Journal of Theoretical Physics

Founded and Edited by M. Apostol **100** (2005)

ISSN 1453-4428

On distributions with isolated points and a condensation phenomenon

M. Apostol
Department of Theoretical Physics, Institute of Atomic Physics,
Magurele-Bucharest MG-6, POBox MG-35, Romania
email: apoma@theory.nipne.ro

Abstract

A condensation of data is always associated with a singular behaviour of probability distributions in an isolated point. Finite distributions are extracted from such singular distributions, or, conversely, singular distributions are associated to finite distributions, providing the condensation phenomenon is allowed for.

Let a set of N_0 data x be distributed from 0 to ∞ by a probability distribution $p(x)$, and let N_{c0} data be "condensed" at $x = 0$, such that these condensed data are separated by a gap x_c from the rest. We further assume that N_{c0}/N_0 is finite for $N_0 \rightarrow \infty$. A histogram plots the frequency distribution

$$p(x) = \Delta N/N_0\Delta x \ \ , \tag{1}$$

where ΔN is the number of data lying in the range x to $x+\Delta x$, *i.e.* the probability distribution of having x in this range. On increasing the division sharpness, *i.e.* making Δx smaller and smaller, we get a family of functions $p(x)$, with $p(x)$ higher and higher for fixed x, up to values of Δx at $x = 0$ smaller than x_c when, since N_{c0}/N_0 is finite, $p(0)$ goes to infinite. Including, therefore, both the N- and the N_c-data, we get a distribution $p(x)$ which is singular at $x = 0$. Let us assume that the cutoff x_c is sufficiently small, and have $\Delta x = x_c$ at $x = 0$. We may write then the distribution probability as

$$\frac{dN}{N_0 dx} = p(x) \ \ , \tag{2}$$

where $\int_0^\infty dxp(x) = 1$, and

$$p_0 = p(0) = \frac{N_{c0}}{N_0 x_c} \ . \tag{3}$$

We call the N_{c0} isolated data principal data, to be distinguished from the remaining N_0, and we define the ratio

$$R = \frac{N_0}{N_{c0}} = 1/p_0 x_c \tag{4}$$

as describing the number of remainder for one principal data.

Let us further expand $p(x)$ as $p_0 - p_1 x$ for $x < x_c \ll x_0 = p_0/p_1$, where $-p_1 = p'(0) < 0$ is the first derivative of $p(x)$ at $x = 0$. Eqution (2) becomes then

$$\frac{dN}{N_0 dx} = p_0 - p_1 x = p_0(1 - x/x_0) \ . \tag{5}$$

We look now for getting a representation of $p(x)$ for small values of x, when the histogram division becomes sharper than x_c at $x = 0$. Since $\Delta x \sim x/n$ in this case, where n is an integer, equation (1) suggests that $p(x)$ might be represented as $a/x - b$, where a and b are constants to be determined. The most convenient relationship between the two approximating functions $p_0(1 - x/x_0)$ given by (5) and $a/x - b$ is obtained by requiring them to be equal at the vanishing point x_0, and have also the same derivative at that point. We get immediately $a = p_0 x_0$ and $b = p_0$, so that (5) can also be written as

$$\frac{dN}{N_0 dx} = p_0 x_0/x - p_0 \ . \tag{6}$$

In addition, in order to have a more satisfactory representation, we may require the curves described by the two approximating functions to enclose the same area. This area is $\simeq p_0 x_c$ for the function given by (5) and $p_0 x_0 \ln(x_c/c) - p_0 x_c$ for the function given by (6), where the lower-bound cutoff c is used. By equalling them, we get

$$\ln(x_c/c) = 2x_c/x_0 \ , \tag{7}$$

whose solution is

$$c = x_c e^{-2x_c/x_0} \ . \tag{8}$$

It is worth noting that the probability distribution $p(x)$ given by (6) has indeed a singularity, of the type $\sim 1/x$, for $x \rightarrow 0$, which follows from its definition (1).

We proceed now to disentangle the principal data from the remainder, by re-writing (6) as

$$\frac{d(\widetilde{N} - N_c)}{N_0 dx} = p_0 x_0/x - p_0 \tag{9}$$

and requiring

$$\frac{dN_c}{N_0 dx} = p_0 \ , \tag{10}$$

where N_c represents the principal data and $\widetilde{N}$ is another representation of the data N, including the condensed ones N_c. Obviously,

$$N_0 p_0 x_c = \widetilde{N}_0 - N_{c0} \ , \tag{11}$$

where $\widetilde{N}_0$is the total number of data $\widetilde{N}$ and N_{c0} is the total number of principal data, as the total number of data in (5) is $N_0 p_0 x_c$. It is worth noting that, according to (9) we have included the principal data in the $\widetilde{N}$-data, and subtracted at the same time the same amount of principal data, in order to separate the former from the latter. By (9) and (10) we get also

$$\frac{d\widetilde{N}}{N_0 dx} = p_0 x_0/x \ . \tag{12}$$

We normalize now the probabilities in (10) and (12) over the range 0 to x_c, and obtain the total numbers

$$N_{c0} = N_0 p_0 x_c \tag{13}$$

and

$$\widetilde{N}_0 = N_0 p_0 x_0 \ln(x_c/c) \tag{14}$$

so that (11) is verified, by (7). Making use of (8) we get the ratio R given by (4) as $R = 1/p_0 x_c = \ln(D/x_c)$, where D is an upper-bound cutoff used to normalize the distribution $p_0 x_0/x$ in the range $x_c < x < D$. We may therefore define $\widetilde{p}(x) = 1/x \ln(D/x_c)$as a distribution associated to $p(x)$, and condition $p(x_c) = \widetilde{p}(x_c)$ implies also $N_{c0} = N_0 p_0 x_c = 1$.

In conclusion, one might say that by sharpening the division of the histogram that defines a probability distribution $p(x)$ with a finite value $p_0$at $x = 0$ this probability distribution becomes singular at $x = 0$, behaving like $\sim 1/x$, providing the frequency is finite at $x = 0$ and isolated from the rest of frequencies. The $\sim 1/x$ law is also known as Omori's distribution. It is worth noting that it can also be derived by starting from (5) written as

$$\frac{1}{p_0(1 - x/x_0)} \frac{dN}{N_0 dx} = \frac{1}{p_0}(1 + x/x_0)\frac{dN}{N_0 dx} = 1 \quad , \tag{15}$$

or

$$\frac{x}{p_0 x_c}(1 + x_0/x)\frac{dN}{N_0 dx} = 1 \quad , \tag{16}$$

and defining

$$\frac{d\widetilde{N}}{N_0 dx} = (1 + x_0/x)\frac{dN}{N_0 dx} \quad , \tag{17}$$

which transform (16) into

$$\frac{d\widetilde{N}}{N_0 dx} = p_0 x_0/x \quad , \tag{18}$$

as given by (12). Indeed, making use of (5), equation (17) becomes

$$\frac{d\widetilde{N}}{N_0 dx} = p_0(1 + x_0/x)(1 - x/x_0) = p_0 x_0/x - p_0 x/x_0 \simeq p_0 x_0/x \quad , \tag{19}$$

which coincides with (18) to the same order of approximation. In addition, function given by (18) or (19) differs by p_0 at x_0 from function given by (5), which may define a lower-bound cutoff c by equalling the corresponding areas. It leads to

$$p_0 x_0 \ln(x_c/c) = p_0 x_c + p_0 x_c \quad , \tag{20}$$

which is equation (7). It is now easily to read on this equation the number $N_{c0} = N_0 p_0 x_c$ of principal data contained in the total number $\widetilde{N}_0 = N_0 p_0 x_0 \ln(x_c/c)$.

The present analysis can be generalized to $p(x) = p_0(1 - h)$, where $h \ll 1$ for $x < x_c$, leading to a singular distribution $\sim p_0/h$, and conversely, a singular distribution of the form p_0/h can always be associated to a finite distribution of the form $p_0(1-h)$, providing the condensation phenomenon is accounted for. It amounts to a full distribution of the form $p_0 x_c \delta(x) + p(x)$ for $x < x_c$, and to $\widetilde{p}(x)$ for $x_c < x < D$. Function h may also behave like an x-power, $h \sim x^\beta$, in the neighbourhood of the origin, $\beta > 0$and $\beta \neq 1$, with the same conclusion: a condensation phenomenon in origin.