Cosinusoidal and kindred densities (pdf)

MIGUEL A. S. CASQUILHO
Technical University of Lisbon, 1049-001 Lisboa, Portugal

A simple probability density function is presented, created as a model and having in mind to make simulation easy, as well as are shown a (symmetrical) triangular and a truncated Gaussian.

Key words: *cosinusoidal probability density, triangular, truncated Gaussian.*

1. Introduction

A simple probability density function (pdf) is presented, created as a model for other user-defined functions and having in mind both to make Monte Carlo simulation easy, and the proximity to the (symmetrical) triangular and the truncated Gaussian distributions, which are shown.

2. A cosinusoidal pdf

The basic idea of this study is to take advantage of the proximity in shape of a sine or cosine to the central part of a Gaussian or truncated Gaussian distribution. The cosine was chosen, instead of the sine, just because its integral, which will be needed for the cumulative distribution, is simpler (avoiding the sign change). Several criteria of proximity may be envisaged, and the one (arbitrarily) adopted here is the selection of half a cycle of the cosine, suggested by the typical bell-shaped Gaussian curve.

Intuitively, let the pdf have the form of Eq. {1}

$$f(x) = A\cos[\omega(x - x_0)] \qquad \{1\}$$

for $x \in (\mu \pm a)$, its form being apparently difficult, but adopted just for simplicity. Indeed, it is equivalent to Eq. {2},

$$f(x) = A\cos(\omega x - \omega x_0) = A\cos(\omega x - \omega x_0) = A\cos(\omega x) - B \qquad \{2\}$$

and it is intended that such a form will be able to: be 0 for the left and right-hand extremes of x; be always non-negative; and have an integral of 1. (These properties obey those of a "legitimate" probability density function.) It is the establishment of these conditions that will completely define the function. Thus,

$$\begin{cases} f(\mu - a) = A\cos[\omega(\mu - a - x_0)] = 0 \\ f(\mu + a) = A\cos[\omega(\mu + a - x_0)] = 0 \end{cases} \qquad \{3\}$$

Noting that the cosine function is periodic, thus having infinite solutions, our (convenient) solution is

---

Prof. Dr. Casquilho is an Assistant Professor in the Department of Chemical Engineering at Instituto Superior Técnico, Technical University of Lisbon, Portugal. His email address is mcasquilho@ist.utl.pt.

MC IST File=*Cosinusoidals.doc* (Approximates the *JQT* style.)

$$\begin{cases} \cos[\omega(\mu - a - x_0)] = 0 \\ \cos[\omega(\mu + a - x_0)] = 0 \end{cases} \qquad \{4\}$$

We have used the two first properties mentioned.

$$\begin{cases} \omega(\mu - a - x_0) = -\dfrac{\pi}{2} \\ \omega(\mu + a - x_0) = \dfrac{\pi}{2} \end{cases} \qquad \{5\}$$

The unknowns are ω and x_0. Adding the two expressions in Eq. {5}, it is

$$2\omega(\mu - x_0) = 0 \qquad \{6\}$$

and (since it must obviously be $\omega \neq 0$)

$$\boxed{x_0 = \mu} \qquad \{7\}$$

as expected. Using this and subtracting the two expressions in Eq. {5}, it is

$$\omega(\mu + a - x_0) - \omega(\mu - a - x_0) = \pi \qquad \{8\}$$

$$\omega(a) - \omega(-a) = \pi \qquad \{9\}$$

$$\boxed{\omega = \frac{\pi}{2a}} \qquad \{10\}$$

Important Note always —in these derivations and anywhere else[1]— the dimensional homogeneity of these formulas. Suppose that *x* has units g (gram) and consider Eq. {1}. Then, of course, so have μ and *a*. As a matter of fact, ω would have to be g^{-1}, so that the argument of the cosine may be dimensionless. Verify that the pdf has units inverse of those of its variable, whereas the cdf (a probability) is dimensionless.

Eq. {1} then becomes

$$f(x) = A\cos\left[\frac{\pi}{2a}(x - \mu)\right] \qquad \{11\}$$

Applying now the 3.rd property mentioned, it is

$$F(x) = A\int_{\mu-a}^{x} \cos\left(\frac{\pi}{2}\frac{t-\mu}{a}\right) \mathrm{d}t = A\frac{2a}{\pi}\left[\sin\left(\frac{\pi}{2}\frac{t-\mu}{a}\right)\right]_{\mu-a}^{x} \qquad \{12\}$$

whence it is

$$\begin{aligned} F(x) &= A\frac{2a}{\pi}\sin\left(\frac{\pi}{2}\frac{x-\mu}{a}\right) - A\frac{2a}{\pi}\sin\left(\frac{\pi}{2}\frac{\mu-a-\mu}{a}\right) = \\ &= A\frac{2a}{\pi}\sin\left(\frac{\pi}{2}\frac{x-\mu}{a}\right) + A\frac{2a}{\pi} \end{aligned} \qquad \{13\}$$

or

[1] ...while I have never found this statement anywhere !

$$F(x) = A\frac{2a}{\pi}\left[1+\sin\left(\frac{\pi}{2}\frac{x-\mu}{a}\right)\right] \quad \{14\}$$

Probability goes from 0 to 1:

$$\begin{cases} F(\mu-a) = A\frac{2a}{\pi}\left[1+\sin\left(\frac{\pi}{2}\frac{\mu-a-\mu}{a}\right)\right] = 0 \\ F(\mu+a) = A\frac{2a}{\pi}\left[1+\sin\left(\frac{\pi}{2}\frac{\mu+a-\mu}{a}\right)\right] = 1 \end{cases} \quad \{15\}$$

so

$$\begin{cases} A\frac{2a}{\pi}(1-1) \equiv 0 \\ A\frac{2a}{\pi}(1+1) = 1 \end{cases} \quad \{16\}$$

The first relation becomes a confirmation. The second gives

$$A = \frac{\pi}{4a} \quad \{17\}$$

Finally, we get

$$\boxed{f(x) = \frac{\pi}{4a}\cos\left(\frac{\pi}{2}\frac{x-\mu}{a}\right)} \quad \{18\}$$

$$\boxed{F(x) = \frac{1}{2}\left[1+\sin\left(\frac{\pi}{2}\frac{x-\mu}{a}\right)\right]} \quad \{19\}$$

It is prudent to check the extremes (now only the second):

$$F(\mu+a) = \frac{1}{2}\left[1+\sin\left(\frac{\pi}{2}\frac{\mu+a-\mu}{a}\right)\right] = \frac{1}{2}(1+1) = 1 \quad \{20\}$$

The graph of the pdf (μ = 1020, a = 10) is given in Fig. 1, together with a Gaussian of the same μ.

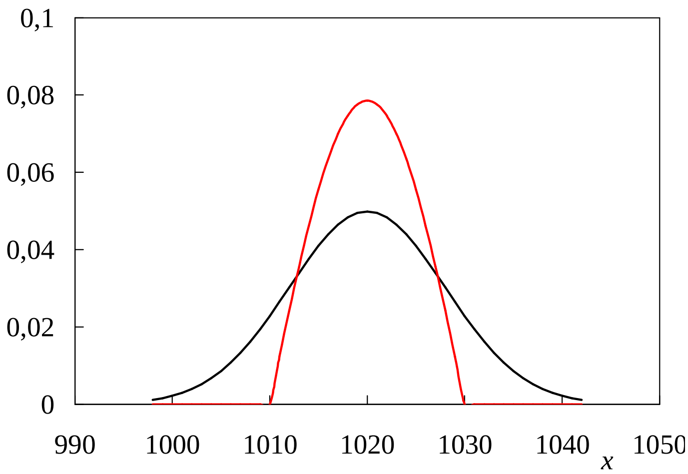

Fig. 1

3. Monte Carlo simulation

Eq. {19}, namely if convenient for simulation, can be written as

$$F(x) = \cos^2\left[\frac{\pi}{4}\left(\frac{x-\mu}{a} - 1\right)\right] \qquad \{21\}$$

In the context of simulation, inverting Eq. {19} and Eq. {21}, it is

$$\boxed{\frac{x-\mu}{a} = \frac{2}{\pi}\arcsin(2F-1)} \qquad \{22\}$$

$$\frac{x-\mu}{a} = 1 - \frac{4}{\pi}\arccos\sqrt{F} \qquad \{23\}$$

Of course, these two expressions are equivalent. For $F = 0$ (with arccos 0 = $\pi/2$), both give –1, i.e., $x = \mu - a$; for $F = 1$ (with arccos 1 = 0), both give +1, i.e., $x = \mu + a$; and, e.g., for $F = 0.9$, we get 0.59. The formula of Eq. {22} looks computationally more economical (avoiding to compute a supplementary square root), so it is preferable. The formula can be written as follows, with r, a (uniform) random number, instead of F, as implied in simulation:

$$\boxed{x = \mu + a\left[\frac{2}{\pi}\arcsin(2r-1)\right]} \qquad \{24\}$$

Illustration 3-A

A short simulation (500 values) of the cosinusoidal variable produced the (relative) frequency histogram of Fig. 2, drawn without the traditional vertical bars.

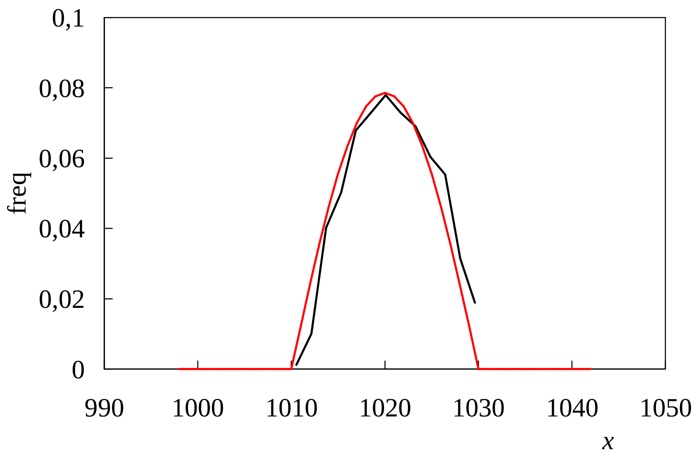

Fig. 2

Important[2] There are *two* kinds of frequency histograms. One of the few places to know them is the NIST/Sematech online handbook ([NIST/Sematech, 2010]), and unfortunately the description therein is confusing. The first type of histogram represents "the count (c_k) in a class (k) divided by the total number of observations (N)" (*sic*) ($N = \sum_{k=1}^{K} c_k$, K classes), i.e., $f_k = c_k / N$; the second type

---

[2] **Rarely** mentioned !

represents "idem, times the class width (w)", i.e., $f_k = c_k / (N\, w)$. Notice the ambiguous "times" (above or below ?). (May it serve as a mnemonic that this has the right units.) This second type permits the comparison with the pdf, to which it must tend in probability.

4. Some properties of the cosinusoidal

The calculation of the mean, μ_C, and standard deviation, σ_C, can be done from Eq. {18}, according to the definitions.

$$\mu_C = \int_X x f(x)\,\mathrm{d}\,x = \frac{\pi}{4a}\int_{\mu-a}^{\mu+a} x \cos\left(\frac{\pi}{2}\frac{x-\mu}{a}\right)\mathrm{d}\,x \qquad \{25\}$$

Making $t = (x - \mu) / a$, it is (the integral of f being 1),

$$\mu_\mathrm{C} = \mu + \frac{\pi}{4} a \int_{-1}^{+1} \underbrace{t}_{u} \underbrace{\cos\left(\frac{\pi}{2} t\right)\mathrm{d}\,t}_{v'} \qquad \{26\}$$

$$\begin{aligned} \frac{4}{\pi a}(\mu_\mathrm{C} - \mu) &= \left[t\frac{2}{\pi}\sin\left(\frac{\pi}{2}t\right)\right]_{-1}^{1} - \int_{-1}^{+1}\sin\left(\frac{\pi}{2}t\right)\mathrm{d}\,t = \\ &= \frac{2}{\pi}0 - \frac{2}{\pi}\left[\cos\left(\frac{\pi}{2}t\right)\right]_{-1}^{1} = 0 \end{aligned} \qquad \{27\}$$

or, as expected,

$$\mu_\mathrm{C} = \mu \qquad \{28\}$$

For the variance, it is

$$\sigma_\mathrm{C}^2 = \int_X (x-\mu)^2 f(x)\,\mathrm{d}\,x = \frac{\pi}{4a}\int_{\mu-a}^{\mu+a}(x-\mu)^2 \cos\left(\frac{\pi}{2}\frac{x-\mu}{a}\right)\mathrm{d}\,x \qquad \{29\}$$

$$\sigma_\mathrm{C}^2 = \frac{\pi}{4}\int_{-1}^{1} a^2 t^2 \cos\left(\frac{\pi}{2}t\right)\mathrm{d}\,t \qquad \{30\}$$

Using some previous results, it is

$$\frac{4}{\pi}\frac{\sigma_\mathrm{C}^2}{a^2} = \int_{-1}^{1} t^2 \cos\left(\frac{\pi}{2}t\right)\mathrm{d}\,t \qquad \{31\}$$

Using the Integrator ([2010]), it is

$$\frac{4}{\pi}\frac{\sigma_\mathrm{C}^2}{a^2} = \frac{2}{\pi^3}\left[\left(\pi^2 t^2 - 8\right)\sin\left(\frac{\pi}{2}t\right)\right]_{-1}^{1} + \frac{8}{\pi^2}\left[t\cos\left(\frac{\pi}{2}t\right)\right]_{-1}^{1}$$

$$\frac{4}{\pi}\frac{\sigma_\mathrm{C}^2}{a^2} = \frac{2}{\pi^3}\left[\left(\pi^2 - 8\right) + \left(\pi^2 - 8\right)\right] = \frac{4}{\pi^3}\left(\pi^2 - 8\right)$$

or

$$\frac{\sigma_C^2}{a^2} = \frac{\pi^2 - 8}{\pi^2} = 1 - \frac{8}{\pi^2} \quad \{32\}$$

$$\frac{\sigma_C}{a} = \sqrt{1 - \frac{8}{\pi^2}} \cong 0.435... \quad \{33\}$$

5. Other pdf's with kindred features

Triangular (symmetrical triangular), with $z_x = \dfrac{x-\mu}{a}$, $z_x \in (-1, +1)$, and r a random number

$$f(x) = \frac{1}{a}(1 \pm z_x) \qquad z_x \le, \ge 0 \quad \{34\}$$

$$F(x) = \pm \frac{1}{2} z_x^2 + z_x + \frac{1}{2} = \frac{1}{2}\left[1 + z_x (2 \pm z_x)\right] \quad \{35\}$$

$$\begin{cases} z_x = -1 + \sqrt{2r} & r \le 1/2 \\ z_x = 1 - \sqrt{2(1-r)} & r \ge 1/2 \end{cases} \quad \{36\}$$

$$\sigma = \frac{a}{\sqrt{6}} \cong 0.408\,a$$

Truncated Gaussian, with $x' = \dfrac{x-\mu}{\sigma}$, etc., and $\Delta\Phi = \Phi(b') - \Phi(a')$

$$f(x; \mu, \sigma, a, b) = \frac{1}{\sigma} \frac{\phi(x')}{\Delta\Phi} \quad \{37\}$$

$$F(x) = \frac{\Phi(x') - \Phi(a')}{\Delta\Phi} \quad \{38\}$$

$$x' = \Phi^{\text{inv}}\left(\Phi(a') + \underbrace{r}_{\text{var.}} (\Delta\Phi) \right) \quad \{39\}$$

In Eq. {39}, only r varies. With

$$\delta = \frac{\phi(a') - \phi(b')}{\Delta\Phi} \quad \{40\}$$

it is

$$\mu_{\text{T}} = \mu + \sigma\delta$$

$$\left(\frac{\sigma_{\text{T}}}{\sigma} \right)^2 = 1 + \frac{a'\phi(a') - b'\phi(b')}{\Delta\Phi} - \delta^2 \quad \{41\}$$

6. Conclusions

A simple pdf was presented and calculated, created as a model for other user-defined functions to show the essential steps to perform. It had in mind the ease of Monte Carlo simulation, because of the simple form (trigonometric functions), as well as the proximity to a truncated Gaussian distribution.

The (symmetrical) triangular distribution and the truncated Gaussian were also briefly presented.

Acknowledgements

This work was done at "Centro de Processos Químicos do IST" (Chemical Process Research Center of IST), in the Department of Chemical Engineering, Technical University of Lisbon. Computations were done on the central system of CIIST (Computing Center of IST).

References

– Integrator (Online Integrator), (2010), Wolfram Mathematica, http://integrals.wolfram.com/index.jsp, accessed 10.th Nov.
– NIST / SEMATECH e-Handbook of Statistical Methods (2010) (SEmiconductor MAnufacturing TECHnology), http://www.itl.nist.gov/div898/handbook/, accessed 10.th Nov., 1.3.3.14. Histogram

❖