The Normal Distribution: A derivation from basic principles

Dan Teague
The North Carolina School of Science and Mathematics

Introduction

Students in elementary calculus, statistics, and finite mathematics classes often learn about the normal curve and how to determine probabilities of events using a table for the standard normal probability density function. The calculus students can work directly with the normal probability density function $p(x)=\frac{1}{s\sqrt{2p}}e^{-\frac{1}{2}\left(\frac{x-m}{s}\right)^2}$ and use numerical integration techniques to compute probabilities without resorting to the tables. In this article, we will give a derivation of the normal probability density function suitable for students in calculus. The broad applicability of the normal distribution can be seen from the very mild assumptions made in the derivation.

Basic Assumptions

Consider throwing a dart at the origin of the Cartesian plane. You are *aiming* at the origin, but random errors in your throw will produce varying results. We assume that:

- the errors do not depend on the orientation of the coordinate system.
- errors in perpendicular directions are independent. This means that being too high doesn't alter the probability of being off to the right.
- large errors are less likely than small errors.

In Figure 1, below, we can argue that, according to these assumptions, your throw is more likely to land in region A than either B or C, since region A is closer to the origin. Similarly, region B is more likely that region C. Further, you are more likely to land in region F than either D or E, since F has the larger area and the distances from the origin are approximately the same.

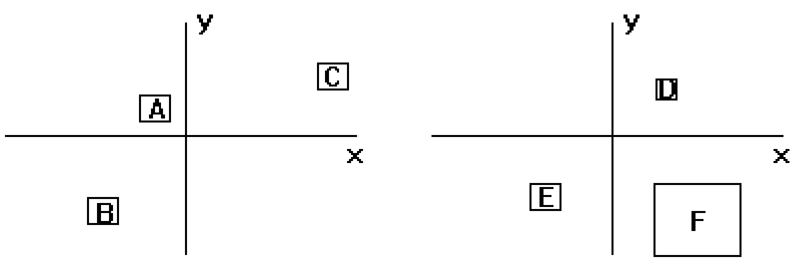

Figure 1

Determining the Shape of the Distribution

Consider the probability of the dart falling in the vertical strip from x to $x+\Delta x$. Let this probability be denoted $p(x)\Delta x$. Similarly, let the probability of the dart landing in the horizontal strip from y to $y+\Delta y$ be $p(y)\Delta y$. We are interested in the characteristics of the function *p*. From our assumptions, we know that function *p* is not constant. In fact, the function *p* is the normal probability density function.

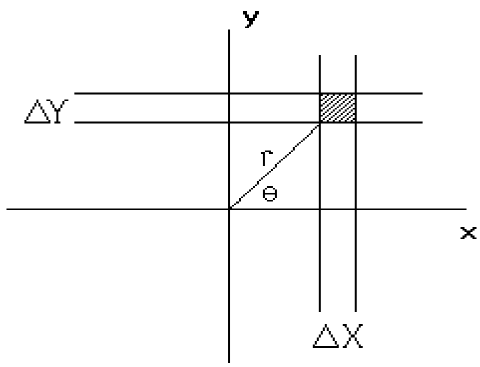

Figure 2

From the independence assumption, the probability of falling in the shaded region is $p(x)\Delta x \cdot p(y)\Delta y$. Since we assumed that the orientation doesn't matter, that any region *r* units from the origin with area $\Delta x \cdot \Delta y$ has the same probability, we can say that

$$p(x)\Delta x \cdot p(y)\Delta y = g(r)\Delta x \Delta y.$$

This means that

$$g(r) = p(x)p(y).$$

Differentiating both sides of this equation with respect to q, we have

$$0 = p(x)\frac{dp(y)}{dq} + p(y)\frac{dp(x)}{dq},$$

since g is independent of orientation, and therefore, q.

Using $x = r\cos(q)$ and $y = r\sin(q)$, we can rewrite the derivatives above as

$$0 = p(x)p'(y)\left(r\cos(q)\right) + p(y)p'(x)\left(-r\sin(q)\right).$$

Rewriting again, we have $0 = p(x)\,p'(y)\,x - p(y)\,p'(x)\,y$. This differential equation can be solved by separating variables,

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

This differential equation is true for any x and y, and x and y are independent. That can only happen if the ratio defined by the differential equation is a constant, that is, if

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

Solving $\frac{p'(x)}{x\,p(x)} = C$, we find that $\frac{p'(x)}{p(x)} = Cx$ and $\ln\big(p(x)\big) = \frac{Cx^2}{2} + c$ and finally,

$$p(x) = Ae^{\frac{C}{2}x^2}.$$

Since we assumed that large errors are less likely than small errors, we know that C must be negative. We can rewrite our probability function as

$$p(x) = Ae^{-\frac{k}{2}x^2},$$

with k positive.

This argument has given us the basic form of the normal distribution. This is the classic bell curve with maximum value at $x = 0$ and points of inflection at $x = \pm\frac{1}{\sqrt{k}}$. We now need to determine the appropriate values of A and k.

Determining the Coefficient *A*

For p to be a probability distribution, the total area under the curve must be 1. We need to adjust A to insure that the area requirement is satisfied. The integral to be evaluated is

$$\int_{-\infty}^{\infty} Ae^{-k\frac{x^2}{2}}\,dx.$$

If $\int_{-\infty}^{\infty} Ae^{-k\frac{x^2}{2}}\,dx = 1$, then $\int_{-\infty}^{\infty} e^{-k\frac{x^2}{2}}\,dx = \frac{1}{A}$. Due to the symmetry of the function, this area is twice that of $\int_{0}^{\infty} e^{-k\frac{x^2}{2}}\,dx$, so

$$\int_{0}^{\infty} e^{-k\frac{x^2}{2}}\,dx = \frac{1}{2A}.$$

Then,

$$\left(\int_{0}^{\infty} e^{-k\frac{x^2}{2}}\,dx\right)\cdot\left(\int_{0}^{\infty} e^{-k\frac{y^2}{2}}\,dy\right) = \frac{1}{4A^2},$$

since x and y are just dummy variables. Recall that x and y are also independent, so we can rewrite this product as a double integral

$$\int_0^\infty \int_0^\infty e^{-\frac{k}{2}\left(x^2+y^2\right)} dy\, dx = \frac{1}{4A^2}\,.$$

(Rewriting the product of the two integrals as the double integral of the product of the integrands is a step that needs more justification than we give here, although the result is easily believed. It is straightforward to show that

$$\left(\int_0^M f(x)\,dx\right)\left(\int_0^M g(y)\,dy\right) = \int_0^M \int_0^M f(x)g(y)\,dy\,dx$$

for finite limits of integration, but the infinite limits create a significant challenge that will not be taken up.)

The double integral can be evaluated using polar coordinates.

$$\int_0^\infty \int_0^\infty e^{-\frac{k}{2}\left(x^2+y^2\right)} dx\, dy = \int_0^{p/2} \int_0^\infty e^{-\frac{k}{2}r^2}\, r\, dr\, dq\,.$$

To evaluate the polar form requires a u-substitution in an improper integral. Performing the integration with respect to r, we have

$$\int_0^{p/2} \int_0^\infty e^{-\frac{k}{2}r^2}\, r\, dr\, dq = \int_0^{p/2} \frac{-1}{k}\left[\int_0^{-\infty} e^u du\right] dq = \int_0^{p/2} \frac{dq}{k} = \frac{p}{2k}\,.$$

Now we know that $\dfrac{1}{4A^2} = \dfrac{p}{2k}$, and so $A = \sqrt{\dfrac{k}{2p}}$. The probability distribution is

$$p(x) = \sqrt{\frac{k}{2p}}\, e^{-\frac{k}{2}x^2}\,.$$

Determining the Value of *k*

A question often asked about probability distributions is "what are the mean and variance of the distribution?" Perhaps the value of k has something to do with the answer to these questions. The mean, m, is defined to be the value of the integral $\int_{-\infty}^{\infty} x\, p(x)\, dx$. The variance, s^2, is the value of the integral $\int_{-\infty}^{\infty} (x-m)^2\, p(x)\, dx$. Since the function $x\, p(x)$ is an odd function, we know the mean is zero. The value of the variance needs further computation.

To evaluate $\int_{-\infty}^{\infty} x^2\, p(x)\, dx = s^2$, we proceed as before, integrating on only the positive *x*-axis and doubling the value. Substituting what we know of $p(x)$, we have

$$2\sqrt{\frac{k}{2p}}\int_0^{\infty} x^2 e^{-\frac{k}{2}x^2}\, dx = s^2 .$$

The integral on the left is evaluated by parts with $u = x$ and $dv = xe^{-\frac{k}{2}x^2}$ to generate the expression

$$2\sqrt{\frac{k}{2p}}\left[\lim_{M\to\infty}\frac{-x}{k}e^{-\frac{k}{2}x_2}\Bigg|_0^M + \frac{1}{k}\int_0^{\infty} e^{-\frac{k}{2}x^2}\, dx\right].$$

Simplifying, we know that $\lim_{M\to\infty}\frac{-x}{k}e^{-\frac{k}{2}x_2}\Big|_0^M = 0$ and we know that $\frac{1}{k}\int_0^{\infty} e^{-\frac{k}{2}x^2}\, dx = \frac{1}{k}\frac{\sqrt{2p}}{2\sqrt{k}}$ from our work before. So $2\sqrt{\frac{k}{2p}}\int_0^{\infty} x^2 e^{-\frac{k}{2}x^2}\, dx = 2\frac{\sqrt{k}}{\sqrt{2p}}\cdot\frac{1}{k}\cdot\frac{\sqrt{2p}}{2\sqrt{k}} = \frac{1}{k}$ so that $k = \frac{1}{s^2}$.

The Normal Probability Density Function

Now we have the normal probability distribution derived from our 3 basic assumptions:

$$p(x) = \frac{1}{s\sqrt{2p}}e^{-\frac{1}{2}\left(\frac{x}{s}\right)^2} .$$

The general equation for the normal distribution with mean m and standard deviation s is created by a simple horizontal shift of this basic distribution,

$$p(x) = \frac{1}{s\sqrt{2p}}e^{-\frac{1}{2}\left(\frac{x-m}{s}\right)^2} .$$

References:

Grossman, Stanley, I., *Multivariable Calculus, Linear Algebra, and Differential Equations*, 2nd., Academic Press, 1986.

Hamming, Richard, W. *The Art of Probability for Engineers and Scientists*, Addison-Wesley, 1991.