Probability Distribution Functions of the Sum of Squares of Random Variables in the Non-zero Mathematical Expectations

Yuri L. Fateev[*]
Military Engineering Institute
Siberian Federal University
Svobodny, 79, Krasnoyarsk, 660041
Russia

Vladimir V. Shaidurov[†]
Institute of Computational Modelling
Siberian Branch of the Russian Academy of Sciences
Akademgorodok, 50/44, Krasnoyarsk, 660036
Russia

Evgeny N. Garin[‡]
Dmitry D. Dmitriev[§]
Valeriy N. Tyapkin[¶]
Military Engineering Institute
Siberian Federal University
Svobodny, 79, Krasnoyarsk, 660041
Russia

The article concluded probability distribution functions sum of the squares of the random variables in the non-zero expectations. The resulting distribution function is possible to create an efficient single-step phase ambiguity resolution algorithm in determining the spatial orientation of the signals of satellite radio navigation systems. Obtained thresholds at rejecting false solutions, as well as statistical data of the algorithm.

Keywords: probability distribution functions, normal distribution, variance, standard deviation square.

Most problems in designing, research and operation of electronic equipment operate on a random variable having a normal distribution of probabilities. The value of the normal distribution is characterized by two parameters - the expectation of m and variance D. In practice, very often have to deal with the sum of normally distributed variables. In the case of the independence of these variables as a result of the addition is normally distributed, and the expectations and variances are added. Given that the variance is the square of the standard deviation, the resulting dispersion is the sum of squares of the standard variance of the initial terms.

[*]fateev_yury@inbox.ru
[†]shaidurov04@mail.ru
[‡]EGarin@sfu-kras.ru
[§]dmitriev121074@mail.ru
[¶]tyapkin58@mail.ru

In the analysis of the resulting error is usually assumed that the individual components are independent and have the expectation of zero. On this basis, the resulting error is determined as the sum of squares of the components of the error or the square root of this value. For the statistical analysis used probability distribution of the sum of squares of random variables with expectation of zero. This is a known distribution χ^2. However, in some cases, it may be that error components are the expectation, not equal to zero and a priori unknown. For example, such a situation is encountered in the resolution of phase ambiguity in the measurement of the spatial orientation of the signals of satellite navigational systems [1–3]. When iterating cycles phase is obtained only right decision and many false. In case the right decision expectation error is zero and the measurement error of spatial orientation is determined solely by the errors of measurement of phase shifts. The false solutions, an additional error a result of incorrect decision that the redundancy of systems of equations leads to additional the residuals measuring the phase shift. These residuals are constant for given values of phase ambiguity and have the character of a priori unknown mathematical expectation.

Another example is the ionosphere and troposphere delay error of measurement signals in satellite navigation systems [4, 5]. They are usually taken into account in the overall budget as a measurement error variance, ie a sum of squares, although these errors are always positive, ie, have a non-zero expectation. In these cases χ^2 distribution function can be used.

Despite the prevalence of cases with zero expectation, distribution function χ^2 with nonzero mathematical expectation is unknown.

1.   The conclusion probability distribution functions sum of the squares of the aundom variables in the non-zero mathematical expectations

So you need to get the value of the probability distribution function

$$Z = \sum_{i=1}^{n} x_i^2, \tag{1}$$

where x_i is independent random variables distributed normally with different mathematical expectations m_i and standard deviations σ_i. Consider the special case when $\sigma_1 = \sigma_2 = \ldots = \sigma_n = \sigma$.

Because the Z value is always greater than zero, then the probability density

$$p(Z < 0) = 0, \tag{2}$$

The values of x_i are independent and normally distributed

$$p(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i - m_i)^2}{2\sigma^2}}, \tag{3}$$

The joint probability density is the product of the probability densities (3) each of independent variables:

$$p(x_1, x_2, \ldots x_n) = p(x_1) \cdot p(x_2) \ldots p(x_n) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n e^{-\frac{\sum_{i=1}^{n}(x_i - m_i)^2}{2\sigma^2}}, \tag{4}$$

To get the probability density value of Z, you must first obtain the probability density function of the square of a random variable, and then, under the law of addition of random variables, can be obtained and the desired probability density.

For one, we have a random variable

$$Z_1 = x^2, \quad x = \pm\sqrt{Z_1}, \quad p_1(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad D = \frac{1}{2\sqrt{Z_1}},$$

If $Z_1 < 0$ then $p(Z_1) = 0$, if $Z_1 \geqslant 0$ then

$$p(Z_1) = |D| \cdot [p_1(Z_1) + p_1(-Z_1)] = \frac{1}{2\sigma\sqrt{2\pi Z_1}}\left[e^{-\frac{(\sqrt{Z_1}-m_1)^2}{2\sigma^2}} + e^{-\frac{(-\sqrt{Z_1}-m_1)^2}{2\sigma^2}}\right] =$$

$$= \frac{1}{2\sigma\sqrt{2\pi Z_1}}e^{-\frac{Z_1+m_1^2}{2\sigma^2}}\left[e^{\frac{\sqrt{Z_1}m_1}{\sigma^2}} + e^{-\frac{\sqrt{Z_1}m_1}{\sigma^2}}\right] = \frac{1}{\sigma\sqrt{2\pi Z_1}}e^{-\frac{Z_1+m_1^2}{2\sigma^2}}\,\mathrm{ch}\left(\frac{\sqrt{Z_1}m_1}{\sigma^2}\right) \qquad (5)$$

This result is known, and shown for example in [6]. Clearly, when $m = 0$, $p(Z_1)$ becomes χ^2 distribution with one degree of freedom.

To calculate the probability density function of the sum of squares of random variables we use characteristic functions [7]. It is known that the characteristic function of the sum of independent random variables is equal to the the characteristic functions of the terms, while the density of the probability distribution of the sum is the convolution operation using the probability density function of terms.

Distribution function of a random variable and the characteristic function associated Fourier transform:

$$\Theta(v) = \int_{-\infty}^{\infty} p(x) \cdot e^{ivZ}dZ, \qquad (6)$$

$$p(Z) = \frac{1}{\sigma\sqrt{2\pi Z}}e^{-\frac{Z}{2\sigma^2}}, \quad Z \geqslant 0,$$

$$\Theta(v) = \int_{-\infty}^{\infty} p(Z) \cdot e^{ivZ}dZ = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{Z}} \cdot e^{-\frac{Z}{2\sigma^2}} \cdot e^{ivx}dZ. \qquad (7)$$

By replacing $y = \sqrt{Z}$, obtain $dy = \frac{dZ}{2\sqrt{Z}}$, from

$$\Theta(v) = \frac{2}{\sigma\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{y^2}{2\sigma^2}+ivy^2}dy = \frac{2}{\sigma\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{y^2-2i\sigma^2vy^2}{2\sigma^2}}dy = \frac{2}{\sigma\sqrt{2\pi}}\int_0^{\infty} e^{-\frac{y^2(1-2i\sigma^2v)}{2\sigma^2}}dy. \quad (8)$$

Making the change of variable retransmission $t = \frac{y \cdot \sqrt{1-2i\sigma^2v}}{\sigma}$, obtain $dt = \frac{\sqrt{1-2i\sigma^2v}}{\sigma}dy$, from

$$\Theta(v) = \frac{2}{\sigma\sqrt{2\pi}} \cdot \frac{\sigma}{\sqrt{1-2i\sigma^2v}}\int_0^{\infty} e^{-\frac{t^2}{2}}dt = \frac{2}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{1-2i\sigma^2v}}\frac{\sqrt{2\pi}}{2} = \frac{1}{\sqrt{1-2i\sigma^2v}}. \qquad (9)$$

It is also a known expression for the characteristic function of the distribution χ^2.

We make the conclusion of the characteristic function at nonzero mathematical expectation $m \neq 0$.

$$p(Z) = \frac{1}{2\sigma\sqrt{2\pi Z}}\left[e^{-\frac{(\sqrt{Z}-m)^2}{2\sigma^2}} + e^{-\frac{(-\sqrt{Z}-m)^2}{2\sigma^2}}\right],$$

$$\Theta(v) = \int_{-\infty}^{\infty} p(Z) \cdot e^{ivZ}dZ = \frac{1}{2\sigma\sqrt{2\pi}}\int_0^{\infty}\frac{1}{\sqrt{Z}}\left[e^{-\frac{(\sqrt{Z}-m)^2}{2\sigma^2}} + e^{-\frac{(-\sqrt{Z}-m)^2}{2\sigma^2}}\right] \cdot e^{ivZ}dZ. \qquad (10)$$

We make the change of variable $y = \sqrt{Z}$, where $dy = \frac{1}{2\sqrt{Z}}dZ$. Then we have:

$$\Theta\left(v\right) = \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \left[e^{-\frac{(y-m)^2}{2\sigma^2}+ivy^2} + e^{-\frac{(y+m)^2}{2\sigma^2}+ivy^2}\right]dy =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \left[e^{-\frac{y^2-2my+m^2-2i\sigma^2vy^2}{2\sigma^2}} + e^{-\frac{y^2+2my+m^2-2i\sigma^2vy^2}{2\sigma^2}}\right]dy =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\times$$

$$\times\int_0^\infty\left[e^{-\frac{y^2\left(1-2i\sigma^2v\right)-2my+\frac{m^2}{1-2i\sigma^2v}+m^2-\frac{m^2}{1-2i\sigma^2v}}{2\sigma^2}} + e^{-\frac{y^2\left(1-2i\sigma^2v\right)+2my+\frac{m^2}{1-2i\sigma^2v}+m^2-\frac{m^2}{1-2i\sigma^2v}}{2\sigma^2}}\right]dy = \quad (11)$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty\left[e^{-\frac{\left(y\sqrt{1-2i\sigma^2v}-\frac{m}{\sqrt{1-2i\sigma^2v}}\right)^2-\frac{2i\sigma^2m^2v}{1-2i\sigma^2v}}{2\sigma^2}} + e^{-\frac{\left(y\sqrt{1-2i\sigma^2v}+\frac{m}{\sqrt{1-2i\sigma^2v}}\right)^2-\frac{2i\sigma^2m^2v}{1-2i\sigma^2v}}{2\sigma^2}}\right]dy =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}\left[\int_0^\infty e^{-\frac{\left(y\sqrt{1-2i\sigma^2v}-\frac{m}{\sqrt{1-2i\sigma^2v}}\right)^2}{2\sigma^2}}dy + \int_0^\infty e^{-\frac{\left(y\sqrt{1-2i\sigma^2v}+\frac{m}{\sqrt{1-2i\sigma^2v}}\right)^2}{2\sigma^2}}dy\right].$$

Given that $c = \frac{m}{\sqrt{1-2i\sigma^2v}}$, $z = y\sqrt{1-2i\sigma^2v}$, $dz = \sqrt{1-2i\sigma^2v}\cdot dy$, we arrive at the following expression:

$$\Theta\left(v\right) = \frac{1}{\sigma\sqrt{2\pi}}\cdot\frac{1}{\sqrt{1-2i\sigma^2v}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}\left[\int_0^\infty e^{-\frac{(z-c)^2}{2\sigma^2}}dz + \int_0^\infty e^{-\frac{(z+c)^2}{2\sigma^2}}dz\right] \quad (12)$$

and replacing $t = \frac{z\pm c}{\sigma}$, where $dt = \frac{1}{\sigma}dz$, we arrive at the following expression

$$\Theta\left(v\right) = \frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{1-2i\sigma^2v}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}\left[\int_{-c}^\infty e^{-\frac{t^2}{2}}dt + \int_c^\infty e^{-\frac{t^2}{2}}dt\right] =$$

$$= \frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{1-2i\sigma^2v}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}\left[\int_{-c}^0 e^{-\frac{t^2}{2}}dt + \int_0^\infty e^{-\frac{t^2}{2}}dt + \int_0^\infty e^{-\frac{t^2}{2}}dx - \int_0^c e^{-\frac{t^2}{2}}dt\right] = \quad (13)$$

$$= \frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{1-2i\sigma^2v}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}\cdot 2\int_0^\infty e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}}\cdot\frac{1}{\sqrt{1-2i\sigma^2v}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}\cdot\sqrt{2\pi} =$$

$$= \frac{1}{\sqrt{1-2i\sigma^2v}}\cdot e^{\frac{im^2v}{1-2i\sigma^2v}}.$$

With zero expectation $m = 0$ coincides with the formula obtained (9). The characteristic function of the sum of squares of independent normal random variables with nonzero expectation is equal to the product of the characteristic functions of the terms:

$$\Theta_n\left(v\right) = \left(1-2i\sigma^2v\right)^{-\frac{n}{2}}\cdot e^{i\frac{v\cdot\sum_{k=1}^n m_k^2}{1-2i\sigma^2v}}. \quad (14)$$

If $m_k = 0$ This formula becomes the known expression for the characteristic function distributed according to χ^2 with n degrees of freedom

$$\Theta_n\left(v\right) = \left(1-2i\sigma^2v\right)^{-\frac{n}{2}}, \quad (15)$$

From the expression (14) shows a remarkable property of the characteristic function - with equal accuracy terms ($\sigma_1 = \sigma_2 = \ldots = \sigma_n = \sigma$), it does not depend on the distribution of the expectations of random variables separately, and from the sum of their squares. The same property must have a probability density function and the sum of squares of random variables. Let the total expectation

$$m^2 = \sum_{k=1}^{n} m_k^2. \tag{16}$$

The density of the probability distribution can be obtained by performing the inverse Fourier transform of the characteristic function.

$$p_n(Z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Theta_n(v) \cdot e^{-ivZ} dv = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(1 - 2i\sigma^2 v\right)^{-\frac{n}{2}} \cdot e^{i\frac{v \cdot m^2}{1-2i\sigma^2 v}} \cdot e^{-ivZ} dv =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(1 - 2i\sigma^2 v\right)^{-\frac{n}{2}} \cdot e^{\frac{iv \cdot m^2}{1-2i\sigma^2 v} - ivZ} dv =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(1 - 2i\sigma^2 v\right)^{-\frac{n}{2}} \cdot e^{\frac{m^2\left(2i\sigma^2 v - 1 + 1\right)}{2\sigma^2\left(1-2i\sigma^2 v\right)} - \frac{\left(2i\sigma^2 v - 1 + 1\right)Z}{2\sigma^2}} dv = \tag{17}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(1 - 2i\sigma^2 v\right)^{-\frac{n}{2}} \cdot e^{-\frac{m^2}{2\sigma^2} + \frac{m^2}{2\sigma^2\left(1-2i\sigma^2 v\right)} - \frac{Z}{2\sigma^2} + \frac{\left(1-2i\sigma^2 v\right)Z}{2\sigma^2}} dv =$$

$$= \frac{1}{2\pi} e^{-\frac{Z+m^2}{2\sigma^2}} \int_{-\infty}^{\infty} \left(1 - 2i\sigma^2 v\right)^{-\frac{n}{2}} \cdot e^{\frac{m^2}{2\sigma^2}\frac{1}{1-2i\sigma^2 v} + \frac{Z}{2\sigma^2}\left(1-2i\sigma^2 v\right)} dv.$$

We make the change of variable $y = 2\sigma^2 v$, where $dy = 2\sigma^2 dv$, $v = \frac{y}{2\sigma^2}$. Then we have:

$$p_n(Z) = \frac{1}{4\pi\sigma^2} e^{-\frac{Z+m^2}{2\sigma^2}} \int_{-\infty}^{\infty} \left(1 - iy\right)^{-\frac{n}{2}} \cdot e^{\frac{m^2}{2\sigma^2}\frac{1}{1-iy} + \frac{Z}{2\sigma^2}\left(1-iy\right)} dy. \tag{18}$$

From the expression (19) shows that the probability density function depends on the ratio of the total of the expectation of a standard deviation m/σ.

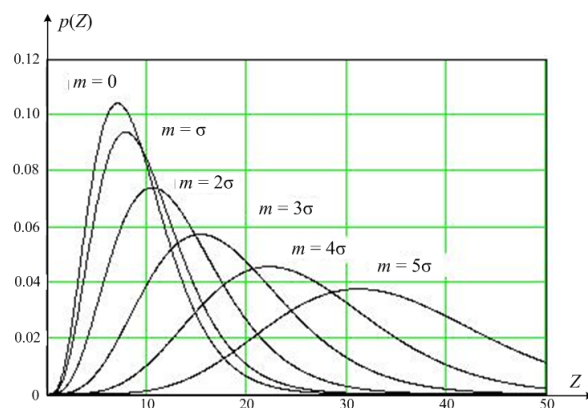Charts probability density for various values are m/σ shown in Fig. 1.

Fig. 1. The density of the probability distribution of the sum of squares of random variables for different values of m

The probability density (18) cannot be expressed in terms of elementary and tabulated functions, complicating the task of calculating the probability of finding a given value in a particular

area. For these purposes, you can get a cumulative distribution function

$$P(Z) = \frac{1}{2\pi} \int_{-\infty}^{Z} \int_{-\infty}^{\infty} \Theta(v) \cdot e^{-ivx} dv dZ = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\Theta(v)}{iv} \cdot \left(1 - e^{-ivZ}\right) dv =$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\left(1 - 2i\sigma^2 v\right)^{-\frac{n}{2}}}{iv} \cdot e^{\frac{im^2 v}{1 - 2i\sigma^2 v}} \left(1 - e^{-ivZ}\right) dv. \tag{19}$$

Conclusion

The resulting probability distribution function is possible to create an efficient single-step algorithm for phase ambiguity resolution in the measurement of the spatial orientation of the signals of satellite radio navigation systems [8–12]. Obtained thresholds at rejecting false solutions, as well as statistical data of the algorithm.

References

[1] Y.L.Fateev, D.D.Dmitriev, V.N.Tyapkin, E.N.Garin, V.V.Shaidurov, The phase ambiguity resolution in the angle-measuring navigation equipment, *AIP Conference Proceedings*, **1611**(2014), no. 12, 12–15.

[2] G.Giorgi, P.J.G.Teunissen, S.Verhagen, P.J.Buist, Improving the GNSS Attitude Ambiguity Success Rate with the Multivariate Constrained LAMBDA Method, In Geodesy for Planet Earth, *International Association of Geodesy Symposia*, **136**(2012), 941–948.

[3] P.Buist, The baseline constrained LAMBDA method for single epoch, single frequency attitude determination applications, In Proceedings of the 20th International Technical Meeting of the Satellite Division of the Institute of Navigation. Fort Worth, Tex, USA, **9**(2007), no. 3, 2962–2973.

[4] Y.L.Fateev, A.S.Kurnosov, Ionosphere parameters definition, International Siberian Conference on Control and Communications (SIBCON–2013), Proceedings, Krasnoyarsk, 2013 (doi: 10.1109/SIBCON.2013.6693622).

[5] A.Borka, G.Even-Tzur, Effect of ionospheric delay modelling on long range VRS, *Survey Review*, (2014), no. 46, 122–131.

[6] V.I.Tikhonov, Statistical radiotechnics, Moscow, Sovivetskoe Radio, 1966 (in Russian).

[7] B.R.Levin, Theoretical Foundations of Statistical Radio Engineering, Moscow, Sovivetskoe Radio, 1974 (in Russian).

[8] Y.L.Fateev, D.D.Dmitriev, V.N.Tyapkin, I.N.Ishchuk, E.G.Kabulova, The phase ambiguity resolution by the exhaustion method in a single-base interferometer, *ARPN Journal of Engineering and Applied Sciences*, **8**(2015), no. 10, 8264–8270.

[9] Yu.L.Fateev, D.D.Dmitriev, V.N.Tyapkin, I.N.Kartsan, A.E.Goncharov, Phase methods for measuring the spatial orientation of objects using satellite navigation equipment, IOP Conference Series: Materials Science and Engineering, vol. 94, 2015 (doi:10.1088/1757-899X/94/1/012022).

[10] Yu.L.Fateev, D.D.Dmitriev, V.N.Tyapkin, N.S.Kremez, V.N.Bondarev, Phase ambiguity resolution in the GLONASS/GPS navigation equipment, equipped with antenna arrays, International Siberian Conference on Control and Communications (SIBCON–2015), Proceedings, Tomsk, 2015.

[11] V.N.Ratushnyak, Yu.L.Fateev, V.N.Tyapkin, D.D.Dmitriev, E.N.Garin, E.A.Veisov, Pat. 2564523 Russian Federation C1 MPC G01S1/00. The method of the angular orientation of the object on the navigation signals spacecraft, Applicant and patentee Siberian Federal University, **10**(2015), no. 2014129573/07.

[12] Yu.L.Fateev, E.N.Garin, V.N.Tyapkin, D.D.Dmitriev, V.N.Ratushnyak, Pat. 2529649 Russian Federation C1 MPC G01S1/00. The method of the angular orientation of the object on the navigation signals spacecraft, *applicant and patentee Siberian Federal University.* **10**(2015), no. 2013132328/07.

Функция распределения плотности вероятностей суммы квадратов случайных величин при ненулевых математических ожиданиях

Юрий Л. Фатеев
Владимир В. Шайдуров
Евгений Н. Гарин
Дмитрий Д. Дмитриев
Валерий Н. Тяпкин

В статье произведен вывод функции распределения плотности вероятностей суммы квадратов случайных величин при ненулевых математических ожиданиях. Полученная функция распределения позволила создать эффективный одномоментный алгоритм разрешения фазовой неоднозначности при измерении пространственной ориентации по сигналам спутниковых радионавигационных систем. Получены пороговые значения при отбраковке ложных решений, а также статистические характеристики полученного алгоритма.

Ключевые слова: функция распределения плотности вероятностей, нормальное распределение, дисперсия, квадрат среднеквадратического отклонения.