The determination of sample size in sampling survey is very important, and it is directly related to the survey cost and accuracy. The survey accuracy requirements for the parameters θ to be estimated are generally expressed in absolute error *d* or relative error *r*, that is, under the confidence of $1-\alpha$, ensure that the difference between the estimated quantity $\hat{\theta}$ and the parameter θ is within the error limit:

$$P\left(\left|\hat{\theta}-\theta\right| \leqslant d\right) = 1-\alpha \tag{5}$$

or

$$P\left(\frac{\left|\hat{\theta}-\theta\right|}{\theta} \leqslant r\right) = 1-\alpha \tag{6}$$

When the sample size n_s is large (e.g., >30), according to the central limit theorem, $\hat{\theta}$ approximately follows a normal distribution:

$$P\left(\frac{\left|\hat{\theta}-\theta\right|}{\sqrt{V\left(\hat{\theta}\right)}} \leqslant t\right) = 1-\alpha \tag{7}$$

where $V(\hat{\theta})$ is the sampling variance; *t* is the two-sided α quantile of the standard normal distribution.

In the mean value estimation, the parameter θ to be estimated is the population mean value $\overline{Y}$, and the estimated amount is the sample mean value $\overline{y}$, N is the total sample amount. Through $d = t\sqrt{V\left(\overline{y}\right)}$ and $V\left(\overline{y}\right) = \left(1-\frac{n_s}{N}\right)\frac{S^2}{n_s}$, it can be solved

$$n_s = \frac{Nt^2S^2}{Nd^2+t^2S^2} = \frac{\dfrac{t^2S^2}{d^2}}{1+\dfrac{t^2S^2}{Nd^2}} = \frac{Nt^2S^2r^2\overline{Y}^2}{Nr^4\overline{Y}^4+t^2S^2r^2\overline{Y}^2} \tag{8}$$

where n_s is the MNS required; S is the standard deviation of the overall JRC sample; $\overline{Y}$ is the average of the overall JRC samples; and r is the statistically acceptable relative percentage deviation.

3 Case studies

3.1 Rock joint sample selection

The calcareous slate rock joint is selected from Changshan county, Zhejiang province, as shown in Fig.1. The original rock joint is a neutral tuff without recrystallization, which can be peeled into thin slices along the direction of the slab for a good rock joint. The selected joint is hard and intact, and the joint surface is smooth to rough, which fully meets the test requirements. Table 2 shows the main physical and mechanical indexes of the calcareous slate[25].

Table 2 Physical and mechanical indices of calcareous slate[25]

Lithology	Density /(g • cm^{-3})	Compressive strength /MPa	Elastic Modulus /GPa	Softening coefficient
Calcareous slate	2.68	78.6	43	0.92

Fig.1 Natural rock structural surface

3.2 Extraction of geometric information of the joint

Considering the exposed condition of the natural rock joint and the difficulty of measuring the field roughness, different scholars have adopted different data collection methods, which can be divided into contact and non-contact types. The former mainly includes single-row needle-shaped profile ruler[26], single-needle automatic profiler[27], simple longitudinal profiler[28], etc.; the latter mainly includes photogrammetry[29], three-dimensional laser scanning method, etc.[30–31]. In this paper, the three-dimensional laser scanning method is used to obtain the surface data of the joint. This method has the advantages of high precision, fast speed, and significant time and cost savings.

Metra SCAN 750|Elite handheld 3D laser scanning system from Canadian Creaform company was used in this study. By recording a large number of dense point cloud 3D coordinates, reflectivity, texture and other information on the surface of the measured object, the system can quickly create a 3D model of the measured target. The maximum scanning accuracy of the instrument is 0.003 cm. It has the unique advantages of high efficiency, high precision, and non-stratification of point clouds. Figure 2 shows the three-dimensional laser scanning of the joint. The scanned point cloud data is combined in Geomagic software to generate the corresponding rock joint spatial model, as shown in Fig.3. Excise the unnecessary data, intercept the 100 cm×100 cm joint, and uniformly intercept 2 001 joint contour curves along the y-axis direction with an accuracy of 0.05 cm[32], and convert it into 2 001 txt text data. It is assumed that the JRC obtained under this accuracy is the true.

Fig.2 3D laser scanning of structural surface

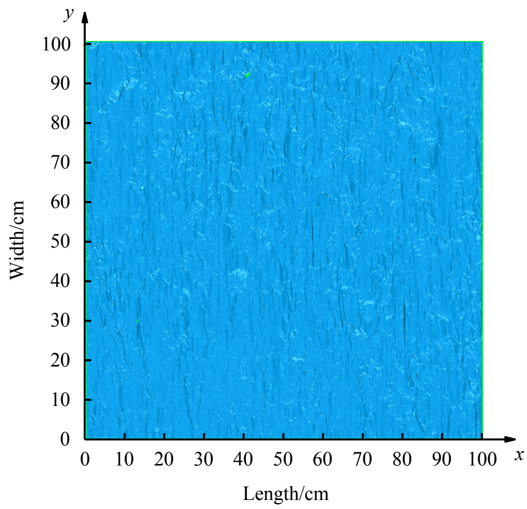


Fig.3 Spatial modelling of structural surface

3.3 Calculation of JRC

Convert the obtained 2 001 txt text data to the corresponding Matlab data and load it into the Matlab calculation program. The calculation program can automatically intercept contour curves of 10, 20, 30, 40, 50, 60, 70, 80 cm and 90 cm length from contour curves of 100 cm length, and automatically read the coordinate data of each contour curve. The uniform segmentation method is used to carry out series size sampling for a single contour curve with a length of 100 cm. This method is commonly used to study the size effect. Bandis[33] and Bahaaddini[34] et al. used this method to examine the effect of size effect on the shear behavior of rock joints. The uniform segmentation method refers to the production of small-size samples through uniform division within the maximum sample size range. Small-size samples with original sizes of 1/2, 1/5, and 1/10 can be obtained. Figure 4 shows the sampling of the contour curve of the joints under the series size.

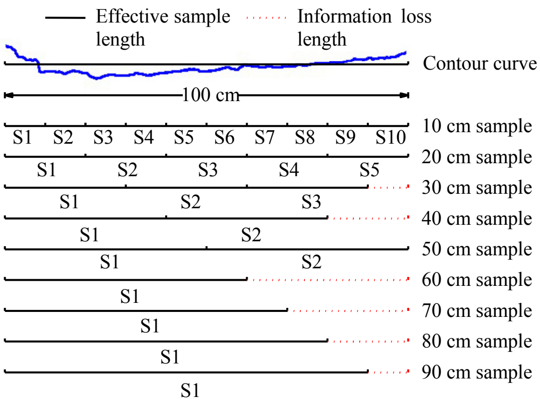


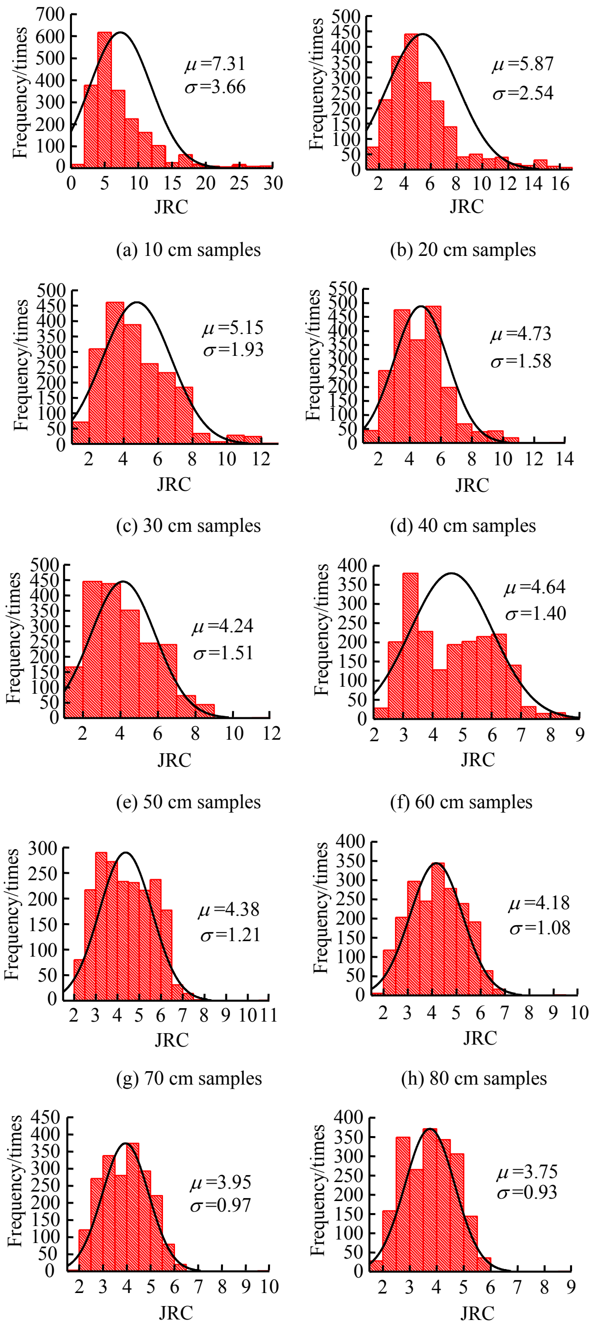
Fig.4 Sampling of structural profile curves in different sizes of samples

The JRC values under different sampling sizes are calculated by the simple formula of Barton's straight edge method (9)[35]. Equation (9) has better adaptability to the structural plane of rock mass with large fluctuation.

$$\mathrm{JRC_n} = 400\frac{R_y}{L_n} \tag{9}$$

where R_y is the width of the tooth profile of the surface profile curve of the joint (cm); and L_n is the length of the profile curve (cm).

The JRC frequency distribution histogram and the total number of JRC samples for series size are shown in Fig.5 and Table 3, respectively. It can be seen that the JRC frequency under the series size basically follows a normal distribution; and the overall JRC mean value and standard deviation under the total sample decrease as the sampling size increases. The average JRC of 60 cm sampling size is greater than the 50 cm sampling size because the uniform segmentation method can equally divide a 100cm long curve into two 50 cm samples, but for the 60 cm sample, only the first 60 cm is taken as the representative, which can reflect the roughness fluctuation information of 60% of the joint. Therefore, the average value of the JRC with a sampling size of 60 cm is not the true value of the JRC of the original rock joint, causing 40% of the information loss. However, the focus of this study is to determine the method of statistically measuring the MNS. The sample mean generally meets the JRC size effect law. The information loss degree of the joint roughness fluctuation under the series size is shown in Fig.6.



(a) 10 cm samples (b) 20 cm samples (c) 30 cm samples (d) 40 cm samples (e) 50 cm samples (f) 60 cm samples (g) 70 cm samples (h) 80 cm samples (i) 90 cm samples (j) 100 cm samples

Fig.5 Histogram of JRC frequencies for different sizes of samples

Table 3 Total number of JRC samples with different sizes

Series size /cm	JRC total number of samples	Series size /cm	JRC total number of samples
10	20 010	60	2 001
20	10 005	70	2 001
30	6 003	80	2 001
40	4 002	90	2 001
50	4 002	100	2 001

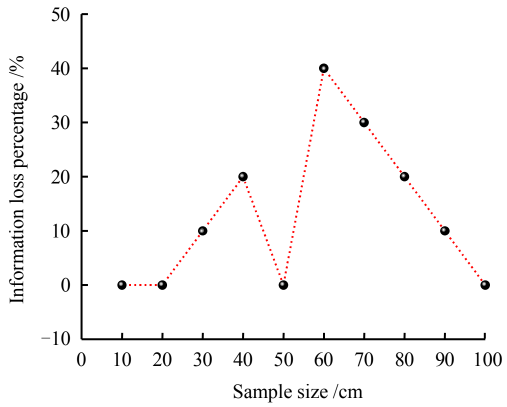


Fig.6 The information loss of joint roughness fluctuation for different size of samples

3.4 Determination of MNS for statistical measurement of joint roughness based on the class ratio analysis of COV

Group JRC samples by different sampling sizes from 10 to 100 cm. In order to ensure its maximum accuracy, the first group has 1 sample, the second group has 2 samples, the third group has 3 samples, ..., the m-th group has k samples (where $m=k$), a total of m groups. Bring each group of corresponding JRC samples into Eqs. (1) and (2) to calculate the average and standard deviation. Based on the average value and standard deviation, the COV of each group is obtained by Eq. (3). Due to the large number of selected samples, this study analyzed the first 100 sets of samples at each sampling size. Figure 7 shows the relationship between the first 100 sets of COV's and the number of samples in the series size.

It can be seen from Figure 7 that for samples of different sizes with a length of 10 to 100 cm, the COV of each group gradually stabilizes as the number of samples increases. The COV value of the small size sample is obviously larger than that of the large size sample, and the COV value decreases with the increase of the sample size. The COV value with a sampling size of 10 to 50 cm is basically stable between 0.31 and 0.47, while the COV value with a sampling size of 60 to 100 cm is basically stable between 0.21 and 0.31. The fluctuation range of the COV value generated by the small sample is larger than that of the large sample, and the fluctuation is weakened as the number of samples increases. When the sample size is large enough, the COV tends to be stable, and the stable value will vary depending on the sample size. Therefore, it is important to determine the reasonable number of samples based on the stable COV for different sampling sizes.

The problem can be solved by class ratio analysis. Substituting the COV of each group into Eq. (4) can obtain the coefficients of class ratio of each group under the series size. The variation rules of the coefficients of class ratios under the series size are almost the same. The overall fluctuation gradually stabilizes with the increase of the number of samples. The class ratio coefficient data points are gradually stabilized within the error range of $1+\varepsilon(\varepsilon=\pm2\%)$ within the first 100 samples; and with the increase of the sample size, the stable state of fluctuations is more and more advanced, the MNS also gradually decreases. Figure 8 is a distribution diagram of the class ratio coefficient at a sampling size of 10 cm.

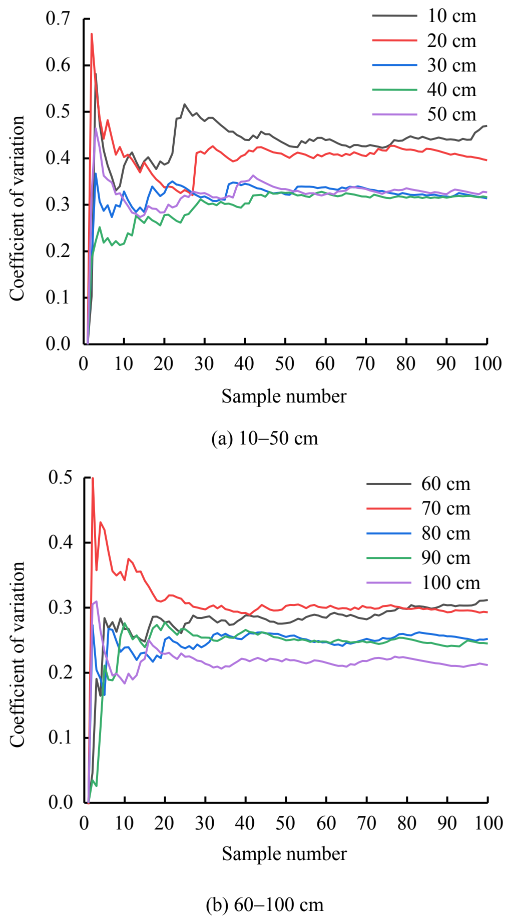


Fig.7 Relationship between COV and sample number of the first 100 groups of different sample sizes

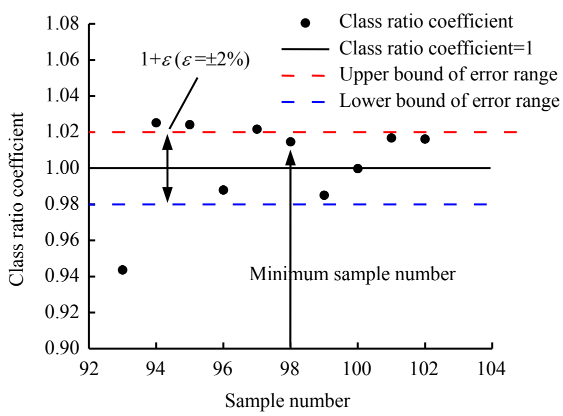


Fig.8 Class ratio coefficient distributions for 10 cm sampling size

As shown in Fig.8, the class ratio coefficient gradually approaches unity as the number of samples increases. Use the error range of $\varepsilon = \pm 2\%$ to determine a reasonable MNS. Observing the scatterplot, we can see that all the data points after the 98 groups fall within the range of 0.98 to 1.02, and the class ratio coefficient meets the error range requirements, indicating that the number of samples represented by this point is the minimum necessary number of samples under the sampling length of 10 cm, that is, the MNS required for statistical measurement of the joint roughness under a sample length of 10 cm is 98.

In this study, the results of the class ratio analysis with a sampling size of 20 to 100 cm in length are similar to the sampling size of 10 cm. The MNS for each sampling size is within 100 groups, and the overall decreases with the increase of the sampling size. Table 4 shows the MNS n_c of the statistical measurement of the joint roughness under the series size. Non-linear fitting is performed on the MNS obtained from the series size, and the fitting equation can be obtained:

$$y = 383.77x^{-0.58} \tag{10}$$

The fitting results show that the MNS corresponding to the series size and the sample size basically satisfy the power function relationship. Fig.9 is the fitting curve of the MNS obtained by the series size.

Table 4 Minimum number of samples for statistical measurement of joint roughness in different sizes

Sampling size /cm	10	20	30	40	50	60	70	80	90	100
n_c	98	77	54	47	43	37	26	33	25	26

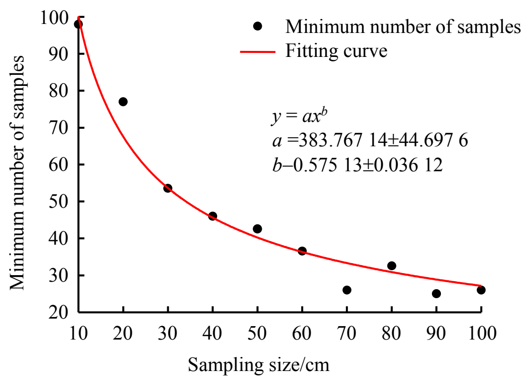


Fig.9 Curve fitting for the minimum number of samples in different sizes

3.5 Determination of MNS for statistical measurement of joint roughness based on simple random sampling principle

The average and standard deviation of the overall JRC sample under the series size are shown in Table 5. According to the maximum relative allowable error (r=10%) between the JRC sampling sample mean and the JRC overall sample mean, the confidence level 95%, and the upper quantile t=1.96 of the standard normal distribution table corresponding to the given confidence level, the MNS for the statistical measurement of the joint roughness under the series size is calculated. The sampling size of 10 cm length is analyzed as an example, where the average value of the overall sample $\bar{Y}$ is 7.31; the standard deviation S is 3.66; t=1.96; and r=0.1. These parameters are taken into Eq. (8) to obtain:

$$n_{s1} = \frac{Nt^2S^2r^2\bar{Y}^2}{Nr^4\bar{Y}^4 + t^2S^2r^2\bar{Y}^2} = \frac{20\,010 \times 1.96^2 \times 3.66^2 \times 0.1^2 \times 7.31^2}{20\,010 \times 0.1^4 \times 7.31^4 + 1.96^2 \times 3.66^2 \times 0.1^2 \times 7.31^2} \approx 96$$

Thus, the MNS required for statistical measurement of the joint roughness at a sample length of 10 cm is 96. Considering the data in Table 5, the MNS n_{s1} for the statistical measurement of the joint roughness under series sizes can be calculated using Eq.(8).

Different allowable errors and different confidence levels in the sampling survey will lead to differences in the calculated sample size. Considering the JRC maximum relative allowable error r=8%, 95% confidence level, and JRC maximum relative allowable error r=10%, 90% confidence level, respectively, the MNS for statistical measurement of the joint roughness are recorded as n_{s2} and n_{s3}, respectively. The mean and standard deviation of overall sample are substituted into Eq.(8) for the cases of t=1.96, r=0.08 and t=1.64, r=0.1, the n_{s2} and n_{s3} under the series size can be calculated, as shown in Table 5.

Table 5 Calculated parameters and minimum number of samples in different sizes

Sampling /cm	Mean value	Standard deviation	n_{s1}	n_{s2}	n_{s3}
10	7.31	3.66	96	150	68
20	5.87	2.54	72	112	51
30	5.15	1.93	54	84	38
40	4.73	1.58	43	66	30
50	4.24	1.51	49	75	34
60	4.64	1.40	35	54	25
70	4.38	1.21	29	45	21
80	4.18	1.08	26	40	18
90	3.95	0.97	23	36	17
100	3.75	0.93	24	37	17

3.6 Calculation results and comparative analysis of different MNS determination methods

The comparison between the calculation results of the MNS based on the COV class ratio analysis and the simple random sampling principle is shown in Fig.10.

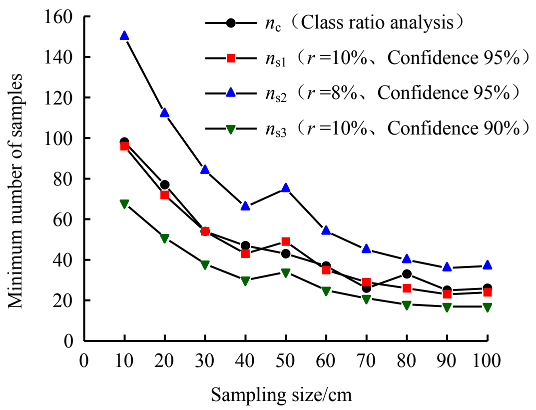


Fig.10 Comparison of minimum numbers of samples

The MNS n_c determined based on the COV class ratio is the same as the MNS's n_{s1}, n_{s2}, and n_{s3} based on the simple random sampling principle considering different relative allowable errors and confidence levels. The MNS generally decreases with increasing sample size.

In order to quantitatively compare the similarity of the two MNS determination methods, the vector similarity measurement method is now used for comparative analysis of the four sets of calculation results. The similarity functions of general vectors include angle cosine method[36−37], generalized Dice coefficient method[37−39], generalized Jaccard coefficient method[37, 40], correlation coefficient method[36] and so on.

Among them, the generalized Dice coefficient method is similar to the angle cosine method. The advantage is that its numerator fully considers the influence of the common term value between the vectors X and Y, and its value is doubled. The denominator takes into account the weights of the scoring terms between the vectors X and Y, and squares the non-zero terms to make the similarity calculation result more accurate. Therefore, in this paper, the generalized Dice coefficient method is used to measure the MNS in the series size obtained by different determination methods, the expression is

$$D(X,Y)=\frac{2X\cdot Y}{\|X\|_2^2+\|Y\|_2^2}=\frac{2\sum_{i=1}^{n}x_i y_i}{\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}y_i^2} \tag{11}$$

where $X=(x_1,x_2,\ldots,x_n)$ and $Y=(y_1,y_2,\ldots,y_n)$ are two n-dimensional space vectors; $X\cdot Y=\sum_{i=n}^{n}x_i y_i$ is the inner product of the two vectors; $\|X\|_2=\sqrt{\sum_{i=1}^{n}x_i^2}$ and $\|Y\|_2=\sqrt{\sum_{i=1}^{n}y_i^2}$ are the second norm of the vectors X and Y.

The MNS's n_c, n_{s1}, n_{s2}, and n_{s3} obtained under four sets of series sizes are converted into corresponding vector data:

$\vec{n}_c=[98,77,54,47,43,37,26,33,25,26]$

$\vec{n}_{s1}=[96,72,54,43,49,35,29,26,23,24]$

$\vec{n}_{s2}=[150,112,84,66,75,54,45,40,36,37]$

$\vec{n}_{s3}=[68,51,38,30,34,25,21,18,17,17]$

Suppose the vector $\vec{n}_c$ is a standard vector, $\vec{n}_{s1}$, $\vec{n}_{s2}$ and $\vec{n}_{s3}$ are the vectors to be compared. The vectors are calculated by Eq.(11). The calculation results of similarity are shown in Table 6.

Table 6 Vector similarity calculation results

Similarity	$\vec{n}_c$ and $\vec{n}_{s1}$	$\vec{n}_c$ and $\vec{n}_{s2}$	$\vec{n}_c$ and $\vec{n}_{s3}$
$D(X, Y)$	0.997 1	0.918 0	0.932 3

It can be seen that the MNS n_c determined based on the COV class ratio analysis is the most similar to the MNS n_{s1} when the simple random sampling principle considers the maximum relative allowable error r=10% and the confidence level 95%. The similarity is greater than 0.997.

Different error ranges and confidence levels will lead to differences in the number of calculated samples. The maximum allowable error range and confidence level of the MNS can be obtained through the similarity measurement of the calculation results of the two MNS determination methods, which provides a basis for accurate evaluation of JRC statistical measurement results and for accurate acquisition of rock mass mechanical parameters.

4 Discussion

It can be seen from Fig.7 that the COV of the 10 cm sample is significantly larger than that of other sizes and has the largest fluctuations. Yong et al.[14] concluded that if the number of samples is large enough with a stable COV for small samples, then for large samples, the number of samples is also sufficient. According to the requirements of the statistical sample number, the statistical analysis of the sample sizes is performed, and the obtained statistical results can also meet the statistical requirements, but the effective number of samples required for different sampling sizes is often different. It is not likely relate the sample size with universal investigation purpose. Therefore, it is not appropriate to simply think the minimum sample size obtained from a small size sample as to be effective. In contrast, it is necessary to determine the MNS for statistical measurement separately for the series size.

In the study of rock shear strength, Xia et al. [41] obtained nine contour curves with 30 cm long at equal intervals of 1.5 cm along the shear direction durin the JRC sampling of a 30 cm× 15 cm structural plane. This sampling number only accounts for 16.7% of the sample number obtained by class ratio analysis method. The number of samples is obviously insufficient, which will cause distortion of the JRC calculation results. WANG et al.[42] studied the effect of shear rate on the mechanical behavior of rock joints. When sampling JRC at a 20 cm×10 cm rock joint, in order to obtain a more accurate average value of JRC, the JRC at different contour intervals was calculated and compared to determine a better contour interval. This idea is similar to the class ratio analysis method, but no specific method and quantified sample number were given. Therefore, it is necessary to determine the MNS standard for statistical measurement for different sampling sizes.

This paper proposes a method of determining the MNS based on the class ratio analysis of the COV and the principle of simple random sampling, which can quantitatively obtain the MNS for the statistical measurement of the roughness of a series of joints within the guaranteed range of error. This research method started with selection of the representative joint samples in the tested area, which can be promoted and used on the joint with the same properties. When the lithology has obvious changes or the roughness properties are obviously different, the results need to be re-calculated to determine the new statistical measurement patterns for MNS.

5 Conclusions

Aiming at the problem that it is impossible to determine a reasonable number of samples in the statistical measurement of joint roughness at present, a method for determining the MNS based on the COV class ratio analysis and the principle of simple random sampling is proposed. By comparing the calculation results of the two methods, the following conclusions can be drawn:

The JRC statistical results of the series size showed that the COV value of the small size sample was significantly larger than that of the large size sample, and the COV value decreased with the increase of the sample size. The COV value of the sample size of 10−50 cm was basically stable at 0.31−0.47, and that of 60−100 cm was basically stable between 0.21−0.31.

The two methods for determining the minimum sample size revealed that the MNS of the statistical measurement corresponding to the series size were different. The MNS and the sample size basically satisfied the power function relationship, and the MNS decreased as the sample size increased.

The comparison of the calculation results between the two MNS determination methods showed that the MNS determined by the class ratio analysis method when the allowable error was ±2% was consistent with the that calculated by the simple random sampling principle when the maximum allowable error was 10% and the confidence level was 95%, and the similarity was greater than 0.997.

References

[1] LI Hua, HUANG Run-qiu. Study on quantitative determination method of roughness coefficient of rock structure surface[J]. Chinese Journal of Rock Mechanics and Engineering, 2014, 33(Suppl.2): 3489−3497.

[2] YE Jun, YONG Rui, LIANG Qi-feng, et al. Neutrosophic Functions of the joint roughness coefficient and the shear strength: a case study from the pyroclastic rock Mass in Shaoxing City, China[J]. Mathematical Problems in Engineering, 2016(1): 1−9.

[3] ZHOU Hui, CHENG Guang-tan, ZHU Yong, et al. A new method to originally reproduce rock joint by integrating 3D scanning and 3D carving techniques and mechanical characteristics of reproduced structural planes[J]. Rock and Soil Mechanics, 2018, 39(2): 417−425.

[4] DU Shi-gui. Accuracy evaluation method for slope stability of large open pit mines[J]. Chinese Journal of Rock Mechanics and Engineering, 2018, 37(6): 1301−1331.

[5] BARTON N. Review of a new shear strength criterion for rock joints[J]. Engineering Geology, 1973, 7(4): 287−332.

[6] DU Shi-gui. Study on mechanical effect of joint surface of rock mass[J]. Geoscience, 1994(2): 198−208.

[7] DU Shi-gui. Orientational statistical study on the roughness coefficient JRC of rock mass structure[J]. Journal of Engineering Geology, 1994, 2(3): 62−71.

[8] DU Shi-gui, TANG Hui-ming. Anisotropy study on fracture roughness coefficient of rock mass[J]. Journal of Engineering Geology, 1993, 1(2): 32−42.

[9] DU Shi-gui, YAN Yu-ren, HU Xiao-fei, et al. Average slope method for estimating shear strength of JRC-JCS model[J]. Journal of Engineering Geology, 2005, 13(4): 489−493.

[10] DU Shi-gui. Method for determining roughness coefficient of rock mass structure surface: China, 200410084467.5[P]. 2006-05-31.

[11] JI Feng. Quantitative determination of roughness coefficient of hard structure surface and its engineering application[J]. Hydrogeology and Engineering Geology,

2010, 37(3): 84−86.

[12] WU Yue-xiu, LIU Quan-sheng, LIU Xiao-yan. Study on correlation between roughness coefficient of rock mass and statistical parameters[J]. Chinese Journal of Rock Mechanics and Engineering, 2011, 30(Suppl.1): 2593−2598.

[13] MA Cheng-rong, HUANG Man, DU Shi-gui, et al. Statistical method for structural surface roughness coefficient under optimal sampling interval: China, 201710822964.8[P]. 2018-01-09.

[14] YONG Rui, YE Jun, LI Bo, et al. Determining the maximum sampling interval in rock joint roughness measurements using Fourier series[J]. International Journal of Rock Mechanics and Mining Sciences, 2018, 101: 78−88.

[15] YONG Rui, QIN Ji-bo, HUANG Man, et al. An innovative sampling method for determining the scale effect of rock joints[J]. Rock Mechanics and Rock Engineering, 2019, 52: 935−946.

[16] MI Jun-nan. Study on influence of parameter variability and sample capacity on slope reliability[D]. Hangzhou: Zhejiang University, 2007.

[17] GAO Da-zhao. Discussion on calculation formula of standard value of geotechnical design parameters[J]. Geotechnical Investigation & Surveying, 1996(3): 5−8.

[18] DONG Jiao-she, CHENG Li-min. Discussion on the reasonable sample number in geotechnical engineering investigation[J]. Port Engineering Technology, 1999(2): 52−54.

[19] LI Zheng-yi. Discussion on statistical analysis of geotechnical samples[J]. Geotechnical Engineering Technique, 2008, 22(3): 109−112.

[20] ZHAO Jing, JIANG Liang-wei, LUO Qiang, et al. Discussion on the number of reasonable samples of rock and soil in slope reliability analysis[J]. Journal of Railway Science and Engineering, 2016, 13(6): 1053−1060.

[21] Ministry of Construction of People′s Republic of China. GB 50021－94 Geotechnical survey specification[S]. Beijing: China Architecture & Building Press, 1995.

[22] DENG Ju-long. On judging the admissibility of grey modelling via class ratio[J]. The Journal of Grey System, 1993, 5: 249−252.

[23] DENG Ju-long. Grey prediction and grey decision[M]. Wuhan: Huazhong University of Science and Technology Press, 2002.

[24] HUANG Qiang. Talking about several technical problems in system sampling[J]. Tatistical Science and Practice, 1997(3): 17−18.

[25] LUO Zhan-you, DU Shi-gui, HUANG Man. Push-pull test study on size effect of rock surface roughness coefficient[J]. Rock and Soil Mechanics, 2015, 36(12): 3381−3386.

[26] BARTON N, CHOUBEY V. The shear strength of rock joints in theory and practice[J]. Rock Mechanics, 1977, 10(1): 1−54.

[27] DEVELI K, BABADAGLI T, COMLEKCI C. A new computercontrolled surface-scanning device for measurement of fracture surface roughness[J]. Computers and Geosciences, 2001, 27(3): 265−277.

[28] DU Shi-gui. Simple longitudinal profiler and its application in the study of coarse sugar content coefficient of rock mass structure[J]. Geological Science and Technology Information, 1992, 11(3): 91−95.

[29] BAE D S, KIM K S, KOH Y K, et al. Characterization of joint roughness in granite by applying the scan circle technique to images from a borehole televiewer[J]. Rock Mechanics and Rock Engineering, 2011, 44(4): 497−504.

[30] FARDIN N, STEPHANSSON O, JING L. The scale dependence of rock joint surface roughness[J]. International Journal of Rock Mechanics and Mining Sciences, 2001, 38(5): 659−669.

[31] MLYNARCZUK M. Description and classification of rock surface by means of laser profilometry and mathematical morphology[J]. International Journal of Rock Mechanics and Mining Sciences, 2010, 47(1): 138−149.

[32] SONG Lei-bo, JIANG Quan, LI Yuan-hui, et al. Study on stability of statistical parameters of structural surface morphology and anisotropy at different sampling intervals[J]. Rock and Soil Mechanics, 2017, 38(4): 1121−1132, 1147.

[33] BANDIS S C, LUMSDEN A C, BARTON N R. Experimental studies of scale effects on the shear behaviour of rock joints[J]. International Journal of Rock Mechanics and Mining Science & Geomechanics Abstracts, 1981, 18(1): 1−21.

[34] BAHAADDINI M, HAGAN P C, MITRA R, et al. Scale effect on the shear behaviour of rock joints based on a numerical study[J]. Engineering Geology, 2014, 181: 212−223.

[35] HU Xiao-fei, DU Shi-gui. Concise formula of Barton straight edge method for structural surface roughness coefficient[J]. Journal of Engineering Geology, 2008, 16(2): 196–200.

[36] ZHANG Yu, LIU Yu-dong, JI Zhao. Vector similarity measure[J]. Technical Acoustics, 2009, 28(4): 532–536.

[37] YONG Rui, YE Jun, LIANG Qi-feng, et al. Estimation of the joint roughness coefficient (JRC) of rock joints by vector similarity measures[J]. Bulletin of Engineering Geology and the Environment, 2018, 77(2): 735–749.

[38] DICE L R. Measures of the amount of ecologic association between species[J]. Journal of Ecology, 1945, 26: 297–302.

[39] TAO Yue-hua. Vector-based similarity calculation scheme[J]. Journal of Yunnan Normal University (Natural Sciences Edition), 2001, 21(5): 17−19.

[40] JACCARD PAUL. Distribution de la flore alpine dans le Bassin des Drouces et dans quelques regions voisines[J]. Bulletin de la Société Vaudoise des Sciences Naturelles. 1901, 37(140): 241–272.

[41] XIA C C, TANG Z C, XIAO W M, et al. New peak shear strength criterion of rock joints based on quantified surface description[J]. Rock Mechanics and Rock Engineering, 2014, 47(2): 387–400.

[42] WANG G, ZHANG X P, JIANG Y J, et al. Rate-dependent mechanical behavior of rough rock joints[J]. International Journal of Rock Mechanics and Mining Sciences, 2016, 83: 231–240.