XXI IMEKO World Congress ”Measurement in Research and Industry”
August 30 - September 4, 2015, Prague, Czech Republic

# NUMERICAL INVERSION OF A CHARACTERISTIC FUNCTION: AN ALTERNATIVE TOOL TO FORM THE PROBABILITY DISTRIBUTION OF OUTPUT QUANTITY IN LINEAR MEASUREMENT MODELS

Viktor Witkovský [1]

[1] Institute of Measurement Science, Slovak Academy of Sciences, Bratislava, Slovakia
witkovsky@savba.sk

***Abstract -*** The exact (statistical) inference frequently require evaluation of the probability density function (PDF), the cumulative distribution function (CDF), and/or the quantile function (QF) of a random variable from its (known) characteristic function (CF), which is defined as a Fourier transform of its probability distribution function. Working with CFs provides an alternative (frequently more simple) route, than working directly with PDFs and/or CDFs. In particular, derivation of the CF of a weighted sum of independent random variable is a very simple and trivial task (given the CFs of the random variables). However, the analytical derivation of the PDF and/or CDF by using the Fourier transform is available only in special cases. Thus, in most practical situation, a numerical derivation of the PDF/CDF from the CF is an indispensable tool. In metrological applications, such approach can be used to form the probability distribution for the output quantity of a measurement model of additive, linear or generalized linear form. In this paper we shall present a brief overview of some efficient approaches for numerical inversion of the characteristic function, which are especially suitable for typical metrological applications. The suggested numerical approaches are based on the Gil-Pelaez inverse formula and on the approximation by discrete Fourier transform (DFT) and the FFT algorithm for computing PDF/CDF of (univariate) continuous random variables. We also present a sketch of the MATLAB implementation, together with several examples to illustrate its applicability.

***Keywords:*** characteristic function, probability density function, numerical inversion, Fast Fourier Transform (FFT), Gil-Pelaez inverion formula, GUM.

## 1. INTRODUCTION

In metrology, a number of measurement models used in uncertainty evaluation are, at least approximately (up to reasonable level), of the additive linear form

$$Y = c_1X_1 + \cdots + c_nX_n, \tag{1}$$

where the input quantities $X_1, \ldots, X_n$ are independent random variables with known probability distributions, $X_j \sim F_{X_j}$ for $j = 1, \ldots, n$, possibly parametrized by $\theta_j$, and $c_1, \ldots, c_n$ are known constants, and $Y$ represents the univariate output quantity (a random variable with an unknown distribution to be determined). For more details and discussion on applicability of the uncertainty evaluation methods, based on the *GUM — Guide to the expression of uncertainty in measurement* [1] and its *Supplement 1 — Propagation of distributions using a Monte Carlo method* [2], see e.g. [3, 4, 5, 6].

Here we shall discuss alternative tools to form the probability distribution of the output quantity in linear measurement model. We shall assume that the considered characteristic functions (CFs)[1] of the input and/or output quantities, the random variables $X_1, \ldots, X_n$ and $Y$, are known or can be easily derived. Then, by the Fourier inversion theorem, the PDF of the random variable $Y$ is given by

$$\mathrm{pdf}_Y(y) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ity}\,\mathrm{cf}_Y(t)\,dt, \quad \text{for } y \in \mathbf{R}. \tag{2}$$

Derivation of the CF of a weighted sum of independent random variable is a very simple and trivial task. Let $\mathrm{cf}_{X_j}(t)$ denote the characteristic function of $X_j$. The characteristic function of $Y$ defined by (1) is

$$\mathrm{cf}_Y(t) = \mathrm{cf}_{X_1}(c_1t)\cdots\mathrm{cf}_{X_n}(c_nt). \tag{3}$$

However, analytical derivation of the PDF by using the (inverse) Fourier transform (2) is available only in special cases. Thus, in most practical situation, a numerical derivation of the PDF/CDF from the CF is an indispensable tool.

In [7], Gil-Pelaez derived the alternative inversion formulae, suitable for numerical evaluation of the PDF and/or the CDF, which require integration of a real-valued functions, only. The PDF is given by

$$\mathrm{pdf}_Y(y) = \frac{1}{\pi}\int_0^{\infty} \Re\left(e^{-ity}\,\mathrm{cf}_Y(t)\right)\,dt. \tag{4}$$

If $y$ is a continuity point of the cumulative distribution function (CDF) of $Y$, defined by $\mathrm{cdf}_Y(y) = \Pr\{Y \leq y\}$, then it is given by

$$\mathrm{cdf}_Y(y) = \frac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \Im\left(\frac{e^{-ity}\,\mathrm{cf}_Y(t)}{t}\right)\,dt. \tag{5}$$

[1]The characteristic function, $\mathrm{cf}_Y(t)$, of a continuous univariate random variable $Y \sim F_Y$, with its probability density function $\mathrm{pdf}_Y(y) = F'_Y(y)$, is defined as a Fourier transform of its PDF, $\mathrm{cf}_Y(t) = \mathrm{E}\left[e^{itY}\right] = \int_{\mathbf{R}} e^{ity}\,\mathrm{pdf}_Y(y)\,dy$.

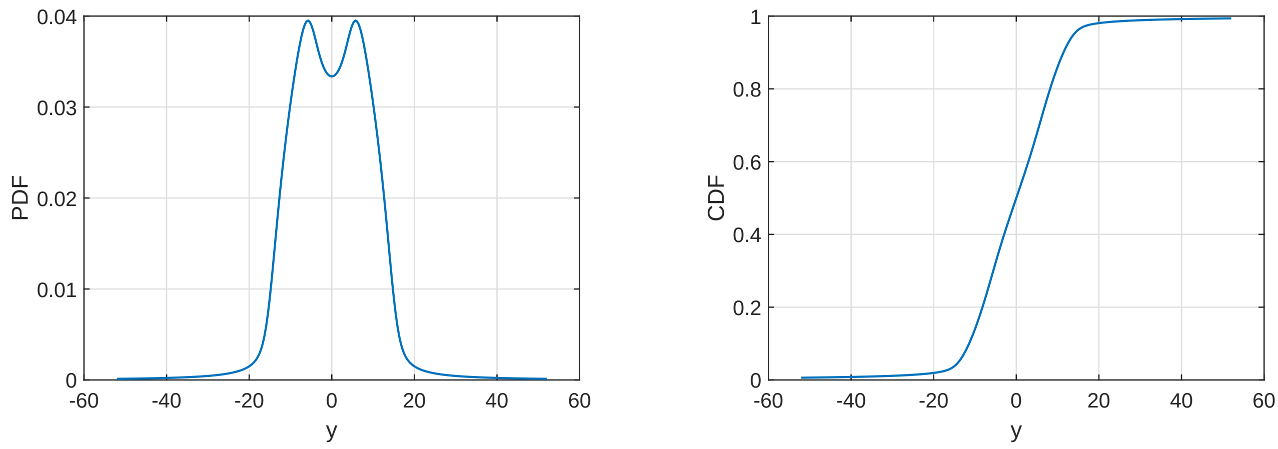


Fig. 1. The probability density function (PDF) and the cumulative distribution function (CDF) of a random variable $Y = \sum_{j=1}^{5} c_j X_j$, with $X_1 \sim N(0,1)$, $X_2 \sim t_{\nu=1}$, $X_3 \sim R(-1,1)$, $X_4 \sim T(-1,1)$, $X_5 \sim U(-1,1)$ and coefficients $c = (c_1,\ldots,c_5) = (1,1,5,1,10)$, evaluated by numerical inversion of its characteristic function by the MATLAB algorithm `tdist`, see the Example 1.

The inversion of the characteristic function based on the procedure [7] have been implemented for numerical evaluation of the distribution function of a linear combination of independent chi-squared RVs by Imhof in [8] and by Davies in [9]. The method was used also for computing the distribution of a linear combination of independent inverted gamma random variables suggested by [10] and the distribution of a linear combination of independent log-Lambert W $\times\, \chi^2_\nu$ RVs, [11].

Table 1. Characteristic functions of selected univariate symmetric (zero-mean) distributions. Here, $K_\nu(z)$ denotes the modified Bessel function of the second kind, and $J_\nu(z)$ is the Bessel function of the first kind.

| Probability distribution | Characteristic function (CF) |
|---|---|
| Gaussian $N(0,1)$ | $\mathrm{cf}(t) = \exp\left(-\frac{1}{2}t^2\right)$ |
| Student's $t_\nu$ | $\mathrm{cf}(t) = \frac{1}{2^{\frac{\nu}{2}-1}\Gamma(\frac{\nu}{2})}\left(\nu^{\frac{1}{2}}\lvert t\rvert\right)^{\frac{\nu}{2}} K_{\frac{\nu}{2}}\left(\nu^{\frac{1}{2}}\lvert t\rvert\right)$ |
| Rectangular $R(-1,1)$ | $\mathrm{cf}(t) = \frac{\sin(t)}{t}$ |
| Triangular $T(-1,1)$ | $\mathrm{cf}(t) = \frac{2-2\cos(t)}{t^2}$ |
| Arcsine $U(-1,1)$ | $\mathrm{cf}(t) = J_0(t)$ |

The Gil-Pelaez's method have been successfully implemented (by using the Gaussian quadrature for numerical integration) in the algorithm TDIST, see [12] and [13], for computing the distribution of a linear combination of independent Student's $t$ random variables and/or other symmetric (zero-mean) random variables with specific distributions. In particular, the algorithm includes the Student's $t$, normal (Gaussian), symmetric rectangular, symmetric triangular, and symmetric arcsine (U-shaped) distributions.[2] Such distributions are common and useful for uncertainty evaluation based on the GUMs uncertainty framework and its *Supplement 1*, see [2]. Table 1 presents their explicit characteristic functions, which can be numerically evaluated by standard software packages for scientific and technical computing, as e.g. MATLAB, R, SAS.

In this paper we also present a simple version of an efficient approach based on approximation by discrete Fourier transform (DFT) and by application of the FFT (fast Fourier transform) algorithm for computing PDF/CDF of (univariate) continuous random variables.

In general, numerical approximations of the continuous Fourier transform by the discrete Fourier Transform, and in particular by using the FFT algorithm, are well known and widely used, in particular, in different fields of engineering. However, FFT applications used for numerical evaluation of the PDF/CDF from the characteristic function are not so much widespread, in particular, among statisticians. An exception in this direction is the contribution by Korczynski, Cox, and Harris in [14], who suggested and illustrated the use of convolution principles, implemented using the Fast Fourier Tranform (FFT) and its inverse, to form the probability distribution for the output quantity of a measurement model of additive, linear or generalized linear form. However, from the FFT/PDF-application point of view, a more advanced field here is the field of financial mathematics and econometrics.

[2] Notice that the symmetric trapezoidal distribution is a convolution of two independent symmetric rectangular distributions.

The here considered approach for computing the PDF and/or CDF by using DFT of the characteristic function is based on the results and approaches recently suggested by Hürlimann in [15]. For an alternative approach, focused mainly on applications of the characteristic function transform methods in financial option pricing modeling and based on using the fractional fast Fourier transform (FRFT), see [16] and [17, 18, 19, 20].

Finally, we would like to mentioned also another fast growing and useful tool for numerical technical computing, the open-source package CHEBFUN, see [21], for fast numerical computing with functions. The CHEBFUN can be naturally and easily integrated for further computation, e.g. for computing CDF from PDF and QF from CDF.

## 2. BASIC RELATIONSHIPS BETWEEN THE CFT, DFT, AND FFT

We shall approximate the continuous Fourier transform (CFT), say $F(y) = \int_{-\infty}^{\infty} f(u)\, e^{-i2\pi uy}\, du$, by a discrete Fourier transform (DFT). DFT can be efficiently evaluated by using the FFT algorithm, that computes the same result as evaluating the DFT definition directly, but much faster. For complex numbers $f_0, \ldots, f_{N-1}$ the DFT is defined by the formula

$$F_k = \sum_{j=0}^{N-1} f_j e^{-i2\pi k \frac{j}{N}} \qquad k = 0, \ldots, N-1, \tag{6}$$

Formally, here we shall use notation $\mathbf{F}_N = \mathrm{FFT}(\mathbf{f}_N)$, where $\mathbf{f}_N = (f_0, \ldots, f_{N-1})$ and $\mathbf{F}_N = (F_0, \ldots, F_{N-1})$.

The relationship between the CF and the PDF is given by the inverse CFT defined by (2). For a sufficiently large interval $(-T, T)$, it is possible to approximate a PDF by means of a numerical Fourier inversion as follows

$$\mathrm{pdf}_Y(y) \approx \frac{1}{2\pi} \int_{-T}^{T} e^{-ity}\, \mathrm{cf}_Y(t)\, dt. \tag{7}$$

For simplicity, further we shall use (repeatedly) only the most simple integral approximation, based on the left-point rule (LPR), $\int_a^b f(x)\, dx \approx f(a)(b-a)$. For more sophisticated approaches see e.g. [15].

So, let $(a, b)$ is a sufficiently large interval, where the distribution of $Y$ is (mostly) concentrated. A reasonable rule for choosing $(a, b)$ can be, for example, the six-sigma rule: $(\mathrm{mean}(Y) - 6\,\mathrm{std}(Y), \mathrm{mean}(Y) + 6\,\mathrm{std}(Y))$.

Let $j, k = 0, \ldots, N-1$, $\delta_y = (b-a)/N$, and $y_k = a + k\delta_y$. For $N$ large, also $T = \pi/\delta_y$ is large, and from (7), by using $t = 2\pi u$, $dt = 2\pi\, du$, and $du = \frac{1}{b-a}$, we get

$$\mathrm{pdf}_Y(y_k) \approx \int_{-\frac{1}{2\delta_y}}^{\frac{1}{2\delta_y}} e^{-i2\pi u y_k}\, \mathrm{cf}_Y(2\pi u)\, du. \tag{8}$$

Now, we shall approximate the integral (8) by (repeatedly) using the approximate LPR, for each of the $N$ sub-intervals. Thus,

$$\mathrm{pdf}_Y(y_k) \approx \frac{1}{(b-a)} \sum_{j=0}^{N-1} e^{-i2\pi u_j y_k}\, \mathrm{cf}_Y(2\pi u_j), \tag{9}$$

where $u_j = \frac{j - \frac{N}{2}}{(b-a)}$, $j = 0, \ldots, N-1$. Hence, by using $e^{i\pi} = -1$, the expressions $u_j$ and $y_k$, and the DFT defined by (6), we finally get the formal relationship

$$\mathbf{pdf} = \mathbf{C} \odot \mathrm{FFT}(\boldsymbol{\phi}) \tag{10}$$

where $\odot$ denotes the dot product (element-wise multiplication), $\mathbf{pdf} = (\mathrm{pdf}_Y(y_0), \ldots, \mathrm{pdf}_Y(y_{N-1}))$, $\boldsymbol{\phi} = (\phi_0, \ldots, \phi_{N-1})$ with $\phi_j = (-1)^{\frac{-2a}{b-a} j}\, \mathrm{cf}_Y(\frac{2\pi}{b-a}(j - \frac{N}{2}))$, and $\mathbf{C} = (C_0, \ldots, C_{N-1})$, with $C_k = (b-a)^{-1}(-1)^{(\frac{a}{b-a} + \frac{k}{N})N}$.

For further details and improved quadrature rules on the FFT approximation of the PDF/CDF with known characteristic functions see e.g. [15, 16, 17, 18, 19, 20].

## 3. EXAMPLES

```
%% EXAMPLE 1: (MATLAB ALGORITHM TDIST)
%% http://www.mathworks.com/matlabcentral/
%% /fileexchange/4199-tdist
% PDF and CDF of a linear combination of RVs
% Y = c1*X1 + c2*X2 + c3*X3 + c4*X4 + c5*X5
% with:
% X1 ~ Normal(0,1) [we set df1=Inf] with c1=1,
% X2 ~ Student's t with  1 d.f. [set df2=1], c2=1,
% X3 ~ Rectangular on (-1,1) [set df3=-1], c3=5,
% X4 ~ Triangular on (-1,1) [set df4=-2], c4=1,
% X5 ~ U-distribution on (-1,1) [df5=-3], c5=10

df      = [Inf 1 -1 -2 -3];
const   = [1 1 5 1 10];

[pdf,y] = tdist([],df,const,'PDF');
cdf     = tdist(y,df,const,'CDF');
% plot(y,pdf)
% plot(y,cdf)
```

```
%% EXAMPLE 2:  (MATLAB ALGORITHM TDIST)
%% PDF and CF of a linear combination of RVs
% Y = c1*X1 + c2*X2
% with:
% X1 ~ Student's t with  1 d.f. [set df1=1], c1=1,
% X2 ~ Rectangular on (-1,1) [set df2=-1], c2=10,

df      = [1 -1];
const   = [1 10];

[pdf,y] = tdist([],df,const,'PDF');
% plot(y,pdf)

options.isPlot = false;
t           = linspace(-pi,pi,501);
cf          = tdist(t,df,const,'CF',options);
plot(t,cf)
```

```
%% EXAMPLE 3: (CF2PDF by FFT method)
% Characteristic function of the random variable
% Y = c1*X1 + c2*X2 + c3*X3
% with:
% c  = [1,2,3]
% X1 ~ Normal(0,1)
% X2 ~ Triangular on (-1,1)
% X3 ~ U-distribution on (-1,1)
```

```
c1  = 1; c2  = 2; c3  = 3;
cf1 = @(t) exp(-t.^2/2);
cf2 = @(t) min(1,(2-2*cos(t))./t.^2);
cf3 = @(t) besselj(0,t);
cf  = @(t) cf1(c1*t) .* cf2(c2*t) .* cf3(c3*t);

N   = 2^10;      % Number of sub-intervals
a   = -10;       % Approximate lower limit of Y
b   = 10;        % Approximate upper limit of Y
k   = 0:(N-1);   % k indices 0 : N-1
j   = 0:(N-1);   % j indices 0 : N-1

dy  = (b-a)/N;           % delta_y
y   = a + k * dy;        % y_k
u   = (j - N/2)/(b-a);   % u_j

phi = (-1).^(-(2*a/(b-a))*j) .* cf(2*pi*u);
C   = ((-1).^((a/(b-a) + k/N)*N))/(b-a);
pdf = real(C .* fft(phi)); % FFT

plot(y,pdf)
```

## 4. CONCLUSIONS

Here we suggest to consider numerical methods for derivation of the PDF/CDF from the characteristic function. Such approach can be used to form the probability distribution for the output quantity of a measurement model of additive, linear or generalized linear form, and can be considered as an alternative tool to the uncertainty evaluation based on the Monte Carlo methods. Here we have presented a brief overview of some efficient approaches for numerical inversion of the characteristic function, which are especially suitable for metrological applications. The suggested numerical approaches are based on the Gil-Pelaez inverse formula and on the approximation by discrete Fourier transform (DFT) and the FFT algorithm for computing the PDF/CDF of (univariate) continuous random variables. We have presented simple MATLAB examples in order to illustrate applicability of the suggested methods.

## ACKNOWLEDGMENTS

The work was supported by the Scientific Grant Agency VEGA of the Ministry of Education of the Slovak Republic and the Slovak Academy of Sciences, by the projects VEGA 2/0047/15 and VEGA 2/0043/13.

## REFERENCES


[1] JCGM100:2008 (GUM), Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM 1995 with minor corrections), in *JCGM - Joint Committee for Guides in Metrology*, (ISO (the International Organization for Standardization), BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP and OIML, 2008)

[2] JCGM101:2008 (GUM S1), Evaluation of measurement data – Supplement 1 to the Guide to the expression of uncertainty in measurement – Propagation of distributions using a Monte Carlo method, in *JCGM - Joint Committee for Guides in Metrology*, (ISO (the International Organization for Standardization), BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP and OIML, 2008)

[3] M. G. Cox and P. M. Harris, Validating the applicability of the GUM procedure, *Metrologia* **51**, p. S167S175 (2014).

[4] P. M. Harris and M. G. Cox, On a Monte Carlo method for measurement uncertainty evaluation and its implementation, *Metrologia* **51**, p. S176S182 (2014).

[5] R. Willink, An efficient procedure for combining type A and type B components of uncertainty, in *Advanced Mathematical and Computational Tools in Metrology and Testing X*, eds. F. Pavese and A. Chunovkina (World Scientific Publ. Co., Singapore, St. Petersburg, Russia, September 9-12, 2014, 2014) p. Submitted for publication.

[6] R. Willink, *Measurement Uncertainty and Probability* (Cambridge University Press, 2013).

[7] J. Gil-Pelaez, Note on the inversion theorem, *Biometrika* **38**, 481 (1951).

[8] J. Imhof, Computing the distribution of quadratic forms in normal variables, *Biometrika* **48**, 419 (1961).

[9] R. Davies, Algorithm AS 155: The distribution of a linear combinations of $\chi^2$ random variables, *Applied Statistics* **29**, 232 (1980).

[10] V. Witkovský, Computing the distribution of a linear combination of inverted gamma variables, *Kybernetika* **37**, 79 (2001).

[11] V. Witkovský, G. Wimmer and T. Duby, Logarithmic Lambert $W \times \mathcal{F}$ random variables for the family of chi-squared distributions and their applications, *Statistics & Probability Letters* **96**, 223 (2015).

[12] V. Witkovský, On the exact computation of the density and of the quantiles of linear combinations of t and F random variables, *Journal of Statistical Planning and Inference* **94**, 1 (2001).

[13] V. Witkovský, Matlab algorithm TDIST: The distribution of a linear combination of Student's t random variables, in *COMPSTAT 2004 Symposium*, ed. J. Antoch (Physica-Verlag/Springer, Heidelberg, Germany, 2004) pp. 1995–2002.

[14] M. J. Korczynski, M. Cox and P. Harris, Convolution and uncertainty evaluation, *Series on Advances in Mathematics for Applied Sciences* **72**, p. 188 (2006).

[15] W. Hürlimann, Improved FFT approximations of probability functions based on modified quadrature rules, *International Mathematical Forum* **8**, 829 (2013).

[16] D. H. Bailey and P. N. Swarztrauber, The fractional Fourier transform and applications, *SIAM review* **33**, 389 (1991).

[17] P. Carr and D. Madan, Option valuation using the fast Fourier transform, *Journal of Computational Finance* **2**, 61 (1999).

[18] K. Chourdakis, Option pricing using the fractional FFT, *Journal of Computational Finance* **8**, 1 (2004).

[19] M. Held, *CFH Toolbox (Characteristic Function Option Pricing)*. Matlab Central File Exchange, file 46489 (May 5 (updated Aug 12), 2014).

[20] Y. S. Kim, S. Rachev, M. L. Bianchi and F. J. Fabozzi, Computing VaR and AVaR in infinitely divisible distributions, *Probability and Mathematical Statistics* **30**, 223 (2010).

[21] T. A. Driscoll, N. Hale and L. N. Trefethen, *Chebfun Guide* (Pafnuty Publications, Oxford, 2014).