

Math3810 - Probability
Section 001 - Fall 2025
Notes: The Gaussian Normal Distribution

University of Colorado Denver / College of Liberal Arts and Sciences

Department of Mathematics - Dr. Robert Rostermundt

Introduction: In these notes we discuss a derivation of the density function for the Gaussian normal distribution.

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ for all } x \in \mathbb{R}$$

For many students the origins of this density function seem even more mysterious than the origins of the densities for the Poisson distributions. While we usually demonstrate the Poisson distribution as a limiting case of the binomial distribution most courses do not give a demonstration of how the density for the normal distribution arises. In these notes we will give a straightforward such development, keeping in mind that this development will be different from both those original ones given by Abraham de Moivre (1783) or Carl Frederick Gauss (1809). Moreover, we will develop the density for the standard normal distribution where $\mu = 0$ and $\sigma = 1$, from which a change of variables will give rise to the general normal distribution.

The Experiment: Consider an infinite dart board. Our experiment is to throw a dart, aiming at the origin $(0, 0)$. We are not able to throw with enough accuracy to hit the origin each time and so our dart will land at some random point (X, Y) with joint density function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. We make the following reasonable assumptions:

1. The density is rotationally invariant. That is, the distribution of where my dart lands only depends on the distance of the dart to the center.
2. The random variables X and Y are independent and identically distributed. How much the dart misses left-to-right makes no difference to how much the dart misses up-and-down.

The Math: Given assumption(1) we must be able to express the joint density

$$f(x, y) = g(x^2 + y^2).$$

Moreover, as the random variable X and Y are independent identically distributed we must be able to write

$$f(x, y) = f(x, 0)f(0, y).$$

But then we have

$$g(x^2 + y^2) = f(x, y) = f(x, 0)f(0, y) = g(x^2)g(y^2).$$

But this guarantees that g is an exponential function. That is, we can write

$$g(t) = Ae^{-Bt}$$

for some non-negative constants A and B . Here A will be our normalizing constant and B is just reflecting the units we are measuring on the t -axis. Also, observe that we have argument $-Bt$ as the density should be a decreasing function of distance to the origin. The point (x, y) corresponds to error of the dart throw and we are not throwing the dart at random, but aiming (with some level of skill) at the origin $(0, 0)$. Our goal now is to work out the constants A and B .

Since f is a density function we must have

$$\int \int_{\mathbb{R}^2} f(x, y) dA = 1.$$

Using our assumptions, and a change to polar coordinates, we evaluate the integral as follows.

$$\begin{aligned} \int \int_{\mathbb{R}^2} f(x, y) dA &= \int \int_{\mathbb{R}^2} g(x^2 + y^2) dA \\ &= \left(\int_{\theta=0}^{\theta=2\pi} d\theta \right) \left(\int_{r=0}^{\infty} rg(r^2) dr \right) \\ &= 2\pi \int_{r=0}^{\infty} rg(r^2) dr \\ &= 2\pi \lim_{c \rightarrow \infty} \int_{r=0}^{r=c} rg(r^2) dr \\ &= \pi \lim_{c \rightarrow \infty} \int_{t=0}^{t=c^2} g(t) dt \\ &= A\pi \lim_{c \rightarrow \infty} \int_{t=0}^{t=c^2} e^{-Bt} dt \\ &= -\frac{A\pi}{B} \lim_{c \rightarrow \infty} [e^{-Bc^2} - e^0]_{t=0}^{t=c^2} \\ &= \frac{A\pi}{B} \end{aligned}$$

Therefore we must have $A = B/\pi$. Since B reflects the units in which we are measuring it seems reasonable to let B be a function of the standard deviation σ . For reasons we will justify later in these notes, we will let $B = \frac{1}{2\sigma^2}$ which then implies $A = \frac{1}{2\pi\sigma^2}$ and we have

$$g(t) = \frac{1}{2\pi\sigma^2} e^{-t^2/2\sigma^2}.$$

Lets now define $\tilde{f}(x) = f(x, 0) = f(0, x)$. This gives us the following:

$$f(x, y) = \tilde{f}(x)\tilde{f}(y) \implies \tilde{f}(x)\tilde{f}(y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

We have finally arrived at our goal of developing the standard normal density function.

$$\tilde{f}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}.$$

Returning to our choice of B , we recall that in our construction we were free to choose $B > 0$ as we saw fit. If we compute the variance of the distribution with density function

$$\tilde{f}(x) = \sqrt{\frac{B}{\pi}} e^{-Bx^2} \text{ we get}$$

$$\text{Var}(X) = \sqrt{\frac{B}{\pi}} \int_{-\infty}^{\infty} t^2 e^{-Bt^2} dt = \sqrt{\frac{B}{\pi}} \cdot \frac{\sqrt{\pi}}{2B^{3/2}} = \frac{1}{2B}.$$

Therefore, by letting $B = \frac{1}{2\sigma^2}$ we have the nice property that the σ in the expression of our function \tilde{f} is the actual standard deviation of the distribution. Of course, it also satisfies our original goal of constructing the density function for the standard normal distribution. So while other choices $B > 0$ would have given us a valid density function they would not have given us our intended density for the standard normal distribution. Our construction is now complete.

We observe that by aiming our dart at some point (μ_1, μ_2) , with a simple change of variables we would have obtained the general density for a normal distribution.

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

So where does the exponential fit into the picture? This part of the density follows from our assumptions about independence of X and Y . Where does the π fit in? It follows from our assumptions of rotational invariance. The terms μ and σ follow from the coordinates we aim the dart and the units of how we measure distances. The mystery seems to have unraveled.

Some Extensions: This model can also help to infer some nice properties of normal distributions. Let suppose that the first dart lands at position (x_1, y_1) , and another dart is thrown, but now aimed at the position (x_1, y_1) and landing at position (x_2, y_2) where $X_2 = X + X_1$ and $Y_2 = Y + Y_1$. The position of the second dart is the sum of the two errors and X_2 and Y_2 are still independent. Moreover, the joint density of X_2 and Y_2 is rotationally invariant. Since we satisfy the original assumptions we see that X_2 and X_1 are independent normal random variables and their sum is also a normal random variable.

This suggests that a sum of independent normal random variables is also a normal random variable. And if we take the average of a long sequence of independent random variables we should get something with the same shape, no matter how long the sequence is. This is a useful property, which does not hold true for all random variables. And it suggests that the accumulation of errors which are independent and normal can be assumed to have a normal distribution. This is a very powerful statement and one of the reasons the Gaussian normal distribution is often considered the most important distribution in applied probability and statistics.

The interested student might now ask about the accumulation of errors which do not have a normal distribution. This is an important question and we direct the student to the important Central Limit Theorem for further investigations.