Student's t-Distribution *versus* Zeldovich–Kompaneets Solution of Diffusion Problem

R. Wojnar

Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawiński ego 5B, 02-106 Warszawa

Student's t-distribution is compared to a solution of superdiffusion equation. This t-distribution is a continuous probability distribution that arises in the problem of estimating the mean of a normally distributed population when the sample size is small. Formally it can written in the form similar to the Gaussian distribution, in which, however, instead of usual exponential function, the so called K-exponential - a form of binomial distribution - appears. Similar binomial form has the Zeldovich–Kompaneets solution of nonlinear diffusion-like problems. A superdiffusion process, similar to a Zeldovich–Kompaneets heat conduction process, is defined by a nonlinear diffusion equation in which the diffusion coefficient takes the form $D = a(t)(1/f)^n$, where $a = a(t)$ is an external time modulation, n is a positive constant, and $f = f(x, t)$ is a solution to the nonlinear diffusion equation. It is also shown that a Zeldovich-Kompaneets solution still satisfies the superdiffusion equation if $a = a(t)$ is replaced by the mean value of a. A solution to the superdiffusion equation is given. This may be useful in description of social, financial, and biological processes. In particular, the solution possesses a fat tail character that is similar to probability distributions observed at stock markets. The limitation of the analogy with the Student distribution is also indicated.

PACS: 02.70.Rr, 05.40.Jc, 05.40.Fb, 89.65.Gh

1. Introduction

1.1. Student's t-distribution

Student's t-distribution (or the t-distribution) is a continuous probability distribution that arises in the problem of estimating the mean of a normally distributed population when the sample size is small. It is the basis of the Student's t-tests for the statistical significance of the difference between two sample means, and for confidence intervals for the difference between two population means.

It has been proposed a long time ago that t-distribution is a good candidate for describing heavy tails of stock markets distributions, [1–3].

In statistics, the t-distribution was first obtained as a posterior distribution by Helmert [4–6] and Lüroth, [7]. In the English literature, a derivation of the t-distribution was published in 1908 by William Sealy Gosset [5] while he worked at the Guinness Brewery in Dublin. Due to proprietary issues, the paper was written under the pseudonym Student. The t-test and the associated theory became well-known through the work of R.A. Fisher, who called the distribution "Student's distribution", [11]. Also Pearson's 1895 paper introduced the type IV distribution, which contains Student's t-distribution as a special case, predating Gosset's subsequent use by several years, [12].

Student's distribution arises when (as in nearly all practical statistical work) the population standard deviation is unknown and has to be estimated from the data. Student's t-distribution has the probability density function

$$f_S(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}, \tag{1}$$

where ν is known as the number of degree of freedom of the distribution and Γ is the Gamma function. The second moment or the variance is given for $\nu > 2$ by

$$\sigma^2 = \int_{-\infty}^{\infty} t^2 f_S(t)\,\mathrm{d}t = \frac{\nu}{\nu - 2}. \tag{2}$$

The fourth reduced moment or the kurtosis

$$\kappa = \left(\frac{1}{\sigma^2}\right)^2 \int_{-\infty}^{\infty} t^4 f_S(t)\,\mathrm{d}t = 3\frac{\nu - 2}{\nu - 4} \tag{3}$$

is always, for $\nu > 4$, greater then 3, the value for the case of linear diffusion, appearing when $\nu \to \infty$.

The t-distribution is symmetric and bell-shaped, like the normal distribution, but has heavier tails, meaning that it is more prone to producing values that fall far from its mean. This makes it useful for understanding the statistical behaviour of certain types of ratios of random quantities, in which variation in the denominator is amplified and may produce outlying values when the denominator of the ratio falls close to zero. The Student's t-distribution is a special case of the generalised hyperbolic distribution.

1.2. Superdiffusion

Diffusion with constant coefficient is an idealised process. A nonlinear process in which the diffusion coefficient is a negative power of probability density represents a super-diffusion.

Let j denote an one-dimensional flux of Brownian particle diffusion. According to the Fick law

$$j = -D\frac{\partial f}{\partial x}, \tag{4}$$

where $f = f(x, t)$ is the distribution function which depends on position x and time t. A nonlinear process in which the diffusion coefficient takes the form

$$D = af^{-n}, \tag{5}$$

where a and n are positive constants, can be called the

Zeldovich–Kompaneets nonlinear diffusion as it is analogous to a nonlinear heat flow analysed by those authors in 1950, [16, 17]. Such a dependence of D on f leads for $n > 0$ to a diffusion with heavy tails, that is to superdiffusion.

The superdiffusion is opposite to the case treated in [18] where a problem of diffusion without tails, it is the subdiffusion, with $D = af^n, n > 0$ was discussed, cf. Fig. 1.
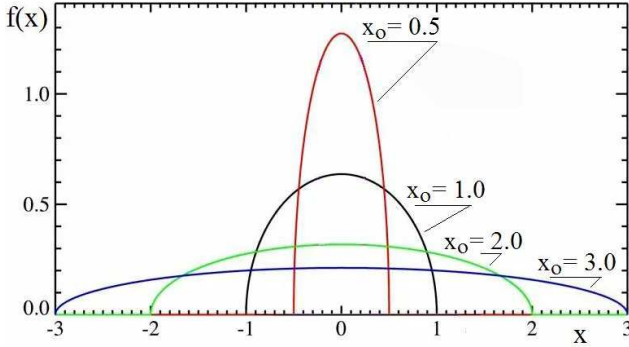

Fig. 1. Wigner semicircle distribution is an example of subdiffusion with the diffusion coefficient $D = af^2$ and $f(x,t) = \frac{2}{\pi x_0^2}\sqrt{x_0^2 - x^2}$. The value of x_0 is varying with time t, as $x_0^2 = (4/\pi)A^{1/2}$, $A = A(t) = \int_0^t a(\tau)\,d\tau$. The value of x_0 is expanding with the time t.

In the present paper a superdiffusion process is considered in which the coefficient a in Eq.(5) is a prescribed function of time $a = a(t)$. Such variation of a can denote modulation of diffusive process by external reasons with flow of the time. For example, it may reflect seasonal changes of moisture or nutrients in biology or societal changes following fashions and trends in the economy.

As always the density $f < 1$, negative power f^{-n} sustains the diffusion coefficient and results in superdiffusion, cf. Eq. (5).

In markets, super-diffusive behaviour is observed typically, [1, 13–15] in contrast to Bachelier's ideal market, for which the kurtosis is equal 3, [19].

In biology, super-diffusion can be the result of active cellular transport processes, while sub-diffusion has been used as a measure of macromolecular crowding in the cytoplasm, cf. [20–23].

2. Zeldovich-Kompaneets diffusion with time modulation

Consider one-dimensional nonlinear diffusion

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial f}{\partial x}\right), \tag{6}$$

where the coefficient D is given by (5). Thus the equation of diffusion reads

$$\frac{\partial f}{\partial t} = a\frac{\partial}{\partial x}\left(f^{-n}\frac{\partial f}{\partial x}\right). \tag{7}$$

Obviously, the time variable t in the description of the

diffusion process, Eq. (7), has nothing common with the variable t in the Student distribution, Eq. (1).

In our case, the dimension of f is $[f]=1/\text{cm}$, the dimension of $\frac{\partial}{\partial x}\left(f^{-n}\frac{\partial f}{\partial x}\right)$ is $1/\text{cm}^{n+3}$. Because the dimension of the left-hand side is $(1/\text{s})(1/\text{cm})$, the dimension $[a]=\text{cm}^{n+2}(1/\text{s})$. Introduce a new variable

$$\xi = \frac{x}{A^{\frac{1}{2-n}}}, \tag{8}$$

where

$$A = A(t) = \int_0^t a(\tau)\,d\tau \tag{9}$$

and look for the solution in the form

$$f = \frac{1}{A^{\frac{1}{2-n}}}\varphi(\xi). \tag{10}$$

As $\partial\xi/\partial t = -(a\xi)/[(2-n)A]$ and $\partial\xi/\partial x = 1/A^{\frac{1}{2-n}}$ we get subsequently $\frac{\partial f}{\partial t} = -\frac{1}{2-n}\frac{a}{A^{\frac{1}{2-n}+1}}\left(\varphi + \xi\frac{\partial\varphi}{\partial\xi}\right)$, $\frac{\partial f}{\partial x} = \frac{1}{A^{\frac{2}{2-n}}}\frac{\partial\varphi}{\partial\xi}$ and $\frac{\partial^2 f}{\partial x^2} = \frac{1}{A^{\frac{3}{2-n}}}\frac{\partial^2\varphi}{\partial\xi^2}$. Further $\frac{\partial}{\partial x}\left(\varphi^n\frac{\partial\varphi}{\partial x}\right) = \frac{1}{A^{\frac{3-n}{2-n}}}\frac{\partial}{\partial\xi}\left(\varphi^{-n}\frac{\partial\varphi}{\partial\xi}\right)$. These relations permit to reduce Eq.(7) to the ordinary differential equation

$$\frac{d}{d\xi}\left[(2-n)\varphi^{-n}\frac{d\varphi}{d\xi} + \xi\varphi\right] = 0. \tag{11}$$

This equation has a solution

$$\varphi = \varphi(\xi) = \left[\frac{n}{2(2-n)}\left(\xi_0^2 + \xi^2\right)\right]^{-1/n}, \tag{12}$$

where ξ_0 is a constant of integration. It is convenient to define

$$x_0 = A^{\frac{1}{2-n}}\xi_0 \tag{13}$$

with A given by (9). If coefficient a in (5) is a constant, then $A = at$ and $x_0 \propto t^{1/(2-n)}$.

3. Initial value problem

Let at the initial instant $t = 0$, the diffusing particle be at x=0 while $f(x,0) = 0$,

$$f(x,0) = \delta(x) \tag{14}$$

and at subsequent instants the density distribution spreads in the space according to (10) and (12). The integration constant ξ_0 in (12) is given by normalization

$$\int_{-\infty}^{\infty} f\,dx = \frac{1}{A^{\frac{1}{2-n}}}\left[\frac{n}{2(2-n)}\right]^{-1/n}$$

$$\times \int_{-\infty}^{\infty}\left(\xi_0^2 + \xi^2\right)^{-1/n}dx = 1. \tag{15}$$

Integration leads to the Euler Beta function $B(x,y)$. Namely $I \equiv \int_{-\infty}^{\infty}\left(\xi_0^2 + \xi^2\right)^{-1/n}dx = 2\xi_0^{-2/n}\int_0^{\infty}\left(1 + \frac{x^2}{x_0^2}\right)^{-1/n}dx$ where definitions (8) and (13) were used, or $I = \xi_0^{-2/n}x_0\int_0^1 t^{\frac{1}{n}-\frac{3}{2}}(1-t)^{-\frac{1}{2}}dt = \xi_0^{-\frac{2}{n}}x_0 B\left(\frac{1}{n}-\frac{1}{2},\frac{1}{2}\right)$ what inserted into (15) gives

$$\xi_0^{2-n} = 2\left(\frac{2}{n}-1\right)B^n\left(\frac{1}{n}-\frac{1}{2},\frac{1}{2}\right). \tag{16}$$

Hence expression (10) takes the form

$$f = \left\{ 2 \left(\frac{2}{n} - 1 \right) A \, B^2 \left(\frac{1}{n} - \frac{1}{2}, \frac{1}{2} \right) \right\}^{-1/(2-n)}$$

$$\times \left(1 + \frac{x^2}{x_0^2} \right)^{-1/n}. \tag{17}$$

Function B can be written in terms of Euler's Γ function $B\left(\frac{1}{n} - \frac{1}{2}, \frac{1}{2} \right) = \frac{\Gamma\left(\frac{1}{n} - \frac{1}{2} \right) \Gamma\left(\frac{1}{2} \right)}{\Gamma\left(\frac{1}{n} \right)}$. From (13) we obtain

$$\xi_0^{2-n} = 2 \left(\frac{2}{n} - 1 \right) \frac{\Gamma^n\left(\frac{1}{n} - \frac{1}{2} \right) \pi^{n/2}}{\Gamma^n\left(\frac{1}{n} \right)}, \tag{18}$$

where relation $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ was used.

The Gamma function is the meromorphic function and may be extended by analytic continuation to all complex numbers except the non-positive integers (where the function has simple poles). According to expression (17) we avoid the infinities in our density distribution f if power $n < 2$.

4. Moments of the distribution

The variance is given by

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f \, \mathrm{d}x, \tag{19}$$

where $f = f(x,t)$ is given by (17). We get

$$\sigma^2 = \frac{n}{2 - 3n} x_0^2. \tag{20}$$

Here we arrive at further limitation of the power n. If we want to keep finite value of variance σ^2, the power n should be less than $2/3$.

The substitution

$$n = \frac{2}{\nu + 1} \tag{21}$$

and $x_0^2 = \nu$, gives the value (2) of the variance for the Student's t-distribution.

The kurtosis

$$\kappa = \left(\frac{1}{\sigma^2} \right)^2 \int_{\infty}^{\infty} x^4 f \, \mathrm{d}x = 3 \frac{2 - 3n}{2 - 5n} \tag{22}$$

is greater than 3 (the value for the case of linear diffusion, $n = 0$), and less than infinity for $n < 2/5$. As could be expected, the values (20) and (22) can be obtained from the values of the variance and kurtosis for the subdiffusion given in [18], if sign of n is there changed for the opposite one.

The substitution (21) for n gives the value (3) of the kurtosis for the Student's t-distribution.

5. Linear diffusion

According to (18), for $n \to 0$ we have $\xi_0 \to 2/\sqrt{n}$. Moreover, in this limit $B\left(\frac{1}{n} - \frac{1}{2}, \frac{1}{2} \right) \to \sqrt{\pi n}$ because $\lim_{n \to 0} \Gamma\left(\frac{1}{n} - \frac{1}{2} \right) / \Gamma\left(\frac{1}{n} \right) = \sqrt{n}$, cf. [24], and the solution given by formula (10) is

$$f = \frac{1}{2} \frac{1}{\sqrt{\pi A}} \lim_{n \to 0} \left(1 - \frac{x^2}{4A} n \right)^{1/n}$$

$$= \frac{1}{2} \frac{1}{\sqrt{\pi A}} \exp\left(-\frac{x^2}{4A} \right). \tag{23}$$

For $n \to 0$ we have $a \to D$ and $A \to Dt$. Hence, we get

$$f = \frac{1}{2} \frac{1}{\sqrt{\pi D t}} \mathrm{e}^{-\frac{x^2}{4Dt}}, \tag{24}$$

which is the solution of diffusion equation for the case of constant coefficient of diffusion, it is for the linear case.

The variance for this case is given by (20) and (13), with $n \to 0$ and $\xi_0 \to 2/\sqrt{n}$. Moreover, because $a = D$ for $n = 0$

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f \, \mathrm{d}x = 2Dt \tag{25}$$

and we regain the Einstein result for the mean square displacement in brownian diffusion.

6. Results

In Figure 2 the density distributions for two different values n in nonlinear diffusion equation (7) are compared and related to the numbers of degrees of Student's distribution. It is visible that the nonlinear diffusion equation can generate distributions that take into account non-gaussian behaviour (so called *fat*, *heavy* or *thick tails*) observed on financial markets, [2].
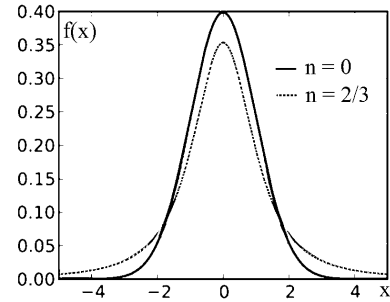
Fig. 2. Superdiffusion for two different powers n of nonlinearity in the diffusion coefficient $D = af^{-n}, n \geq 0$ for the same time. The powers of n are equal to 0 (Gauss' shape) and 2/3. The corresponding numbers of degrees of freedom of Student's distribution ν are equal to ∞ and 2, respectively. This is a consequence of relations $n = 2/(\nu + 1)$ or $\nu = (2/n) - 1$, cf. Eq. (21).

It is shown, in this paper and in [18], that Zeldovich-Kompaneets' type solution describes a full range of nonlinear diffusion processes, from sub- to super-diffusion, depending on the value n. This type of solution was completed by adding a prescribed dependence of the diffusion coefficient on time, $D = a(t)f^{-n}$, Eqs. (5) and (9).

It was shown also that Zeldovich-Kompaneets' type solution of diffusion problem is more general than Student's t-distribution, cf. Eqs. (1) and (17). These two distributions can be compared, if we, at first substitute n of Eq. (17) for $2/(\nu + 1)$ of Eq. (1), and next we limit the value x_0^2 to the value ν, cf. Fig. 2. Here the analogy with the

Student distribution finishes, as in the case of diffusion x_0 is not the parameter of the process but varies with the time, cf. Eqs. (13), (16) and (9).

Further, it was shown that nonlinear diffusion equation leads to Student's t-distribution, if power of nonlinearity is less than 2. Finite value of variance σ^2 is secured if the power n is less than $2/3$, and the finite value of super-diffusion kurtosis demands $n < 2/5$.

References

[1] P.D. Praetz, *Journal of Business* **45**, 49 (1972).

[2] R.N. Mantegma, H.E. Stanley, *An introduction to econophysics. Correlations and complexity in finance*, Cambridge University Press 2000.

[3] S. Hurst, *Financial Mathematics Research* Report No. FMRR006-95, Statistics Research Report No. SRR044-95.

[4] F.R. Helmert, *Zeitschrift für Mathematik und Physik* **20**, 300 (1875).

[5] F.R. Helmert, *Z. Math. Phys* **21**, 192 (1876).

[6] F.R. Helmert, *Astronomische Nachrichten* **88**, 113 (1876).

[7] J. Lüroth, *Astron. Nachr.* **87**, 209 (1876).

[8] J. Pfanzagl, O. Sheynin, *Biometrika* **83**, 891 (1996).

[9] O. Sheynin, *Arch. Hist. Ex. Sci.* **49**, 73 (1995).

[10] Student [William Sealy Gosset], *Biometrika* **6**, 1 (1908), doi:10.1093/biomet/6.1.1 .

[11] R.A. Fisher, *Metron* **5**, 90 (1925).

[12] K. Pearson, *Phil. Trans. R. Soc. London* **186**, 343 (1895).

[13] R. Blattberg, N. Gonedes, *Journal of Business* **47**, 244 (1972).

[14] L. Borland, *Quantitative Finances* **2**, 415 (2002).

[15] H. Kleinert, *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*, 4th ed., World Scientific, Singapore 2006.

[16] Y.B. Zeldovich, A.S. Kompaneets, in: *Collection of papers celebrating the seventieth birthday of Academician A. F. Ioffe*, ed. P.I. Lukirskii, Izdat. Akad. Nauk SSSR, Moskva 1950, pp. 61-71 (in Russian).

[17] L. D. Landau and E. M. Lifshitz, *Fluid mechanics*, Pergamon Press, Oxford 1959.

[18] R. Wojnar, *Acta Physica Polonica A* **114**, 607 (2008).

[19] L. Bachelier, *Annales Scientifiques de l'École Normale Supérieure*, 21 (1900).

[20] A. Caspi, R. Granek, M. Elbaum, *Physical Review E* **66**, 011916 (2002).

[21] M. Weiss, M. Elsner, F. Kartberg, T. Nilsson, *Biophysical Journal* **87** 3518 (2004).

[22] J.-Ph. Bouchaud, *Physics Reports* **195** 127 (1990), http://dx.doi.org/10.1016/0370-1573(90)90099-N.

[23] A. von Kameke, F. Huhn, G. Fernández-García, A.P. Munuzuri, V. Pérez-Munuzuri, *Physical Review E* **81**, 066211 (2010).

[24] I.M.Ryżyk, I.S.Gradsztejn, *Table of integrals, series and products*, PWN, Warszawa 1964 (in Polish).