Approximate Method for Computing the Sum of Independent Random Variables

Andrzej Pownuk

Department of Mathematical Sciences

The University of Texas at El Paso

http://andrzej.pownuk.com

Definition

Let X be a set. Then σ-algebra $\mathcal{F}$ is a nonempty collection of subsets of X such that the following hold:

1. X is in $\mathcal{F}$.

2. If $A \in \mathcal{F}$, then $A^{c} \in \mathcal{F}$.

3. If $A_1 \in \mathcal{F}, A_2 \in \mathcal{F}, \ldots$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition

Let $(\Omega, \mathcal{F}, P)$ **be probability space and** $(E, \mathcal{E})$. **Then an** $(E, \mathcal{E})$**-valued random variable is a function** $X : \Omega \to E$ **which is** $(\mathcal{F}, \mathcal{E})$ **measurable function.**

$$X : \Omega \ni \omega \to X(\omega) \in \mathbb{R}$$

Definition A random variable X with values in a measurable space $(\mathcal{X}, \mathcal{A})$ (usually $\mathbb{R}^n$ with the Borel sets as measurable subsets) has as probability distribution probability distribution the measure $X_* P$ on $(\mathcal{X}, \mathcal{A})$: the density of X with respect to a reference measure μ on $(\mathcal{X}, \mathcal{A})$ is the Radon–Nikodym derivative $f = \frac{dX_* P}{d\mu}$.

That is, f is any measurable function with the property that:

$P\{X \in A\} = \int_{X^{-1}(A)} dP = \int_A f \, d\mu$ for any measurable set $A \in \mathcal{A}$.

Sample PDF

Normal distribution

$$f_X\left(x\right)=\frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{\left(x-\mu\right)^2}{2\sigma^2}\right)$$

Laplace distribution

$$f\left(x;\mu,b\right)=\frac{1}{2b}\exp\left(-\frac{\left|x-\mu\right|}{b}\right)$$

Expectation of the random variable

$$E\left(X\right)=\int_{\mathbb{R}} x f_X\left(x\right)dx$$

$$E\left(g\left(X\right)\right)=\int_{\mathbb{R}} g\left(x\right)f_X\left(x\right)dx$$

$$\varphi_X\left(t\right)=E\left(\exp\left(itX\right)\right)=\int_{\mathbb{R}} \exp\left(itx\right)f_X\left(x\right)dx$$

Sum of Random Variables

$$Y = X_1 + X_2$$

$$F_{Y_{1,2}}(y) = \iint\limits_{x_1 + x_2 < y} f_{X_1, X_2}(x_1, x_2)\, dx_1\, dx_2$$

$$f_{Y_{1,2}}(y) = \frac{d}{dy} F_{Y_{1,2}}(y) = \frac{d}{dy}\left(\int_{-\infty}^{\infty} \left(\int_{-\infty}^{y - x_1} f_{X_1}(x_1)\, f_{X_2}(x_2)\, dx_2 \right) dx_1 \right)$$

$$f_{Y_{1,2}}(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1)\, f_{X_2}(y - x_1)\, dx_1$$

Convolution

$$\left(f_{X_1} * f_{X_2} \right)(x) = \int_{-\infty}^{\infty} f_{X_1}(\tau) \cdot f_{X_2}(x-\tau)\,d\tau$$

$$f_Y(y) = \left(f_{X_1} * \ldots * f_{X_k} \right)(y)$$

Characteristic Function

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

$$\varphi_Y(\omega) = \varphi_{a_1X_1 + a_2X_2 + \ldots + a_nX_n}(\omega) = \varphi_{a_1X_1}(\omega)\varphi_{a_2X_2}(\omega) \cdot \ldots \cdot \varphi_{a_nX_n}(\omega)$$

$$f_Y(y) = \mathcal{F}_\omega^{-1}\left(\varphi_{X_1}(a_1\omega)\varphi_{X_2}(a_2\omega) \cdot \ldots \cdot \varphi_{X_n}(a_n\omega)\right)(y)$$

Least Square Approximation of Characteristic Function

$$\varphi_{X_1}(\omega) \approx \tilde{\varphi}_{X_1}(\omega) = \varphi_{1,1}(\omega) \ldots \varphi_{1,n}(\omega) = \exp\left(\sum_{i=1}^{n} A_{1,i} \left|\omega\right|^{\alpha_i}\right)$$

...

$$\varphi_{X_k}(\omega) \approx \tilde{\varphi}_{X_k}(\omega) = \varphi_{k,1}(\omega) \ldots \varphi_{k,n}(\omega) = \exp\left(\sum_{i=1}^{n} A_{k,i} \left|\omega\right|^{\alpha_i}\right)$$

Least Square Approximation of Characteristic Function

$$\left(A_{11},\ldots,A_{1,n}\right)=\underset{A_{11},\ldots,A_{1,n}}{\arg\min}\frac{1}{2\pi}\left\|\varphi_{X_1}-\tilde{\varphi}_{X_1}\right\|_2^2$$

…

$$\left(A_{k,1},\ldots,A_{k,1}\right)=\underset{A_{k,1},\ldots,A_{k,n}}{\arg\min}\frac{1}{2\pi}\left\|\varphi_{X_k}-\tilde{\varphi}_{X_k}\right\|_2^2$$

Least Square Approximation of Characteristic Function

$$\left(A_{11},\ldots,A_{1,n}\right)=\underset{A_{11},\ldots,A_{1,n}}{\arg\min}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\varphi_{X_1}(\omega)-\tilde{\varphi}_{X_1}(\omega)\right)^2 d\omega$$

$$\ldots$$

$$\left(A_{k,1},\ldots,A_{k,n}\right)=\underset{A_{k,1},\ldots,A_{k,n}}{\arg\min}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\varphi_{X_k}(\omega)-\tilde{\varphi}_{X_k}(\omega)\right)^2 d\omega$$

$$\varphi_Y(\omega) \approx \tilde{\varphi}_Y(\omega) = \tilde{\varphi}_{X_1}(\omega)\ldots\tilde{\varphi}_{X_k}(\omega)$$

$$\tilde{\varphi}_Y(\omega) = \tilde{\varphi}_{X_1}(\omega)\cdot\ldots\cdot\tilde{\varphi}_{X_k}(\omega) = \exp\left(\sum_{j=1}^{k}\sum_{i=1}^{n} A_{j,i}\,|\omega|^{\alpha_i}\right)$$

$$\tilde{\varphi}_Y(\omega) = \exp\left(\sum_{i=1}^{n}\sum_{j=1}^{k} A_{j,i}\,|\omega|^{\alpha_i}\right) = \exp\left(\sum_{i=1}^{n} B_i\,|\omega|^{\alpha_i}\right)$$

$$B_i = \sum_{j=1}^{k} A_{j,i}$$

$$f_Y(y) = \mathcal{F}_{\omega}^{-1}\left(\varphi_Y(\omega)\right)(y)$$

$$\mathcal{F}_{\omega}^{-1}\left(\varphi_Y(\omega)\right)(y) \approx \mathcal{F}_{\omega}^{-1}\left(\tilde{\varphi}_Y(\omega)\right)(y)$$

$$\tilde{f}_Y(y) = \mathcal{F}_{\omega}^{-1}\left(\tilde{\varphi}_Y(\omega)\right)(y)$$

$$\tilde{f}_Y(y) = \mathcal{F}_{\omega}^{-1}\left(\exp\left(\sum_{i=1}^{n} B_i \left|\omega\right|^{\alpha_i}\right)\right)(y)$$

According to the Parseval theorem for the L_2 norm the error can be calculated in the following way

$$error = \left\| f_Y - \tilde{f}_Y \right\|^2 = \int_{-\infty}^{\infty} \left(f_Y(y) - \tilde{f}_Y(y) \right)^2 dy =$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\varphi_Y(\omega) - \tilde{\varphi}_Y(\omega) \right)^2 d\omega$$

Error of the Approximation

$$error = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\varphi_{Y}(\omega) - \exp\left(\sum_{i=1}^{n} B_{i} \left|\omega\right|^{\alpha_{i}}\right)\right)^{2} d\omega$$

Least Square Method

$$E = \left\| f - \varphi \right\|^2 = \left\| f - \sum_i a_i p_i \right\|^2 =$$

$$= \int_a^b \left(f - \varphi \right)^2 dx = \int_a^b \left(f - \sum_i a_i p_i \right)^2 dx$$

$$\begin{bmatrix} \left(p_1, p_1\right) & \dots & \left(p_1, p_n\right) \\ \dots & \dots & \dots \\ \left(p_n, p_1\right) & \dots & \left(p_n, p_n\right) \end{bmatrix} \begin{bmatrix} a_1 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \left(f, p_1\right) \\ \dots \\ \left(f, p_n\right) \end{bmatrix}$$

Approximation
of the Characteristic Function

$$\varphi(\omega) \approx \exp\left(a_1 |\omega|^{\alpha_1} + a_2 |\omega|^{\alpha_2} + \ldots + a_n |\omega|^{\alpha_n}\right)$$

then

$$\ln \varphi(\omega) \approx a_1 |\omega|^{\alpha_1} + a_2 |\omega|^{\alpha_2} + \ldots + a_n |\omega|^{\alpha_n}$$

Specific Form of the Dot Product

$$(p,q)=\sum_{i=1}^{n} p(\omega_i)q(\omega_i)w(\omega_i)$$

$$\begin{bmatrix} \sum_{i=1}^{m} |\omega_i|^{\alpha_1}|\omega_i|^{\alpha_1} w(\omega_i) & \dots & \sum_{i=1}^{m} |\omega_i|^{\alpha_1}|\omega_i|^{\alpha_n} w(\omega_i) \\ \dots & \dots & \dots \\ \sum_{i=1}^{m} |\omega_i|^{\alpha_n}|\omega_i|^{\alpha_1} w(\omega_i) & \dots & \sum_{i=1}^{m} |\omega_i|^{\alpha_n}|\omega_i|^{\alpha_n} w(\omega_i) \end{bmatrix} \begin{bmatrix} a_1 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} \ln\varphi(\omega_i)\,|\omega_i|^{\alpha_1} w(\omega_i) \\ \dots \\ \sum_{i=1}^{m} \ln\varphi(\omega_i)\,|\omega_i|^{\alpha_n} w(\omega_i) \end{bmatrix}$$

$$\begin{bmatrix} \sum_{i=1}^{m} |\omega_i|^{\alpha_1}|\omega_i|^{\alpha_1} \varphi(\omega_i) & \dots & \sum_{i=1}^{m} |\omega_i|^{\alpha_1}|\omega_i|^{\alpha_n} \varphi(\omega_i) \\ \dots & \dots & \dots \\ \sum_{i=1}^{m} |\omega_i|^{\alpha_n}|\omega_i|^{\alpha_1} \varphi(\omega_i) & \dots & \sum_{i=1}^{m} |\omega_i|^{\alpha_n}|\omega_i|^{\alpha_n} \varphi(\omega_i) \end{bmatrix} \begin{bmatrix} a_1 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} \ln\varphi(\omega_i)\,|\omega_i|^{\alpha_1} \varphi(\omega_i) \\ \dots \\ \sum_{i=1}^{m} \ln\varphi(\omega_i)\,|\omega_i|^{\alpha_n} \varphi(\omega_i) \end{bmatrix}$$

Error of Approximation

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{b} f(x)\,dx + \int_{b}^{\infty} f(x)\,dx =$$

$$= \int_{-1}^{0} f\left(-\frac{a}{t}\right)\frac{a}{t^2}\,dt + \int_{a}^{b} f(x)\,dx + \int_{0}^{1} f\left(\frac{b}{t}\right)\frac{b}{t^2}\,dt =$$

$$= \int_{a}^{b} f(x)\,dx +$$

$$+\int_{-1}^{1}\left(f\left(-\frac{a}{t}\right)\frac{a}{t^2}\chi_{[-1,0]}(t) + f\left(\frac{b}{t}\right)\frac{b}{t^2}\chi_{[0,1]}(t)\right)dt$$

Numerical Integration

$$\int_{-1}^{1} f(x)dx \approx w_0 f(x_0) + w_1 f(x_1) + w_2 f(x_2) + w_3 f(x_3) + w_4 f(x_4) =$$

$$= \frac{128}{225} f\left(0\right) + \frac{322 + 13\sqrt{70}}{900} f\left(-\frac{1}{3}\sqrt{5 - 2\sqrt{\frac{10}{7}}}\right) +$$

$$+ \frac{322 + 13\sqrt{70}}{900} f\left(\frac{1}{3}\sqrt{5 - 2\sqrt{\frac{10}{7}}}\right) +$$

$$+ \frac{322 - 13\sqrt{70}}{900} f\left(-\frac{1}{3}\sqrt{5 + 2\sqrt{\frac{10}{7}}}\right) +$$

$$+ \frac{322 - 13\sqrt{70}}{900} f\left(\frac{1}{3}\sqrt{5 + 2\sqrt{\frac{10}{7}}}\right)$$

Example

PDF of the Laplace distribution

$$f\left(x;\mu,b\right)=\frac{1}{2b}\exp\left(-\frac{\left|x-\mu\right|}{b}\right)$$

Characteristic function of the Laplace distribution

$$\varphi\left(\omega\right)=\frac{\exp\left(i\omega\mu\right)}{1+b^2\omega^2}$$

For $\mu = 0, b = 1$

PDF

$$f\left(x;\mu,b\right) = \frac{1}{2}\exp\left(-\left|x\right|\right)$$

Characteristic function

$$\varphi\left(t\right) = \frac{1}{1+\omega^{2}}$$

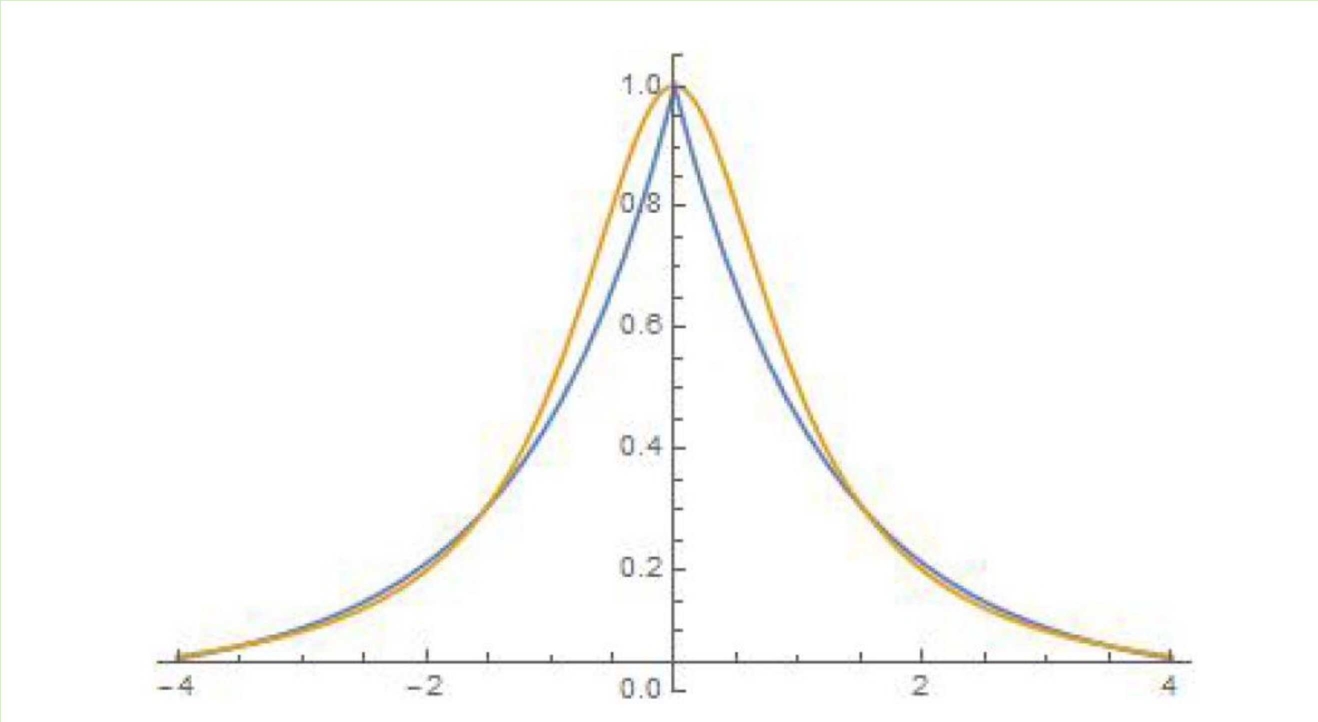
Approximation

$$\varphi(\omega) = \frac{1}{1+\omega^2}$$

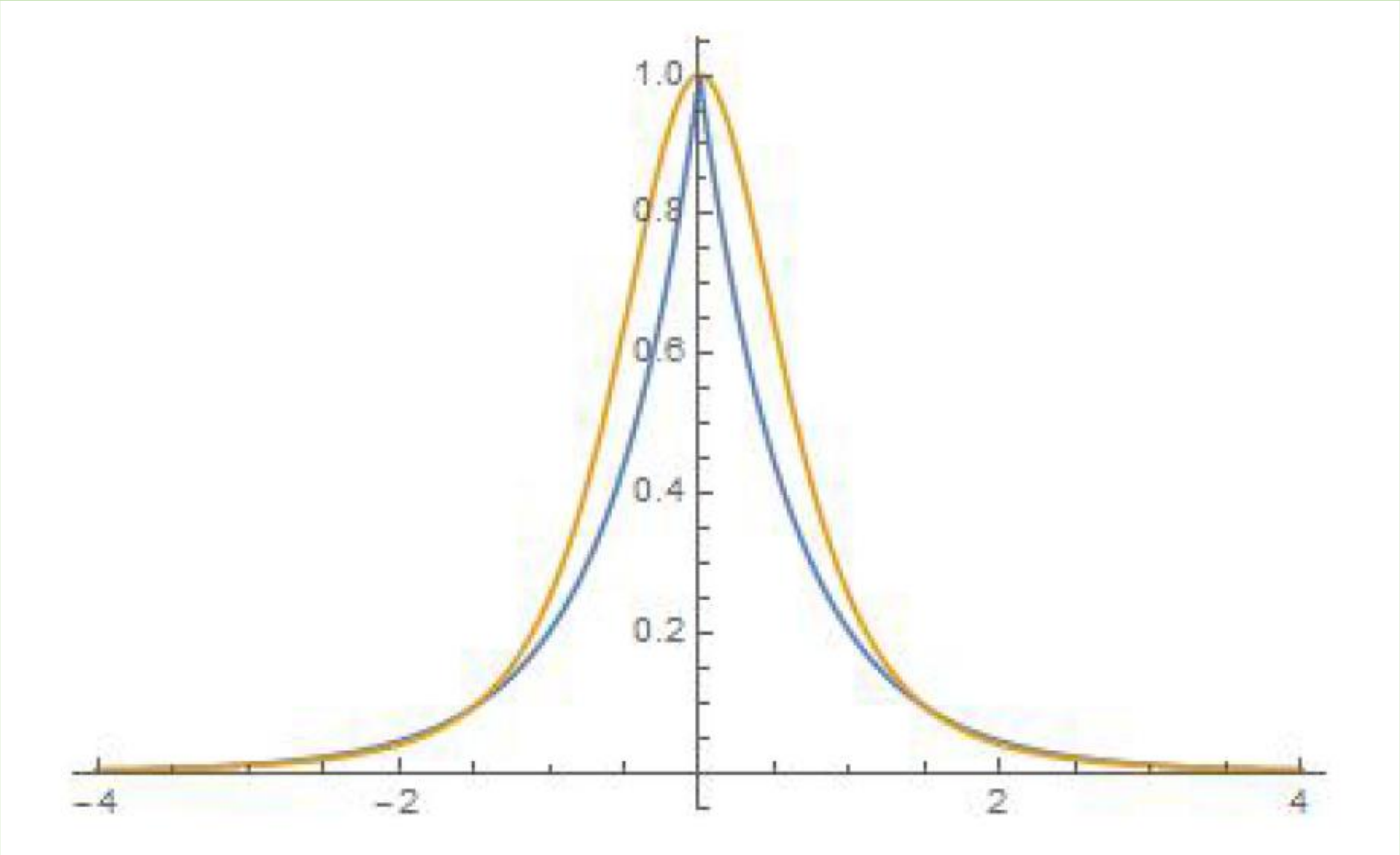
For n=2

$$\varphi_2(\omega) = \exp\left(a_1 \cdot |\omega|^{\alpha_1} + a_2 \cdot |\omega|^{\alpha_2}\right)$$

Approximation of the Characteristic Function



Approximation of PDF



Calculations were done in the framework of the adaptive computational system which use domain specific programming language and several general purpose mathematical tools (e.g. Mathematica, Matlab, Octave).

Theorem

If ϕ is a characteristic function, then ϕ is positive definite.

Bochner Representation Theorem

A continuous function $\phi:\mathbb{R}\to\mathbb{C}$ is positive definite if and only if it is the Fourier transform of a finite positive measure μ on $\mathbb{R}$, that is $\phi\left(x\right)=\int_{\mathbb{R}}\exp\left(-itx\right)d\mu\left(t\right)$.

Let us consider a function $f:\mathbb{R}\to\mathbb{R}$, then the necessary, but not sufficient condition for the function being positive semidefinite is that:

$$f\left(0\right)\geq 0$$

$$\left|f\left(x\right)\right|\leq\left|f\left(0\right)\right|$$

$$f\left(-x\right)=\overline{f\left(x\right)}\text{ for all } x\in\mathbb{R}.$$

Conclusions

In order to speed up the calculations of the sum of the independent random variables it is possible to approximate a given probability density by a sum of infinitely divisible random variables and then use the properties of the infinitely divisible distributions in order to calculate the final result.

Instead of convolution it is possible to use the product of the characteristic functions. Error estimation can be done by the use of the inverse Fourier transform and the Parseval's theorem. Convergence of the method is guaranteed by the appropriate theorems about polynomial approximation of the continuous function by a sum of fractional polynomials.