Study on Sample Size of Small Batch Precast Concrete Components

Lin Gao¹,², Danxuan Liu¹,², Yingli Liu¹,²
¹College of Civil and Architectural Engineering, North China University of Science and Technology, Tangshan 063210, P.R.China;
²Earthquake Engineering Research Center of Hebei Province, Tangshan 063210, P.R.China
gaolin_heut@126.com

Abstract. In the sampling test, in the absence of actual data of the eigenvalues, combined with the actual conditions of the project, the total coefficient of variation under various distribution types can be used to estimate the sample size required for sampling. The article studies the sampling inspection of the strength of small batches of precast concrete components, and obtains the overall coefficient of variation under the normal distribution type, and then it can calculate the sample size required for different total amount of corresponding sampling inspection.

1. Introduction
At present, the Code for Acceptance of Construction Quality of Concrete Structures (GB50204-2015) adopts a percentage sampling method within the inspection lot for the sampling of approach components of prefabricated components. Due to the large number of components in the same batch at the production stage, the percentage sampling has serious shortcomings of "big lot and strict loosening". In addition, the sample quality level does not completely be equal to the batch quality level, and the determination of the sampling proportion of the percentage sampling test also lacks theory in accordance with.

Sampling inspection is to take samples from the overall product and estimate the overall quality of the product by testing product samples. However, the key to inferring the overall quality as accurately as possible through test samples is to use scientific sampling methods to extract reasonable sample sizes. Probability theory and mathematical statistics are the disciplines that study the stochastic phenomena and find their statistical regularity. They have very important applications in the fields of natural sciences, social sciences, and engineering technology. The application of this discipline to the sampling inspection of prefabricated components is more scientific and effective, so that the quality of a certain number of samples of the inspection more accurately represents the overall quality [1].

For the problems existing in the existing experience sampling, for the different detection items, a scientific sampling method for the quality inspection of architectural parts based on the probability theory was proposed for the first time-the overall coefficient of variation method [2]. Without the actual data of eigenvalues, combined with the actual situation of the project, the total coefficient of variation under various distribution types can be used to quickly calculate the size of the required sample size. Because the strength value of concrete is measured by the mean $\mu$ and obeys the normal distribution, the sample's total coefficient of variation is investigated for the type of normal distribution in the...
sampling test and the required sample size for the inspection of the preform components is calculated. This study discusses the use of probabilistic and mathematical statistics for low-volume precast concrete components to obtain a reasonable sample size.

2. Theoretical basis and formula derivation

In the practice of sampling inspection, the first and foremost technical problem is usually the amount of sample required for the investigation. The influencing factors of the sample size mainly include: the time limit, the expense limit and the accuracy requirement. The three are mutually restricted and interconnected, and the general accuracy requirement is often at the forefront. Most of the sampling inspections do not replace simple random sampling[3].

Usually with sample mean \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \) Inferring the overall mean \( \bar{X} \). The variance of \( \bar{X} \) is

\[
V(\bar{X}) = \frac{N - n}{nN} S^2 \quad \text{(Where N is the total number of units, S}^2 \text{ is the overall variance of the target variable, n is the sample size, the same)}.
\]

If the accuracy requirement of the estimator \( \bar{X} \) in advance is that the variance upper limit is \( V \), which is \( V(\bar{X}) = \frac{N - n}{nN} S^2 \leq V \), according to the conservative principle, you can use \( \frac{N - n}{nN} S^2 = V \) to calculate the sample size needed for the survey, get

\[
n = \frac{S^2 / V}{1 + S^2 / NV}.
\]

If \( n_0 = \frac{S^2}{V} \), then there are:

\[
n = \frac{n_0}{1 + n_0 / N} \quad (1)
\]

And when \( N \) is very large, there is \( n \approx n_0 \).

If the accuracy requirement given in advance is the absolute error limit \( d \), there are

\[
v = \left( \frac{d}{u_\alpha} \right)^2,
\]

(\( u_\alpha \) is the standard normal distribution function value \( \Phi(1 - \alpha / 2) \), the same below), get the formula here:

\[
n_0 = \left( \frac{u_\alpha S}{d} \right)^2 \quad (2)
\]

If the accuracy requirement given in advance is relative error limit \( r \), there are \( d = r\bar{X} \), brought into equation (2) to get

\[
n_0 = \left( \frac{u_\alpha S}{r\bar{X}} \right)^2, \quad \text{where} \quad \frac{S}{X} \text{ is the coefficient of variation of the overall indicator, marked as CV, get the formula here:}
\]

\[
n_0 = \left( \frac{u_\alpha}{r} \right)^2 (CV)^2 \quad (3)
\]

Thus, the accuracy of the estimator requires that the upper limit of variance, the absolute error limit, and the relative error limit be given in advance. To obtain the final sample size \( n \), you also need to know the variance of the overall indicator variance \( S^2 \) or the standard deviation \( S \) or the total coefficient of variation CV. However, before the survey, these values are often not known, and they can only be roughly estimated using relevant information. This is the so-called pre-estimation of the degree of overall variation[4].
In sampling surveys, the overall population faced is usually a non-negative population, i.e., the value of survey indicators is generally non-negative, such as concrete strength values. Therefore, this study sets the minimum value of the survey overall to \( t \) and \( t \geq 0 \).

![Figure 1. Normal distribution curve.](image)

The curve shown in Figure 1 is the normal overall curve \( N(\mu, \sigma^2) \), let the overall range difference be \( R \), then its mean \( \mu = t + R / 2 \). According to the conclusion of reference [5], the approximate pre-estimate of the population standard deviation is \( \hat{\sigma} = \frac{R}{0.5 \ln(n) + 3} \). Therefore, the coefficient of variation is:

\[
CV = \frac{R/[0.5 \ln(n) + 3]}{(t + R / 2)}
\]

(4)

Combining with the actual situation of the project, different prefabricated component plants produce different batches of components, their strength values are different, and the extreme value \( R \) and the minimum value \( t \) cannot be determined because the minimum value \( t \geq 0 \), so when the minimum value \( t=0 \) is assumed, the actual normal distribution curve is translated in the coordinate axis and does not have much influence on the actual normal distribution curve. Therefore, simplifying equation (4) results in:

\[
CV = \frac{2}{0.5 \ln(n) + 3}
\]

(5)

The above equations (1)(3)(5) can be obtained by:

\[
n_0 = \left( \frac{u_\alpha}{r} \right)^2 \left( \frac{2}{0.5 \ln(\frac{n_0}{r n_0 / N}) + 3} \right)^2
\]

(6)

Where \( u_\alpha \) is the standard normal distribution function value \( \Phi(1-\alpha / 2) \), this study has a confidence level of 95%, where \( \alpha / 2 \) is 0.25. Taking the rebound tester to measure the strength of concrete as an example, according to the research results of the reference, the relative error limit \( r \) is about 10%, and the total \( N \) value is evenly selected from 50-1000.

When \( N \) approaches infinity, the above formula (6) is a convergence function, and the limit values for calculating \( n_0 \) using MATLAB software are:

\[
\lim_{N \to \infty} n_0 = 60.2687
\]

(7)
When N is large, there is $n \approx n_0$, so the sample size n approaches to 60.2687.

3. Theoretical numerical simulation results
According to the above equation (6), the value of $n_0$ can be obtained, and then the sample amount n can be obtained according to the equation (1). The above operation is performed using MATLAB software, and the operation data is shown in Table 1 below.

| Total amount | n0     | n (Sample size) |
|--------------|--------|-----------------|
| 50           | 69.9666| 29.1609         |
| 100          | 65.6070| 39.6161         |
| 150          | 63.9595| 44.8399         |
| 200          | 63.0904| 47.9610         |
| 250          | 62.5530| 50.0339         |
| 300          | 62.1876| 51.5010         |
| 350          | 61.9231| 52.6144         |
| 400          | 61.7227| 53.4717         |
| 450          | 61.5656| 54.1563         |
| 500          | 61.4392| 54.7158         |
| 550          | 61.3352| 55.1814         |
| 600          | 61.2482| 55.5751         |
| 650          | 61.1743| 55.9122         |
| 700          | 61.1108| 56.2041         |
| 750          | 61.0556| 56.4594         |
| 800          | 61.0072| 56.6845         |
| 850          | 60.9644| 56.8845         |
| 900          | 60.9263| 57.0633         |
| 950          | 60.8921| 57.2242         |
| 1000         | 60.8614| 57.3698         |
| 1050         | 60.8335| 57.5020         |
| 1100         | 60.8081| 57.6227         |
| 1150         | 60.7849| 57.7333         |
| 1200         | 60.7636| 57.8350         |
| 1250         | 60.7440| 57.9289         |
| 1300         | 60.7259| 58.0158         |
| 1350         | 60.7092| 58.0966         |
| 1400         | 60.6936| 58.1717         |
| 1450         | 60.6791| 58.2418         |
| 1500         | 60.6655| 58.3073         |

According to the simulation results can be drawn:
When the total N is selected from 50 to 500, the sample size n increases from 29 to 54 and the growth rate is relatively large. When N is selected from 550 to 1500, the sample size n increases from 55 to 58, and the growth rate is decreases. This trend shows that, as a whole, the sample size n of construction components and fittings increases with the increase in the total number N of inspection lots, but the rate of increase of sample size shows a decreasing trend, indicating that when the total amount N is large enough, its sample size converges to a fixed value, the limit value \( n \approx n_0 = 60.2687 \), and its sample size does not increase with the increase of the total quantity.

4. Conclusion

To sum up, according to the calculation of the coefficient of variation data above, and combined with the actual situation of the project, it can be concluded that when the total amount of the concrete strength of the small batch of precast concrete components and fittings is 50, the minimum sample size is 29. At 1500, the minimum sample size is 58. On the whole, the sample size n of construction components and fittings increases with the increase in the total number N of inspection lots, but the sample size continuously approaches the limit value of 60.2687. When the total amount N is large enough, the sample size no longer increases with the increase in the total amount.

This research comes from the sub project of national key research and development program: Research on defect detection technology of precast components (2016YFC0701801-01).

References

[1] Wu Yefeng. The upper limit of the overall variation coefficient of common surveys and its application in the rapid estimation of sample size[J]. Statistics and Decision, 2011(03):15-18.
[2] Guan Yu, Gu Guangtong. Interval estimation of overall proportion and determination of sample size [J]. Statistics and Information Forum, 2017, 32 (02): 16-20.
[3] Gong Jingqi. The accuracy of measuring concrete strength with a rebound meter[J]. Concrete and Cement Products, 1982(03):46-49+38.
[4] Gong Jingqi. Accuracy problem of measuring the concrete strength with a rebound meter[J]. Harbour Engineering, 1981(01): 43-48.
[5] Du ZF, Wu YF. An indirect method of estimating the variance of normal population[J]. Statistics and Decision, 2005(08):10-13.