Distribution of the Coefficient of Variation of the Continuous Sample in the Electronic Testing of the Raw Silk Size

Jian-Mei Xu*, Lun Bai*

School of Material Science, Soochow University, Suzhou 215021, Jiangsu Province, China

Abstract: As we know, the coefficient of variation (CV) is one of the main quality indexes in the electronic testing for raw silk. Therefore, to study the classification of this index, it's nessary to get the distribution and other statistical information of this variable. In our previous study, we have gotten its distribution when the population is normal and the sampling is random. However, in the real testing, the CV value is obtained by computing a series of continuous silk size tested by the electronic machine. Due to the continuity of the silk filament, the series of raw silk size can be regarded as a continuous sample. The characteristics of the size from the continuous sample will be different from that of the random sample. This paper analyzes the results of the continuous sampling tests of the CV of the raw silk size, and deduces the distribution of CV of the raw silk size and its characteristics in the case of continuous sampling test, which provides a theoretical basis for the development of electronic testing standard.

Keywords: random sampling, continuous sampling, distribution of the coefficient of variation, equivalent sample

1. Introduction

As we know, the electronic testing system has been widely used in the textile industries of chemical fiber, cotton and wool for its objective testing results and high efficiency. However, in the silk industry only some enterpries in the silk consuming countries are using the electric testing method. There is no global use and universal standards in the electronic testing and classification. However, the demand for silk is global; the international trade of raw silk demands industry standards for raw silk quality to enable the buyers to purchase raw silk at internationally accepted grades. Therefore, it is of great importance for us to make series of researches on the electronic testing and classificaiton.

In the electronic testing system, coefficient of variation (CV) is an important quality index as it embodies the relative variations of the raw silk size. Here the size means the linear density of the fiber. In the testing, we sample a bobbin from a silk lot, then sample a long segment of raw silk from the bobbin [1], and test its size in a certain testing length, thus get a size series, from which we can get CV of the raw silk size of the long segment. Repeating this test several times [1], we get the average value of CV of the raw silk size, thus we can use the average value to estimate the CV of the raw silk size of this silk lot.

To scientifically and rationally make the raw silk testing standard that takes the size coefficient of variation as the quality index, we need to find the sampling distribution of the size coefficient of variation and its statistic characteristics of the continuous sample in the electronic testing of raw silk size. In our previous research [2], we have found the distribution of the CV of the random sample from the normal population. In the raw silk sampling test, we can assume that the size of a whole lot of raw silk has normal distribution [3]; and this assumption is based on the real testing results that had been done by many reseachers [4]. In the traditional testing, sample of the raw silk size is a random sample. However, in the electronic testing the sample is a long segment of raw silk, the test is also a continuous process, so what we get is not a random sample, but a continuous sample. Here "continuous" means the data of the raw silk size tested by the machine is one after another. At this time due to the continuous change of the raw silk size, the

* Corresponding author's email: *xujianmei@suda.edu.cn; *bailun@suda.edu.cn

data of the sample have correlation with each other. It is very difficult to get the sampling distribution of the CV when the sample is continuous. By analyzing the statistical character of the CV of the continuous sample, this paper get a random sampling distribution of the CV that is equivalent to a continuous sampling distribution of the CV, thus make the CV of the continuous sample take an equivalent expression of the random sample. Therefore, we can get the distribution of the CV of the continuous sample, as well as its mean and variance. By simulation, the feasibility and the accuracy of the equivalence are verified.

2. Continuous Sampling Method and Random Sampling Method

Assume that we are testing the size of a lot of raw silk, the testing length is meters, then we can get a size series $\{x_1, x_2, x_3, \mathrm{L}\ x_N\}$ of this lot, we may regard the series as a normal population $N(\mu, \sigma^2)$ whose samples have correlation with each other, and μ , $\sigma^2$are the mean and variance of the raw silk size respectively[3].

In the electronic raw silk testing, we sample a long silk segment from this lot of raw silk, and test its size every l meter continuously, and this kind of sampling method is regarded as the continuous sampling method. Assume that n data of the raw silk size are obtained by this method, and the series of data can be expressed as $\{x_1, x_2, x_3, \mathrm{L}\ x_n\}$, then x_i is a variable of the raw silk size of a certain length, thus the expected value and variance of the raw silk size can be written respectively as

$$E(x_i) = \mu, \quad D(x_i) = \sigma^2 .$$

As the raw silk size is tested continuously, the series of the data have a periodic change, from thick to thin, and then from thin to thick, thus change in cycles. Therefore, the individual datum of the series is not independent, and there are certain correlations among them. In the statistics, this correlation is usually expressed by the autocorrelation coefficient

$$\rho_s = \frac{E[(x_i - \mu)(x_{i+s} - \mu)]}{\sigma^2},$$

where s is the potential difference.

As the autocorrelation coefficient of the raw silk size attenuates gradually, and the attenuation process takes on a cosine change, an autocorrelation coefficient model [5] of the raw silk size is established, i.e.

$$\rho_s = e^{-\lambda s} \cos\frac{2\pi s}{T}, \quad (s = 0, 1, 2, \mathrm{L}\ \mathrm{L}\ )$$

where parameter λ is the attenuation rate, which expresses the attenuating speed of the autocorrelation coefficient; T is the mimetic period, which embodies the change period of the raw silk size. In this paper, the values of λ and T are all based on the fact that the testing length is l meters long.

The average and variance of the raw silk size of the continuous sample are as follows

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i , \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 ;$$

as n in the electronic testing is always very large, there will have no big error, when the variance of the raw silk size is written as

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \tag{1}$$

Coefficient of variation of the raw silk size of the sample is

$$c = s / \bar{x} . \tag{2}$$

Repeat this test for N times; we will get the average of the sample $\bar{c} = \frac{1}{N}\sum_{i=1}^{N} c_i$, which can be regarded as the estimation of the size coefficient of variation of this lot of raw silk. When , this value will be approximately equal to the CV of the population,

$$E(c) = \sigma / \mu \tag{3}$$

On the other hand, random sampling is that we get n samples from the size series $\{x_1, x_2, x_3, \cdots x_N\}$