Densities and Senses

Consider a collection X of points x_k in one-dimensional space:

$$X = \{x_1, x_2, x_3, ..., x_k, ..., x_{N-1}, x_N\}$$

Suppose the number of points in this 1-D *points cloud* is N. As a rule, the points x_k are unevenly spaced. They may form a regularly ordered discrete set, an randomly distributed set or a continuous set, the latter being the case when points can no longer be distinguished from each other.

Exact Densities

Let us start with Regular Discrete sets. It may be questioned with such sets what a Density is supposed to mean. Without doubt, by far the simplest example of such a *Density* $P(x)$ is the constant density. There will hardly be any argument about a definition like this:

$$P(x) = \text{constant} \quad \Longleftrightarrow \quad x_k = \frac{k}{N}$$

Next consider a case in which Density is increasing in a linear fashion. What does it mean: linear ? Well, suppose there is one point x_k between 0 and 1, then there are two points x_k between 1 and 2, three points x_k between 2 and 3, and so on. Suppose the initial sampling is Δ, then, if we arrive at $x = k.\Delta$, the number of points has increased to $1 + 2 + 3 + ... + k = k(k+1)/2$. Therefore our basic equation is:

$$x_{k(k+1)/2} = L.\Delta$$

In this way, the array x is only defined for certain values of its index, namely: $1, 3, 6, ...$ In order to generalize for all values of the index, L must be solved from:

$$k(k+1)/2 = L \quad \Longrightarrow \quad k^2 + k - 2L = 0 \quad \Longrightarrow \quad k = \frac{\sqrt{8L+1}-1}{2}$$

A linear increasing density function is thus generated by:

$$x_L = \frac{\sqrt{8L+1}-1}{2}\Delta$$

The first few values (for $\Delta = 1$):

$x_0 = 0\,,\ x_1 = 1\,,\ x_2 = 1.56155\,,\ x_3 = 2\,,\ x_4 = 2.37228\,,\ x_5 = 2.70156\,,\ x_6 = 3\,,\ ...$

Other Exact Densities can be constructed by assuming that the integral over the accumulated density P up to x_k is equal to the number of points involved, divided by N. More precise, while taking care of scaling factors:

$$\int_{-\infty}^{x_k} P(t)\,dt = \frac{k}{N}\int_{-\infty}^{+\infty} P(t)\,dt$$

Repeat for the *Constant Density*, which is given by $P(x) = C$. Let the area of interest be restricted to $[0, 1]$. Resulting in:

$$\int_0^{x_k} C\,dt = C.x_k = \frac{k}{N}.C \quad \Longrightarrow \quad x_k = \frac{k}{N}$$

Another example has been the *Linear Density*, which is given by $P(x) = x$. Again, let the area of interest be restricted to $[0, 1]$. Giving:

$$\int_0^{x_k} t\,dt = \frac{1}{2}x_k^2 = \frac{k}{N}.\frac{1}{2} \quad \Longrightarrow \quad x_k = \sqrt{\frac{k}{N}}$$

But wait, this outcome seems to be different from the one we have obtained in an earlier stage:

$$\text{Compare} \qquad x_k = \frac{\sqrt{8k+1}-1}{2}\Delta \qquad \text{with} \qquad x_k = \sqrt{\frac{k}{N}}$$

Assume that $x_N = 1$, then it follows that $\Delta = 2/(\sqrt{8N+1}-1)$ and, indeed, for large values of k:

$$x_k = \frac{\sqrt{8k+1}-1}{\sqrt{8N+1}-1} \approx \sqrt{\frac{k}{N}}$$

A third example is the Density which is given by:

$$P(x) = \frac{d/\pi}{d^2+x^2}$$

The area of interest is $[-\infty, +\infty]$. Resulting in:

$$\int_{-\infty}^{x_k} \frac{d/\pi}{d^2+t^2}\,dt = \frac{1}{\pi}\int_{-\infty}^{x_k} \frac{d(t/d)}{1+(t/d)^2} = \frac{1}{\pi}arctan(x_k/d) + \frac{1}{2} = \frac{k}{N}.1 \quad \Longrightarrow$$

$$arctan(x_k/d) = \frac{k}{N}\pi - \frac{1}{2}\pi \quad \Longrightarrow \quad x_k = d\,tan\left(\frac{k}{N}\pi - \frac{1}{2}\pi\right)$$

Where the possible values of k need to be be restricted: $0 < k < N$.
It's also possible to formulate the reverse problem: how to find the Density Function $P(x)$ if the point cloud $\{x_k\}$ has been given. An illustrative example is the distribution of the stops on a guitar's neck. It can be shown that it is given by:

$$x_k = L\left[1 - \left(\frac{1}{2}\right)^{k/12}\right]$$

Where L is the length of the strings. The case x_0 for $k = 0$ corresponds with the full length of a string and $k = 1/2$ corresponds with a point halfway on a string (sounds an octave higher). Rewrite the above formula, as follows:

$$\frac{x_k}{L} = 1 - \left(\frac{1}{2}\right)^{k/12} \quad \Longrightarrow \quad 1 - \frac{x_k}{L} = \left(\frac{1}{2}\right)^{k/12} \quad \Longrightarrow$$

$$ln\left(1 - \frac{x_k}{L}\right) = \frac{k}{12}\, ln(\frac{1}{2})$$

The problem is squeezed now into standard form. We see that the total number of points equals 12, which is the number of notes in an octave. And:

$$\int_0^{x_k} P(t)\, dt = ln\left(1 - \frac{x_k}{L}\right) \quad \Longrightarrow \quad \int_0^{x} P(t)\, dt = ln\left|1 - \frac{x}{L}\right| \quad \Longrightarrow$$

$$P(x) = \frac{1/L}{1 - x/L} = \frac{1}{L - x}$$

By differentiation at both sides (and eventually ignoring a minus sign). The Density Function $P(x)$ starts at $(0,1)$ and it has a vertical asymptote for $x = L$ (: the highest notes cannot be reached).

Density Distributions

Consider again the collection X of points x_k in one-dimensional space:

$$X = \{x_1, x_2, x_3, ..., x_k, ..., x_{N-1}, x_N\}$$

A continuous function $\overline{P}(x)$ may be associated with any such a collection of points by allowing each weight m_k of a point x_k to be a continuous and differentiable function of x. We will be even more specific, though, and define:

$$m_k(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}(x-x_k)^2/\sigma^2}$$

Meaning that each point (quickly) gains less weight at a greater distance from its center. For N points x_k, all weights are summed:

$$\overline{P}(x) = \frac{1}{N}\sum_k m_k(x) = \frac{1}{N}\sum_k \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}(x-x_k)^2/\sigma^2} = \frac{1}{\sigma N\sqrt{2\pi}}\sum_{k=1}^{N} e^{-(x-x_k)^2/2\sigma^2}$$

The function $\overline{P}(x)$ will be called a *Density Distribution* associated with the points x_k.

A great advantage of defining a continuous (and even differentiable) function over a discrete set is that the whole apparatus of classical analysis will be at our disposal in this way.

To begin with, density distributions can be *integrated*:

$$\int_{-\infty}^{+\infty} \overline{P}(x)dx = \frac{1}{\sigma N\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sum_{k=1}^{N} e^{-(x-x_k)^2/2\sigma^2} dx =$$

$$\frac{1}{N}\sum_{k=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-(x-x_k)^2/2\sigma^2} dx = \frac{1}{N}\sum_{k=1}^{N} 1 = 1$$

Hence the integral over any Density Distribution happens to be exactly equal to unity:

$$\int_{-\infty}^{+\infty} \overline{P}(x)dx = 1$$

Is this a coincidence? No, of course. We simply have defined our *norming factor* $1/(\sigma N\sqrt{2\pi})$ in such a way, that such behaviour of Density Distributions is to be expected.
We will demonstrate now that the Density Distributions, thus defined here, are a *sensible approximation* to the exact Density Functions, as defined in a previous paragraph. The definition there is repeated:

$$\int_{-\infty}^{x_k} P(t)\,dt = \frac{k}{N}\int_{-\infty}^{+\infty} P(t)\,dt$$

Assuming that the Density Function P is properly *normed*, the expression may be simplified to:

$$\int_{-\infty}^{x_k} P(t)\,dt = \frac{k}{N}$$

Where:

$$\int_{-\infty}^{+\infty} P(x)dx = 1$$

What we will do first is make a *Continuation* of this definition, herewith answering the question what will happen if the number N of points x_k in the points cloud approximates infinity: $N \rightarrow \infty$. The fraction k/N then approximates a continuous variable y, while it is still restricted to the same interval $[0, 1]$:

$$\int_{-\infty}^{x} P(t)\,dt = y(x) \qquad \text{where} \qquad 0 \le y \le 1$$

If we are able to determine the inverse function μ, which is defined by $\mu(y(x)) = x$ or $y(\mu(x)) = x$, then $x = \mu(y)$ will be the continuation of the points cloud x_k.
At the other hand, we have the definition of a Density Distribution. If the latter is continuated, then the result is an integral instead of a sum:

$$\overline{P}(x) = \frac{1}{N}\sum_k m_k(x) = \sum_{k=1}^{N} m_k(x)\frac{1}{N} \quad \Longrightarrow$$

$$\overline{P}(x) = \int_0^1 w_y(x)dy \qquad (\text{because } 0 \le y \le 1\ )$$

Here the discrete variable k (or k/N) has been replaced by the continuous variable y, the corresponding discrete finitesimal $1/N$ has been replaced by the continuous infinitesimal dy and the discrete variable m_k has been replaced by the continuous variable w. Remember that $x = \mu(y)$ represents the continuation of the points cloud x_k. Then:

$$w_y(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}[x-\mu(y)]^2/\sigma^2}$$

The integral becomes:

$$\overline{P}(x) = \int_0^1 w_y(x)\, dy = \int_0^1 \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}[x-\mu(y)]^2/\sigma^2}\, dy$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}[x-\mu(y)]^2/\sigma^2} \left(\frac{dy}{d\mu}\right) d\mu$$

Here the change of independent variable $y \to \mu$ invokes a change of the integration interval from $[0, 1]$ to $[-\infty, +\infty]$. Moreover, the fraction $dy/d\mu$ denotes the derivative of the inverse function of $\mu(y)$, which is $y(\mu)$. And:

$$\frac{dy(\mu)}{d\mu} = \frac{d}{d\mu}\int_{-\infty}^{\mu} P(t)\, dt = P(\mu)$$

Then the integral becomes:

$$\overline{P}(x) = \int_0^1 w_y(x)\, dy = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}[x-\mu]^2/\sigma^2}\, P(\mu)\, d\mu$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, P(x-\mu)\, d\mu$$

We could stop here and simply state that any function P as it is convoluted to $\overline{P}$ with a Gaussian function is an approximation to the original. That is, we would have reduced the Discrete to the Continuous case.
But we won't stop, since there is quite a concise proof for the Continuous case. Develop the function $P(x - \mu)$ into a power series around x:

$$P(x-\mu) = P(x) - \mu P'(x) + \frac{1}{2}\mu^2 P''(x) + ...$$

And substitute into the integral, then:

$$\overline{P}(x) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, P(x-\mu)\, d\mu =$$

$$P(x)\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, d\mu \quad - \quad P'(x)\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, \mu\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, d\mu$$

$$+ \quad \frac{1}{2}P''(x)\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, \mu^2\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, d\mu \quad - \quad ... \quad \Longrightarrow$$

$$\overline{P}(x) \approx P(x) + \frac{1}{2}\sigma^2 . P''(x)$$

Where use has been made of known properties of Gaussian (normal) distribution functions:

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, d\mu = 1 \qquad \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, \mu\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, d\mu = 0$$

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, \mu^2\, e^{-\frac{1}{2}\mu^2/\sigma^2}\, d\mu = \sigma^2$$

Thus we see that, indeed, any exact Density Function is approximated by an accompanying Density Distribution, with an accuracy which quickly increases as the standard deviation σ of the Gaussian weight functions $w_y(x)$ is diminished accordingly. A preliminary necessary condition for the transition from Discrete to Continuous to take place is: $dy < \sigma$ or $\sigma > 1/N$. (And, with help of Shannon's Sampling Theorem, this can be further refined to $dy < \frac{1}{2}\sigma$, as has been demonstrated elsewhere.)

Differentiation

We can differentiate the Density Function $P(x)$ with respect to x:

$$P'(x) = \sum_{k=1}^{N} -\frac{x - x_k}{\sigma^2} e^{-(x-x_k)^2/2\sigma^2} = \frac{1}{\sigma^2}\left[\sum_k m_k(x) x_k - x \sum_k m_k(x)\right]$$

The *mean value* of x has been defined as the following quotient:

$$\overline{x} = \frac{\sum_k m_k x_k}{\sum_k m_k} = \left[\sum_{k=1}^{N} x_k\, e^{-(x-x_k)^2/2\sigma^2}\right] / \left[\sum_{k=1}^{N} e^{-(x-x_k)^2/2\sigma^2}\right]$$

Then it is seen that:

$$P'(x) = \frac{1}{\sigma^2} \sum_{k=1}^{N} e^{-(x-x_k)^2/2\sigma^2} (\overline{x} - x) \quad \Longrightarrow$$

$$P'(x) = \frac{P(x)}{\sigma^2}(\overline{x} - x) = \frac{P(x)}{\sigma^2}(\mu_x - x)$$

The density function may be differentiated twice:

$$P'(x) = \sum_{k=1}^{N} \frac{-(x - x_k)}{\sigma^2} e^{-(x-x_k)^2/2\sigma^2} \quad \Longrightarrow$$

$$P''(x) = \sum_{k=1}^{N} \left[\left(\frac{x - x_k}{\sigma^2} \right)^2 e^{-(x-x_k)^2/2\sigma^2} - \frac{1}{\sigma^2} e^{-(x-x_k)^2/2\sigma^2} \right]$$

Rewrite as follows:

$$P''(x) = \frac{1}{\sigma^4} \sum_{k=1}^{N} (x^2 - 2xx_k + x_k^2 - \sigma^2) e^{-(x-x_k)^2/2\sigma^2}$$

The *mean value* of x^2 has been defined as the following quotient:

$$\overline{x^2} = \sum_k m_k x_k^2 / \sum_k m_k \left[\sum_{k=1}^{N} x_k^2 \, e^{-(x-x_k)^2/2\sigma^2} \right] / \left[\sum_{k=1}^{N} e^{-(x-x_k)^2/2\sigma^2} \right]$$

Then we find:

$$P''(x) = \frac{P(x)}{\sigma^4} (x^2 - 2x\overline{x} + \overline{x^2} - \sigma^2)$$

Rework towards a moment of inertia with respect to the midpoint μ_x:

$$P''(x) = \frac{P(x)}{\sigma^4} \left[\overline{x^2} - \overline{x}^2 + (x - \overline{x})^2 - \sigma^2 \right] = \frac{P(x)}{\sigma^4} \left[\sigma_{xx} + (x - \mu_x)^2 - \sigma^2 \right]$$

It is seen that the second derivative is equal to the Density, times the total moment of inertia minus the spread of the Gass function. The total moment of inertia, in turn, consists of two terms: the moment of inertia of the points cloud with respect to the midpoint plus the moment of inertia of the midpoint with respect to the origin.

Fourier Series

Consider an important limiting case. Suppose there are infinitely many points x_k in the collection X and they are all evenly spaced: $x_k = k\Delta$. Thus, apart from a norming constant, we have the density function:

$$P(x) = \sum_{L=-\infty}^{+\infty} e^{-(x-L\Delta)^2/2\sigma^2}$$

It is easily shown that the above density function is *periodic*. Its period is equal to Δ: $P(x + \Delta) = P(x)$ for arbitrary x. This means that $P(x)$ can be developed into a *Fourier series*. In addition, the function is *even*, meaning that $P(x) = P(-x)$, which results in real-valued Fourier coefficients A_k. They are calculated initially as complex-valued entities.

$$A_k + iB_k = \frac{1}{\Delta/2} \int_{-\Delta/2}^{+\Delta/2} P(x) e^{i\,k\,2\pi/\Delta\,x} dx =$$

In the sequel, kind of an angular frequency ω will stand for the quantity $\omega = 2\pi/\Delta$. Then let calculations continue:

$$\frac{1}{\Delta/2} \int_{-\Delta/2}^{+\Delta/2} \sum_{L=-\infty}^{+\infty} e^{-(x-L\Delta)^2/2\sigma^2} e^{ik\omega x} dx =$$

$$\frac{1}{\Delta/2}\sum_{L=-\infty}^{+\infty}\int_{-\Delta/2}^{+\Delta/2} e^{-(x-L\Delta)^2/2\sigma^2}e^{ik\omega x}dx$$

Substitute $y = x - L\Delta$ and integrate to y:

$$A_k + iB_k = \frac{1}{\Delta/2}\sum_{L=-\infty}^{+\infty}\int_{-\Delta/2-L\Delta}^{+\Delta/2-L\Delta} e^{-y^2/2\sigma^2}e^{ik\omega(y+L\Delta)}dy =$$

Where:

$$e^{ik\omega(y+L\Delta)} = e^{ik\omega y}e^{ikL2\pi} = e^{ik\omega y}.1$$

Next replace y by $-y$ and switch integration bounds:

$$A_k + iB_k = \frac{1}{\Delta/2}\sum_{L=-\infty}^{+\infty}\int_{L\Delta-\Delta/2}^{L\Delta+\Delta/2} e^{-y^2/2\sigma^2}e^{ik\omega(-y)}dy$$

The above integrals are precisely the adjacent pieces of another integral which has bounds reaching to infinity. That is, they sum up to an infinite integral:

$$A_k + iB_k = \frac{1}{\Delta/2}\int_{-\infty}^{+\infty} e^{-y^2/2\sigma^2}e^{-ik\omega y}dy =$$

$$\frac{1}{\Delta/2}\int_{-\infty}^{+\infty} e^{-y^2/2\sigma^2-ik\omega y}$$

The part in the exponential function can be written as follows:

$$-y^2/2\sigma^2 - ik\omega y = -\frac{1}{2}(y^2/\sigma^2 - 2.ik\omega\sigma y/\sigma) =$$

$$-\frac{1}{2}\{y^2/\sigma^2 - 2.ik\omega\sigma y/\sigma + (ik\omega\sigma)^2\} + \frac{1}{2}(ik\omega\sigma)^2 =$$

$$-(y/\sigma - ik\omega\sigma)^2/2 - (k\omega\sigma)^2/2$$

Resulting in:

$$\frac{1}{\Delta/2}\int_{-\infty}^{+\infty} e^{-(y/\sigma-ik\omega\sigma)^2/2}e^{-(k\omega\sigma)^2/2}\,dy =$$

$$e^{-(k\omega\sigma)^2/2}\frac{1}{\Delta/2}\int_{-\infty}^{+\infty} e^{-(y-ik\omega\sigma^2)^2/2\sigma^2}\,dy$$

We know that the integral is equal to $\sigma\sqrt{2\pi}$, giving at last:

$$A_k + iB_k = A_k = \frac{\sigma\sqrt{2\pi}}{\Delta/2}e^{-(k\omega\sigma)^2/2}$$

The Fourier series of any periodic function is given by:

$$P(x) = \frac{1}{2}A_0 + \sum_{k=1}^{\infty} A_k \, cos(k\omega x)$$

We conclude that the Fourier series of a uniform Density Function is given by:

$$P(x) = \sum_{L=-\infty}^{+\infty} e^{-(x-L\Delta)^2/2\sigma^2} = \sigma\sqrt{2\pi}\left[\frac{1}{\Delta} + \frac{1}{\Delta/2}\sum_{k=1}^{\infty} e^{-(k\omega\sigma)^2/2} \, cos(k\omega x)\right]$$

Where we remind of the fact that: $\omega = 2\pi/\Delta$. Re-introduce proper norming factors, at last:

$$P(x) := \frac{1}{\sigma N\sqrt{2\pi}} P(x) = \frac{1}{N}\left[\frac{1}{\Delta} + \frac{1}{\Delta/2}\sum_{k=1}^{\infty} e^{-(k\omega\sigma)^2/2} \, cos(k\omega x)\right]$$