Characteristic functions and the couples method when calculating the system

Evgeny Bulyzhev¹,* Andrey Bogdanov² and Leonid Khudobin¹

¹ Ulyanovsk state technical university, Ulyanovsk, Russia
² Ulyanovsk state university, Ulyanovsk, Russia

* eugbul1946@gmail.com

Abstract. This paper describes the application of the couples method and of characteristic functions when calculating purification systems of cooling lubricants. There are main mathematical models for calculating the characteristic function and statistical couples described.

When resolving practical issues it is often necessary to characterize statistical distributions i.e. to select the identification of theoretical dependencies describing actually observed data. Identification is usually carried out if values of integrated numerical characteristics of distributions accumulating their primary qualities are preserved. Ordinary couples are usually used as such characteristics. As it is noted in [1] it is very likely that if two distributions have a certain number of similar couples, then they are similar to some extent and approximate each other. This argument can be substantiated with the coincidence of their approximations using the least square method at the final range and distributions in orthogonal polynomials. Such practical approximation turns out to be satisfactory even if only the first three or four couples coincide. In the same paper there are necessary and sufficient conditions of the opportunity described to plot the characteristic function using couples (hypotheses of the fore quoted theorem are sufficient) which uniquely determines the distribution at all points of its continuity. These conditions, usually, are met. However, there are some negative examples (the logarithmic normal distribution).

At all its disadvantages (more often potential than real) the couples method still remains easy-to-use and is usually applied for identifying distributions. Number of preserved couples is considered as the approximation degree. It is equal to the number of free parameters of the theoretical curve. Please note that the number of such parameters for the most of distributions does not exceed two (normal, logarithmic-normal, even, exponential, Cauchy and etc.).

It shall be also noted that:

• ramping up the number of recorded couples is not always accompanied by the monotonous improvement of the approximation quality;
• any number of recorded couples cannot provide a good fit onto «the tail» of the distribution (on the calculation experience in the probability zone less than 0,01 and more 0,99);
• the relative accuracy of the highest couples obtained using statistical methods breaks away with the increase of the couple order.

That is why they are practically limited by the identification of not less than two and not more than four couples.
In the applied mathematical statistics there are methods giving an opportunity to automatically select the distribution type in the given family (for example, one of K. Pearson curves is selected using four couples, ref. [2]). Fitting criteria (R. Fisher, N. V. Smirnov, A. N. Kolmogorov) of statistical and theoretical distributions with several general couples are insensitive to the distribution type. This gives an opportunity in the functional class with given several couples to select distribution types which provide these or those computational advantages. Looking at the calculation of the purification systems of cooling lubricants these are the normal distribution and the logarithmic-normal distribution [3, 4].

Suppose \( F_n(x) = \Phi \left( \frac{\ln x - a_n}{s_n^2} \right), x > 0 \) is the distribution function of the original sludge in the cooling lubricant before the purification (at diameters \( x \)). Here \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} \, dt \) is the distribution function of the standard normal distribution (the Laplace function). Then the distribution density of original sludge particles is written as

\[
f_n(x) = \begin{cases} 
\frac{1}{\sqrt{2\pi s_n^2 x}} \exp \left( -\frac{(\ln x - a_n)^2}{2s_n^2} \right), & \text{at } x > 0, \\
0, & \text{at } x \leq 0,
\end{cases}
\]

where \( a_n = \ln \left( \frac{\ln d}{\sqrt{s^2 + \ln d}} \right) \) is the quasi-mathematical expectation,

\[
s_n^2 = \ln \left( \frac{s^2}{\ln d} + 1 \right)
\]

(is the quasi-dispersion)

(here \( d \) is the true mathematical expectation, \( s^2 \) is the dispersion).

The formula for the \( n \)-couple of the logarithmic-normal random value \( \xi \) is written as

\[
E\xi^n = \exp \left( na + \frac{n^2s^2}{2} \right), \quad n \geq 1.
\]

Wherefrom, in particular:

\[
E\xi = \exp \left( a + \frac{s^2}{2} \right), \quad D\xi = \left( \exp \left( s^2 \right) - 1 \right) \exp \left( 2a + s^2 \right). \quad \text{Introduce some the most important characteristics of the logarithmic-normal distribution (ref. Table 1).}
\]

Calculate the approximately characteristics function of some centered random value \( \xi \) with couples of any order, with the mathematical expectation \( E\xi = 0 \) and the dispersion \( s^2 \). Break down the \( \varphi_\xi(t) \) onto the Maclaurin power series based on the property 4 limited by quartic components:

\[
\varphi_\xi(t) \approx \varphi_\xi(0) + \frac{\varphi_\xi'(0)}{2!} t^2 + \frac{\varphi_\xi''(0)}{3!} t^3 + \frac{\varphi_\xi'''(0)}{4!} t^4 = 1 + iE\xi t - \frac{1}{2!} |E\xi^2| t^2 - \frac{1}{3!} |E\xi^3| t^3 + \frac{1}{4!} |E\xi^4| t^4.
\]

Because of \( E\xi = 0, \ E\xi^3 = 0 \), then

\[
\varphi_{\xi}(t) \approx 1 - \frac{1}{2!} \left( \frac{E\xi^2}{2} \right) t^2 + \frac{1}{4!} \left( \frac{E\xi^4}{4} \right) t^4.
\]
Table 1. Characteristics of the logarithmic-normal distribution

| Characteristic          | Value          |
|-------------------------|----------------|
| Median                  | $e^a$          |
| Mode                    | $e^{a-\sigma^2}$ |
| Asymmetry factor        | $e^{-a\sigma^2/2}(e^{\sigma^2}+2\sqrt{e^{\sigma^2}-1})$ |
| Excess                  | $e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$ |
| Characteristics function| $1 + \sum_{n=1}^{\infty} \frac{(it)^n}{n!} e^{n\sigma^2 + n^2\sigma^4 / 2}$ |

Estimating the fourth couple of the random value $\xi$ through the second one based on the Gaussian approximation as amended of the remainder of series in the Cauchy form, we have:

$$\phi_\eta(t) \approx 1 - \frac{1}{2} \sigma^2 t^2 + \frac{1}{16} \sigma^4 t^4.$$  

Suppose that the random value $\eta = \xi + d$ has the mathematical expectation $E\eta = \bar{d}$ and the dispersion $\sigma^2$, the distribution law $\eta$ corresponds to the distribution law of the remainder sludge in the cooling lubricant after passing through the purification system [5]. Based on the property 2 for characteristics functions using the Euler formula, we have:

$$\phi_\eta(t) \approx \left(1 - \frac{1}{2} \sigma^2 t^2 + \frac{1}{16} \sigma^4 t^4\right) e^{i\bar{d}t} = \left(1 - \frac{1}{2} \sigma^2 t^2 + \frac{1}{16} \sigma^4 t^4\right) \left[\cos(\bar{d}t) + isin(\bar{d}t)\right]$$  \hspace{1cm} (1)

That is why the real part and the imaginary part of the characteristic function of the distribution law of the remainder sludge are written as:

$$Re\phi_\eta(t) \approx \left(1 - \frac{1}{2} \sigma^2 t^2 + \frac{1}{16} \sigma^4 t^4\right) \cos(\bar{d}t),$$  \hspace{1cm} (2)

$$Im\phi_\eta(t) \approx \left(1 - \frac{1}{2} \sigma^2 t^2 + \frac{1}{16} \sigma^4 t^4\right) \sin(\bar{d}t).$$  \hspace{1cm} (3)

To obtain the high tide in formulas (1) - (3) it is necessary to specify the variation range of the parameter $t$. It is obvious that the expression $\bar{d}t$ shall belong to the length interval $2\pi$.

That is why further we will consider that $t \in \left[-\frac{\pi}{\bar{d}}, \frac{\pi}{\bar{d}}\right].$

Figure 1. Hodograph of the complex-valued characteristic function $\phi_\eta(t)$ of the distribution of the remainder sludge using the formula (1): $E\eta = \bar{d} = 10$ and the mean square deviation is $\sigma = 4$ micron,
Figure 2. Graph $\text{Im}\phi_\eta(t)$ of the distribution of the remainder sludge using the formula (3): \( E\eta = \overline{d} = 10 \) and the mean square deviation is \( \sigma = 4 \) micron, \( t \in \left[ -\frac{\pi}{10}, \frac{\pi}{10} \right] \).

Please note that in the formula (1) the characteristic function $\phi_\eta(t)$ of the distribution has the wavelike nature and interrelated real part and imaginary part. That is why these variations in the certain sense are similar to the distribution of electromagnetic waves in the space when the variable magnetic field generates the alternating electric field and vice versa.

Figure 3. Graph $\text{Re}\phi_\eta(t)$ of the distribution of the remainder sludge using the formula (2): and the mean square deviation is \( \sigma = 4 \) micron, \( t \in \left[ -\frac{\pi}{10}, \frac{\pi}{10} \right] \).

Assume that the original distribution of the sludge in the cooling lubricant has the mathematical expectation $\overline{d}_\eta$ and the dispersion $\sigma^2_\eta$. Particles remaining in the cooling lubricant after passing through the purification system have the mathematical expectation $\overline{d}_\delta$ and the dispersion $\sigma^2_\delta$ at average purification rate $\overline{\varepsilon}$. Based on the results of the monography [6] we have:

$$\sigma^2_\delta = \overline{d}_\eta \left( 1 - \overline{\varepsilon} \right)^{\sigma^2_\eta}, \quad \sigma_\delta = \sigma_\eta \left( 1 - \overline{\varepsilon} \right)^{\sigma^2_\eta}.$$

(4)

Then, having combined formulas (1) and (4), we have the final expression for calculating the characteristic function of the distribution of remaining particles in the cooling lubricant:

$$\phi_\delta(t) = \left( 1 - \frac{1}{2} \left( \sigma^2_\eta \left( 1 - \overline{\varepsilon} \right)^{\sigma^2_\eta} \right) t^2 + \frac{1}{16} \left( \sigma^2_\eta \left( 1 - \overline{\varepsilon} \right)^{\sigma^2_\eta} \right)^2 t^4 \right) \cos \left( \overline{d}_\delta \left( 1 - \overline{\varepsilon} \right)^{\phi_\eta(t)} \right) + i \sin \left( \overline{d}_\delta \left( 1 - \overline{\varepsilon} \right)^{\phi_\eta(t)} \right) \right).$$
Expressions for $\nu_p^{(d)}$ and $\nu_p^{(\sigma)}$ linearly or nonlinearly depend on variables $\bar{d}_i$, $\sigma_i$, $\bar{\epsilon}$.

The presented method gives an opportunity to determine both the characteristic function of the distribution and main/central distribution couples of the grain-size composition in cooling lubricants. This fact in its turn simplifies the modelling of cleansers and purification systems from the perspective of the probabilistic and statistical approach.

Here are some examples:

[1] Feller V 1984 *Introduction to the probability theory and its applications*: in 2 vol. (Moscow: World)

[2] Kendall M G, Stewart A 1966 *Distribution theory* (Moscow: Science) p 587

[3] Ryzhikov Yu I 1969 *Stock management* (Moscow: Science) p 344

[4] Bulyzhev E M and Khudobin L V 2004 *Resource-saving application of lubricant fluids when metal working* (Moscow: Machine engineering) p 352

[5] Bulyzhev E M and Bogdanov A Yu 2008 *Mathematical model of the fine-zoned gravitational purification process*, Guide. Engineering magazine, v. 10, Annex, pp. 10-14

[6] Bulyzhev E M 2000 *Modelling of the fine surface finishing process in terms of the «algebra of straight lines». Mathematical modelling in scientific studies* (Stavropol: publishing house SSU) pp. 175-178