Gaussian Probability Distribution Derivation

— Joaquín Gómez.

— February 12, 2025.

To start with, we look at a plane and we try to aim it with darts, the bullseye is centered at $(0, 0)$ and we follow these 3 rules:

a. Rotating the frame of reference does not modify the deviation of the points.
b. Deviations in orthogonal axes are independent. This means that a deviation in x keeps the probability in y unchanged, and vice-versa.
c. Larger deviations are less likely than smaller ones. The common sense to this is that we aim to the center of the objective.

We look at an infinitesimal region of our probability distribution for both axes

$$P(x \leq X \leq x + \Delta x) = \int_{x}^{x+\Delta x} p(x)dx \approx p(x)\Delta x \qquad \text{when } \Delta x \to 0$$

$$P(y \leq Y \leq y + \Delta y) = \int_{y}^{y+\Delta y} p(y)dy \approx p(y)\Delta y \qquad \text{when } \Delta y \to 0$$

Now, let the joint probability distribution of x and y be $p(x, y)$. Both variables are independent, this means that we can factor out the distribution as a product of two different functions. The meaning of p is the combined probability in the area given by $\Delta A = |\Delta x \cdot \Delta y|$.

$$p(x, y) = p(x)p(y)$$

Then, the probability in an infinitesimal ΔA is

$$P(X = x, Y = y) \approx p(x, y)\Delta A = p(x)p(y)\Delta x\Delta y$$

(we always assume that $\Delta x\Delta y > 0, \forall x, y \in \mathbb{R}$). This small area can be represented in polar coordinates by replacing $x = r\cos\theta, y = r\sin\theta$. But we want to measure the distribution with respect to the distance to the center, this means that the angle θ can be discarded since it does not modify the distance.

So we replace $g(r) = p(x)p(y) = p(r\cos\theta)p(r\sin\theta)$, for some $\theta \in \mathbb{R}$. We will bound theta to be in the range $[0, 2\pi]$.

Let $g(r)$ be continuous differentiable, then taking $\frac{d}{d\theta}g(r)$ should give zero since the equation sets θ to be constant.

$$\frac{d}{d\theta}g(r) = p(r\cos\theta)\frac{d}{d\theta}p(r\sin\theta) + \frac{d}{d\theta}p(r\cos\theta)p(r\sin\theta)$$

Using the chain rule and going back to Cartesian coordinates

$$p(x)\,\frac{dp}{dy}\frac{dy}{d\theta} + p(y)\,\frac{dp}{dx}\frac{dx}{d\theta}$$

Now, if we return to polar coordinates we find the derivatives of x and y with respect to θ. Thus we can write finally

$$p(r\cos\theta)p'(r\sin\theta)r\cos\theta - p(r\sin\theta)p'(r\cos\theta)r\sin\theta = 0$$

Let's reorder terms and return to Cartesian coordinates.

$$\frac{p'(x)}{p(x)x} = \frac{p'(y)}{p(y)y}$$

This is a separable differential equation. These equations do not depend on x and y. Seeing it from the polar coordinate perspective

$$\frac{p'(r\cos\theta)r\sin\theta}{p(r\cos\theta)} = \frac{p'(r\sin\theta)r\cos\theta}{p(r\sin\theta)}$$

we can easily factor out the r by dividing by r itself, and since θ is not a parameter, we get that the equation does not depend either in x or y.

Now, looking at any function f differentiable at any point. More exactly, let's find some function that satisfies the form $f'(x)/f(x) = c$. If we look at the exponential, say

$$f(x) = a^x$$

and get the first derivative

$$f'(x) = \ln(a) \cdot a^x = ln(a) f(x) \longrightarrow \frac{f'(x)}{f(x)} = ln(a)$$

Note that $ln(a) = c$ is just a constant. So the exponential function satisfies this requirement. For simplicity take $a = e$, the Euler constant. And solve the following equation

$$\frac{p'(x)}{p(x)x} = k \qquad \text{for some } k \in \mathbb{R}$$

Take the indefinite integral

$$\int \frac{p'(x)}{p(x)} dx = \int kx \, dx$$

and note that $p'(x)dx = dp$, thus we get

$$\int \frac{dp}{p} = \frac{1}{2} kx^2$$

then

$$\ln(p(x)) + c = \frac{1}{2} kx^2$$

and solving for $p(x)$

$$p(x) = \exp\left(\frac{1}{2} kx^2 - c\right) = A \exp\left(\frac{1}{2} kx^2\right) \qquad A = e^{-c}$$

We have successfully obtained $p(x)$! Now the problem that remains is to find a way to make $p(x)$ satisfy the next conditions:

a. $p(x)$ must be defined in all $\mathbb{R}$

b. $\int_{\mathbb{R}} p(x) dx = 1$

c. $p(x) \geq 0, \forall x \in \mathbb{R}$

The function we've found complies with (a), we have to find some constant term A such that (b) and (c) are satisfied. To satisfy (b) we have to set k to be

always negative, so we replace k with $-k$ in the exponent and constraint $k \geq 0$
.

$$p(x) = Ae^{-\frac{1}{2}kx^2}$$

Then we have to solve

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

So we can write

$$\int_{\mathbb{R}} e^{-\frac{1}{2}kx^2} dx = \frac{1}{A}$$

Now, this is a hard integral, but we can exploit a trick to solve it. Note that $p(y) = p(x)$, this implies that the probabilities are distributed equally on both axes. Then, if we square the equation we are allowed to write

$$\left(\int_{\mathbb{R}} e^{\frac{1}{2}kx^2} dx\right)\left(\int_{\mathbb{R}} e^{\frac{1}{2}ky^2} dy\right) = \frac{1}{A^2}$$

Combining the integrals

$$\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}kx^2} e^{-\frac{1}{2}ky^2} dxdy\right) = \frac{1}{A^2}$$

Simplifying the integrand

$$\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}k(x^2+y^2)} dxdy\right) = \frac{1}{A^2}$$

This integral is easy to solve in the polar coordinate system. Remember the definition of a double integral in polar coordinates to be $\iint f(r,\theta)r\, drd\theta$. Then $x^2 + y^2 = r^2$ (check it yourself). Note that r is always positive in polar coordinates. The bounds of r are then $(-\infty, 0]$.

$$\int_0^{2\pi}\int_0^{\infty} e^{-\frac{1}{2}kr^2} rdrd\theta = \frac{1}{A^2}$$

Let's change the bounds of integration of θ to be from 0 to $\frac{\pi}{2}$, so scale the integral by 4

$$4\int_0^{\pi/2}\int_0^{\infty} e^{-\frac{1}{2}kr^2} r dr d\theta = \frac{1}{A^2}$$

Divide both sides by 4 and get

$$\int_0^{\pi/2}\int_0^{\infty} e^{-\frac{1}{2}kr^2} r dr d\theta = \frac{1}{4A^2}$$

Now let $u = \frac{1}{2}kr^2$, so $du = krdr$. Replacing this on our integral we get

$$\frac{1}{k}\int_0^{\pi/2}\int_0^{\infty} e^{-u} du d\theta = \frac{1}{4A^2}$$

This solves easily, first integrate with respect to u and get

$$-e^{-u}\Big|_0^{\infty} = e^0 - \lim_{t\to\infty} e^{-t} = 1$$

Plug it in the integral and find that

$$\int_0^{\pi/2} 1 d\theta = \frac{k}{4A^2}$$

Finally, we get that

$$\frac{\pi}{2} = \frac{k}{4A^2} \longrightarrow A = \pm\sqrt{\frac{k}{2\pi}}$$

To satisfy (c) we must only take $A \geq 0$, so only the positive result matters here.

The final expression for $p(x)$ is

$$p(x) = \sqrt{\frac{k}{2\pi}} e^{-\frac{1}{2}kx^2}$$

Now, suppose we want to move the origin to $x = \mu$, where μ will be the mean of the distribution. This means we have to do $p(x-\mu)$, thus we can just replace it on the function and get

$$p(x) = \sqrt{\frac{k}{2\pi}} e^{-\frac{1}{2}k(x-\mu)^2}$$

To complete the derivation, we want to find the inflection points of the function. So, taking the second derivative we get

$$p''(x) = \sqrt{\frac{k^3}{2\pi}} \left(e^{-\frac{1}{2}k(x-\mu)^2} - k(x-\mu)^2 e^{-\frac{1}{2}k(x-\mu)^2} \right)$$

This seems messy, but setting $p''(x) = 0$ simplifies it to

$$e^{-\frac{1}{2}k(x-\mu)^2} = k(x-\mu)^2 e^{-\frac{1}{2}k(x-\mu)^2}$$

This implies that $k(x-\mu)^2 = 1$, solving this simple equation we find out the following expression

$$\frac{1}{k} = (x-\mu)^2 \longrightarrow x = k^{-1/2} + \mu$$

The constant $k^{-1/2}$ called the Standard Deviation, the variance is the square of the SD

$$\sigma = k^{-1/2}, \quad \sigma^2 = k^{-1}$$

The variance tells us the amount of dispersion the values have from the mean. In this case the constant term k is referred in some literature as β, and it is called the precision.

We will rewrite the formula for $p(x)$ in terms of the variance.

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

— END — (February 12, 2025)