Departmental Technical Reports (CS)   Computer Science

10-2020

How to Describe Measurement Errors: A Natural Generalization of the Central Limit Theorem Beyond Normal (and Other Infinitely Divisible) Distributions

Julio Urenda
University of Texas at El Paso, jcurenda@utep.edu

Olga Kosheleva
University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
University of Texas at El Paso, vladik@utep.edu

Comments:

Technical Report: UTEP-CS-20-100a

Published in: Franco Pavese, Alistair B. Forbes, Nien Fan Zhang, and Anna G. Chunovkina (eds.), *Advanced Mathematical and Computational Tools in Metrology and Testing XII*, World Scientific, Singapore, 2021, pp. 418-428.

1

How to Describe Measurement Errors: A Natural Generalization of the Central Limit Theorem Beyond Normal (and Other Infinitely Divisible) Distributions

J. C. Urenda

Departments of Computer Science and Mathematical Sciences
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
E-mail: jcurenda@utep.edu

O. Kosheleva

Department of Teacher Education
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
E-mail: olgak@utep.edu

V. Kreinovich

Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
E-mail: vladik@utep.edu

When precise measurement instruments are designed, designers try their best to decrease the effect of the main factors leading to measurement errors. As a result of this decrease, the remaining measurement error is the joint result of a large number of relatively small independent error components. According to the Central Limit Theorem, under reasonable conditions, when the number of components increases, the resulting distribution tends to Gaussian (normal). Thus, in practice, when the number of components is large, the distribution is close to normal – and normal distributions are indeed ubiquitous in measurements. However, in some practical situations, the distribution is different from Gaussian. How can we describe such distributions? In general, the more parameters we use, the more accurately we can describe a distribution. The class of Gaussian distributions is 2-dimensional, in the sense that each distribution from this family can be uniquely determined by 2 parameters: e.g., mean and standard deviations. Thus, when the approximation of the measurement error by a normal distribution is insufficiently accurate, a natural idea is to consider families with more parameters. What are 3-, 4-, 5-, n-dimensional

2

limit families of this type? Researchers have considered 3-dimensional classes of distributions, which can – under weaker assumptions – be used to describe similar limit cases; distributions from these families are known as infinitely divisible ones. A natural next question is to describe all possible n-dimensional families for all n. Such a description is provided in this paper.

Keywords: measurement error, Central Limit Theorem, infinitely divisible distributions

1. Central Limit Theorem and Distributions of Measurement Error: A Brief Reminder and Formulation of the Problem

Specifics of precise measuring instruments: main idea. Measurements are never absolutely accurate: the measurement result $\widetilde{x}$ is, in general, different from the actual (unknown) value x of the measured quantity; see, e.g., Novitskii and Zograph[1], Orlov[2], and Rabinovich[3].

For most measurement instruments, we can usually identify several main factors that contribute to this measurement error. For many instruments, these factors include thermal noise, interference of nearby electric lines, etc. To increase the measurement accuracy, we need to decrease the effect of these factors. For example:

- to decrease the effect of thermal noise, we need to cool down the measuring instrument;
- to decrease the effect of electric lines, we can place the instrument in a conducting box, etc.

As a result of this thorough decrease, all *major* factor affecting measurement uncertainty have been decreased. Thus, the remaining measurement error is the joint effect of many *small* independent error components.

This idea can help describe the probability distribution of measurement errors. Interestingly, the above seemingly *qualitative* idea can help us describe, in *quantitative* terms, the probability distribution of the corresponding measurement errors.

Specifically, in many cases, there are theorems – they are called *limit theorems* – that state that when the number of components increases, the distribution of the sum of that many independent components tends to distributions from a certain family. Thus, in practice, when the number of components is large, the actual distribution of the measurement error is close to the corresponding limit distribution.

Central Limit Theorem. Historically the first – and most well-known – limit theorem is the Central Limit Theorem that states that under some reasonable conditions, the sum of many similar-size independent random variables tends to Gaussian (normal) distribution; see, e.g., Sheskin[4]. So, under these conditions, the distribution of the resulting measurement error is close to Gaussian; see, e.g., Rabinovich[3].

For many measuring instruments, the distribution of the measurement error is indeed close to Gaussian.

Need for limit theorems beyond Central Limit Theorem. For some measuring instruments, however, the distribution of the measurement error is different from Gaussian – and even when it is close to Gaussian, it is not exactly equal to Gaussian. To describe such distributions, we need to go beyond normal distributions.

How can we do that? The more parameters we use – i.e., in other words, the higher the dimension of the corresponding family – the more accurately we can describe the corresponding distributions. A general Gaussian distribution can be described by two parameters: mean μ and standard deviation. A natural idea is thus to consider more general – e.g., 3-parametric – families of distributions.

First natural property of the class of limit distributions: closeness under addition. Suppose that we have two different families of independent small random variables:

- the first one tends to a random variable X, and
- the second one tends to the random variable Y.

Then, when we combine variables from both families, the resulting limit random variable is simply equal to the sum $X + Y$ of the two limit random variables. So, the probability distribution corresponding to this sum can also appear in the limit.

Thus, the desired family of limit probability distributions must satisfy the following property:

- if X and Y are independent random variables from this family,
- then their sum $X + Y$ also belongs to this same family.

This property can be easily described in terms of the distribution's *characteristic function* $\chi_X(\omega) \stackrel{\text{def}}{=} E[\exp(\mathrm{i} \cdot \omega \cdot X)]$, where X is the corresponding random variable, $\mathrm{i} \stackrel{\text{def}}{=} \sqrt{-1}$, and $E[Z]$ denotes the expected value of the

4

random variable Z. Indeed, here

$$\exp(i \cdot \omega \cdot (X + Y)) = \exp(i \cdot \omega \cdot X) \cdot \exp(i \cdot \omega \cdot Y), \tag{1}$$

thus the expected values of both sides are also equal:

$$E[\exp(i \cdot \omega \cdot (X + Y))] = E[\exp(i \cdot \omega \cdot X) \cdot \exp(i \cdot \omega \cdot Y)]. \tag{2}$$

Since the variables X and Y are independent, the variables $\exp(i \cdot \omega \cdot X)$ and $\exp(i \cdot \omega \cdot Y)$ are also independent. Hence, the expected value of their product is equal to the product of their expected values, so

$$E[\exp(i \cdot \omega \cdot (X + Y))] = E[\exp(i \cdot \omega \cdot X)] \cdot E[\exp(i \cdot \omega \cdot Y)], \tag{3}$$

i.e., indeed

$$\chi_{X+Y}(\omega) = \chi_X(\omega) \cdot \chi_Y(\omega). \tag{4}$$

Thus, in terms of the characteristic function, the above property takes a very simple form: the family of the characteristic functions must be closed under multiplication.

Examples. This property is definitely true for the characteristic functions of the normal distribution $\chi(\omega) = \exp(i \cdot \mu \cdot \omega - \sigma^2 \cdot \omega^2 / 2)$. Indeed:

- if we have two independent normally distributed random variables X_1 and X_2 with means μ_i and variances $V_i = \sigma_i^2$,
- then, as one can easily check, the product of their characteristic functions

$$\chi_1(\omega) = \exp(i \cdot \mu_1 \cdot \omega - V_2 \cdot \omega^2 / 2) \tag{5}$$

and

$$\chi_2(\omega) = \exp(i \cdot \mu_2 \cdot \omega - V_2 \cdot \omega^2 / 2) \tag{6}$$

also had the same form

$$\chi(\omega) = \exp(i \cdot \mu \cdot \omega - V \cdot \omega^2 / 2), \tag{7}$$

with $\mu = \mu_1 + \mu_2$ and $V = V_1 + V_2$.

In general, this property is similarly satisfied for families of the type

$$\chi(\omega) = \exp(C_1 \cdot f_1(\omega) + \ldots + C_n \cdot f_n(\omega)), \tag{8}$$

where the functions $f_i(\omega)$ are fixed, but the parameters C_i can take any value. Indeed:

5

- if we have two distributions X_1 and X_2 from such a family, with characteristic functions

$$\chi_1(\omega) = \exp(C_{11} \cdot f_1(\omega) + \ldots + C_{1n} \cdot f_n(\omega)) \tag{9}$$

and

$$\chi_2(\omega) = \exp(C_{21} \cdot f_1(\omega) + \ldots + C_{2n} \cdot f_n(\omega)), \tag{10}$$

- then the characteristic sum of the sum $X_1 + X_2$ has a similar characteristic function

$$\chi(\omega) = \exp(C_1 \cdot f_1(\omega) + \ldots + C_n \cdot f_n(\omega)), \tag{11}$$

with $C_i = C_{1i} + C_{2i}$.

Comment. It is sufficient to describe the values of the characteristic function only for $\omega > 0$. Indeed, for $\omega < 0$, we have

$$\exp(\mathrm{i} \cdot \omega \cdot X) = \exp(-\mathrm{i} \cdot |\omega| \cdot X) = [\exp(\mathrm{i} \cdot |\omega| \cdot X)]^*, \tag{12}$$

where for each complex number $z = a + \mathrm{i} \cdot b$, the notation $z^* \stackrel{\mathrm{def}}{=} a - \mathrm{i} \cdot b$ denotes its complex conjugate. By taking the expected value of both sides, we conclude that for $\omega < 0$, we have

$$\chi(\omega) = [\chi(|\omega|)]^*. \tag{13}$$

Thus, it is indeed sufficient to consider the values of the characteristic function only for $\omega > 0$.

Towards other natural properties. Other properties of the limit family are related to the fact that numerical values of a physical quantity depend:

- on the choice of a measuring unit and,
- (for many quantities like time or temperature) on the selection of the starting point.

Scale-invariance.

- If we replace the original measuring unit by a new unit which is λ times smaller,
- then all the numerical values of this quantity are multiplied by λ.

Comment. This idea is well-known in metrology: it is, e.g., an important feature to the so-called *interval scales*; see, e.g., Rabinovich[3].

6

Scale-invariance (cont-d). In particular, if we replace the original measuring unit by a new unit which is λ times smaller, then, instead of the original random variable X, we get the random variable $X' = \lambda \cdot X$. In the new units, the characteristic function has the form

$$\chi_\lambda(\omega) = E[\exp(i \cdot \omega \cdot X')] = E[\exp(i \cdot \omega \cdot (\lambda \cdot X))] =$$

$$E[\exp(i \cdot (\omega \cdot \lambda) \cdot X)] = \chi(\lambda \cdot \omega). \qquad (14)$$

We are interested in the universally applicable limit family of distributions, a family that should not depend on the choice of the measuring unit. Thus, we should require that:

- if a function $\chi(\omega)$ belongs to the limit family,
- then, for every $\lambda > 0$, the function $\chi_\lambda(\omega) = \chi(\lambda \cdot \omega)$ should also belong to this family.

This property is called *scale-invariance*.

Comment. One can easily see that the family of characteristic functions corresponding to normal distributions has this property. Indeed:

- if we have a characteristic function

$$\chi(\omega) = \exp(i \cdot \mu \cdot \omega - V \cdot \omega^2/2) \qquad (15)$$

  corresponding to normal distribution,
- then, for each $\lambda > 0$, we have

$$\chi_\lambda(\omega) = \chi(\lambda \cdot \omega) = \exp(i \cdot \mu \cdot \lambda \cdot \omega - V \cdot \lambda^2 \cdot \omega^2/2) =$$

$$\exp(i \cdot \mu_\lambda \cdot \omega - V_\lambda \cdot \omega^2/2), \qquad (16)$$

  where we denoted $\mu_\lambda \stackrel{\text{def}}{=} \lambda \cdot \mu$ and $V_\lambda \stackrel{\text{def}}{=} \lambda^2 \cdot V$.

Shift-invariance.

- If we replace the original starting point with a new one which is x_0 units smaller,
- then to all numerical values, we add x_0.

In particular, instead of the original random variable X, we get the random variable $X' = X + x_0$. In the new units, the characteristic function has the form

$$\chi_{x_0}(\omega) = E[\exp(i \cdot \omega \cdot X')] = E[\exp(i \cdot \omega \cdot (X + x_0))] =$$

$$\exp(\mathrm{i} \cdot \omega \cdot x_0) \cdot E[\exp(\mathrm{i} \cdot \omega \cdot X)] = \exp(\mathrm{i} \cdot \omega \cdot x_0) \cdot \chi(\omega). \qquad (17)$$

We are interested in the universally applicable limit family of distributions, a family that should not depend on the choice of the starting point. Thus, we should require that:

- if a function $\chi(\omega)$ belongs to the limit family,
- then, for every x_0, the function $\chi_{x_0}(\omega) = \exp(\mathrm{i} \cdot \omega \cdot x_0) \cdot \chi(\omega)$ should also belong to this family.

This property is called *shift-invariance.*

Comment. One can easily see that the family of characteristic functions corresponding to normal distributions has this property. Indeed:

- if we have a characteristic function

$$\chi(\omega) = \exp(\mathrm{i} \cdot \mu \cdot \omega - V \cdot \omega^2/2) \qquad (18)$$

  corresponding to normal distribution,
- then, for each x_0, we have

$$\chi_{x_0}(\omega) = \exp(\mathrm{i} \cdot \omega \cdot x_0) \cdot \exp(\mathrm{i} \cdot \mu \cdot \omega - V \cdot \omega^2/2) =$$

$$\exp(\mathrm{i} \cdot (\mu + x_0) \cdot \omega - V \cdot \omega^2/2), \qquad (19)$$

  i.e., the same form with $\mu + x_0$ instead of μ.

What is known: infinitely divisible distributions. For each positive real number a, there is a family of distributions that satisfy all three above-described properties. For $\omega > 0$, the corresponding characteristic functions have the form $\chi(\omega) = \exp(\mathrm{i} \cdot \omega \cdot x_0 + c \cdot \omega^a)$ for a complex value $c = c_r + \mathrm{i} \cdot c_i$.

Distributions described by these formulas are known as *infinitely divisible.*

Comment. It should be mentioned that normal distributions are a particular case of this family corresponding to $a = 2$ and $c_i = 0$.

Remaining problem. Distributions from the 3-parametric family of infinitely divisible distributions – defined by the above types of characteristics functions – do not always provide a precise description of how the measurement errors are distributed; see, e.g., a detailed empitical analysis in Novitskii and Zograph[1] and in Orlov[2]. A natural idea is thus to try 4-parametric, 5-parametric, etc. families.

8

A natural question is: which n-parametric families satisfy the three above-described properties?

What we do in this paper. In this paper, we provide a full description of all the families of this type.

2. Definition and the Main Result

Analysis of the problem. We are looking for families of characteristic functions of the type

$$\chi(\omega) = \exp(C_1 \cdot f_1(\omega) + \ldots + C_n \cdot f_n(\omega)), \tag{20}$$

where functions $f_i(\omega)$ are fixed, and the parameters C_i can take any value. It is reasonable to assume that the functions $f_i(\omega)$ are smooth for $\omega > 0$.

We want to find families which are scale- and shift-invariant. So, we arrive at the following definition:

Definition 1. *By a* limit family *we mean the family F of the functions* (20) *– corresponding to some functions differentiable $f_i(\omega)$, which satisfies the following two properties:*

- *if a function $\chi(\omega)$ belongs to the family F, then, for every $\lambda > 0$, the function $\chi(\lambda \cdot \omega)$ also belongs to the family F; and*
- *if a function $\chi(\omega)$ belongs to the family F, then, for every x_0, the function $\exp(\mathrm{i} \cdot \omega \cdot x_0) \cdot \chi(\omega)$ also belongs to the family F.*

Proposition 1. *For each limit family, each function $\chi(\omega)$ from this family has the form $\exp(\ell(\omega))$, where $\ell(\omega)$ is a linear combination of the functions $(\ln(\omega))^k \cdot \omega^a$ for some non-negative integer k and some complex value a.*

Comment. For $k = 0$ and real a, we get the characteristic functions of the normal distribution and of the infinitely divisible distributions.

Proof. Instead of the characteristic functions of the type (20), it is convenient to consider their logarithms

$$\ell(\omega) = \ln(\chi(\omega)) = C_1 \cdot f_1(\omega) + \ldots + C_n \cdot f_n(\omega). \tag{21}$$

Let us denote the class of all the logarithms corresponding to all the characteristic functions from the class F by L. By S, let us denote the set of all linear combinations of functions from the set L.

In terms of these logarithms, the scale-invariance property has the similar form:

- if a function $\ell(\omega)$ belongs to the family L
- then the re-scaled function $\ell(\lambda \cdot \omega)$ should also belong to the family L.

From this, we can conclude that:

- if a function $\ell(\omega)$ belongs to the family S
- then the re-scaled function $\ell(\lambda \cdot \omega)$ should also belong to the family S.

In particular, since each function $f_i(\omega)$ belongs to the family S, the re-scaled function $f_i(\lambda \cdot \omega)$ also belongs to the family S, i.e., has the form

$$f_i(\lambda \cdot \omega) = C_{i1}(\lambda) \cdot f_1(\omega) + \ldots + C_{in}(\lambda) \cdot f_n(\omega), \tag{22}$$

for some coefficients C_{ij} which are, in general, depending on λ.

Let us fix i and fix λ and consider n different values of ω:

$$\omega_1, \ldots, \omega_j, \ldots, \omega_n. \tag{23}$$

Then, we have n linear equations for n unknowns $C_{i1}(\lambda), \ldots, C_{in}(\lambda)$:

$$f_i(\lambda \cdot \omega_1) = C_{i1}(\lambda) \cdot f_1(\omega_1) + \ldots + C_{in}(\lambda) \cdot f_n(\omega_1),$$

$$\ldots$$

$$f_i(\lambda \cdot \omega_j) = C_{i1}(\lambda) \cdot f_1(\omega_j) + \ldots + C_{in}(\lambda) \cdot f_n(\omega_j), \tag{24}$$

$$\ldots$$

$$f_i(\lambda \cdot \omega_n) = C_{i1}(\lambda) \cdot f_1(\omega_n) + \ldots + C_{in}(\lambda) \cdot f_n(\omega_n).$$

Due to Cramer's rule, the solution to the system of linear equations can be represented:

- as the ratio of two polynomials depending on the coefficients and on the right-hand sides, i.e.,
- as a differentiable function of the coefficients and of the right-hand sides.

Here:

- The coefficients do not depend on *lambda* at all and are, thus, differentiable (namely, constant-valued) functions of λ.
- The right-and sides $f_i(\lambda \cdot \omega_i)$ are also differentiable functions of λ – since the functions f_i are differentiable.

10

Thus, the values $C_{ij}(\lambda)$ are also differentiable function of λ.

So, in the equation (22), all the functions are differentiable. Thus, we can differentiate both sides with respect to λ and get the following equation:

$$\omega \cdot f_i'(\lambda \cdot \omega) = C_{i1}'(\lambda) \cdot f_1(\omega) + \ldots + C_{in}'(\lambda) \cdot f_n(\omega), \qquad (25)$$

where f_i' and C_{ij}', as usual, denotes the derivative of the corresponding function f_i or C_{ij}.

In particular, for $\lambda = 1$, we get

$$\omega \cdot f_i'(\omega) = c_{i1} \cdot f_1(\omega) + \ldots + c_{in} \cdot f_n(\omega), \qquad (26)$$

where we denoted $c_{ij} \stackrel{\text{def}}{=} C_{ij}'(1)$.

The equation (26) can be further simplified if instead of the original variable ω, we introduce:

- a new variable $w = \ln(\omega)$ for which $\omega = \exp(w)$, and
- the corresponding new functions $F_i(w) \stackrel{\text{def}}{=} f_i(\exp(w))$ for which

$$f_i(\omega) = F_i(\ln(\omega)). \qquad (27)$$

In these terms,

$$F_i'(w) = \frac{dF_i}{dw} = \frac{df_i(\exp(w))}{dw} = f_i'(\exp(w)) \cdot \exp(w) = f_i'(\omega) \cdot \omega, \qquad (28)$$

which is exactly the left-hand side of the equation (26). Thus, in terms of the new variables and new functions, the equations (26) corresponding to $i = 1, \ldots, n$ take the form

$$F_1'(w) = c_{11} \cdot F_1(w) + \ldots + c_{1j} \cdot f_j(w) + \ldots + c_{1n} \cdot F_n(w),$$

$$\ldots$$

$$F_i'(w) = c_{i1} \cdot F_1(w) + \ldots + c_{ij} \cdot f_j(w) + \ldots + c_{in} \cdot F_n(w), \qquad (29)$$

$$\ldots$$

$$F_b'(w) = c_{n1} \cdot F_1(w) + \ldots + c_{nj} \cdot f_j(w) + \ldots + c_{nn} \cdot F_n(w).$$

So, the functions $F_1(w), \ldots, F_n(w)$ satisfy a system of linear differential equations with constant coefficients. It is know that a general solution to such a system is a linear combination of the functions $w^k \cdot \exp(a \cdot w)$, where:

- a is an eigenvalue of the matrix $\|c_{ij}\|$, and

- k is a non-negative integer which is smaller than the multiplicity of the eigenvalue.

Thus, each function $f_i(\omega) = F_i(\ln(\omega))$ is a linear combination of the functions $(\ln(\omega))^k \cdot \exp(a \cdot \ln(\omega))$. Here,

$$\exp(a \cdot \ln(\omega)) = (\exp(\ln(\omega)))^a = \omega^a, \tag{30}$$

so we conclude that each function $f_i(\omega)$ is a linear combination of the expressions $(\ln(\omega))^k \cdot \omega^a$.

Since each function $\ell(\omega) = \ln(\chi(\omega))$ is a linear combination of the functions $f_i(\omega)$, it is also equal to the linear combination of the expressions $(\ln(\omega))^k \cdot \omega^a$. The proposition is thus proven.

Acknowledgments

References

1. P. V. Novitskii and I. A. Zograph, *Estimating the Measurement Errors*, (Energoatomizdat, Leningrad, 1991), in Russian.
2. A. I. Orlov, How often are the observations normal?, *Industrial Laboratory* **57**(7), 770–772 (1991).
3. S. G. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice* (Springer, New York, 2005).
4. D. J. Sheskin, *Handbook of Parametric and Non-Parametric Statistical Procedures* (Chapman & Hall/CRC, London, UK, 2011).