THE GAUSSIAN INTEGRAL AND THE GAUSSIAN PROBABILITY DENSITY FUNCTION

DANIEL F. SAVARESE

ABSTRACT. We review moment-generating functions to provide a context for understanding the utility of Gaussian functions and the Gaussian integral. We then explain Gaussian functions as a class of exponential function and demonstrate a common technique for integrating a Gaussian function over $(-\infty, \infty)$ by converting to polar coordinates, Finally, we use a moment-generating function and the Gaussian integral to construct the parameterized form of the Gaussian or *normal* probability density function.

1. CHARACTERIZING PROBABILITY DISTRIBUTIONS

Although we could start by presenting a Gaussian function and proceed by evaluating its integral over the real numbers, that would not provide a context for the exercise. The relevance of the procedure would be lost and the reader would be left with an isolated mathematical process devoid of applicability. Given that the root of Gaussian functions lies in probability theory, where a specific instance defines the so-called *normal distribution*, we will review the necessary statistical principles to understand the utility of the Gaussian integral.

1.1. **Moments.** In probability theory, a parameter associated with a probability distribution, such as its mean or variance, is said to be a *characteristic* of the distribution. A probability distribution is a function that models a so-called *population* of data points. A population is the hypothetical collection of all possible measurements made under a given set of conditions. It is a theoretical set of data and not the actual set of measurements. To distinguish the characteristics of a sample from the parameters of a distribution, we use the term *statistic*. Therefore, the population mean is a *parameter* of the probability distribution and the sample mean is a *statistic* of the sample set.

To illustrate the difference between a parameter and a statistic, let's compare the formula for the expected value of a discrete random variable X with

---

Version: 1.1.1, 2022-02-21.

probability mass function $P(x)$,

$$E(X) = \mu_X = \sum_{n=1}^{N} x_n P(x_n), \tag{1}$$

to the formula for the expected value of a continuous random variable X with probability density function $P(x)$,

$$E(X) = \mu_X = \int_{-\infty}^{\infty} x P(x) \, dx. \tag{2}$$

Equation 1 represents a statistic and may have different values for different sample sets. Equation 2 represents a parameter and has a fixed value. Each is a characteristic of its respective discrete and continuous distribution.

A *moment* is an element of a set of characteristic values that collectively describe the distribution of a random variable, albeit not in a unique manner. Under the right conditions, two random variables with identical moments will have identical probability distributions. That makes it possible to approximate the probability distribution of a random variable using moments.

In the study of statistics, moments fall into two classes: *raw* and *central*. A *raw moment* is computed about the origin and a *central moment* is computed about the mean. Without proof, we state that if X is a discrete random variable with probability mass function $P(x)$ and $g(X)$ is a function of X, then

$$E(g(X)) = \mu_{g(X)} = \sum_{n=1}^{N} g(x_n) \cdot P(x_n); \tag{3}$$

and if X is a continuous random variable with probability density function $P(x)$ and $g(X)$ is a function of X, then

$$E(g(X)) = \mu_{g(X)} = \int_{-\infty}^{\infty} g(x) \cdot P(x) \, dx. \tag{4}$$

We set $g(X) = X^k$ to find the raw moments of a probability distribution.

The raw moments of a discrete probability distribution are defined as

$$E(X^k) = \mu_k' = \sum_{n=1}^{N} x_n^k P(x_n), \tag{5}$$

where $E(X^k)$ is the kth moment of discrete random variable X and written as μ_k'. Similarly, the raw moments of a continuous probability distribution for a continuous random variable X are defined as

$$E(X^k) = \mu_k' = \int_{-\infty}^{\infty} x^k P(x) \, dx. \tag{6}$$

The central moments are defined analogously, setting $g(X) = (X - \mu)^k$ to center the moments about the mean μ. The central moments of a discrete random variable are

$$E([X - \mu]^k) = \mu_k = \sum_{n=1}^{N} (x_n - \mu)^k P(x_n) \tag{7}$$

and for a continuous random variable they are

$$E([X - \mu]^k) = \mu_k = \int_{-\infty}^{\infty} (x - \mu)^k P(x) \, dx. \tag{8}$$

Note that the first raw moment, μ_1', is equal to the mean, μ. The first central moment, μ_1, will always equal zero because Equation 8 reduces to $\mu - \mu = 0$ for $k = 1$. The area under the curve to the left of the mean will always be equal to that to the right for a proper probability density function. The second central moment, μ_2, is equal to the population variance. Knowing that the population variance, σ^2, of random variable X is $E(X^2) - E^2(X)$, the variance can be expressed in terms of raw moments as

$$\sigma^2 = \mu_2' - \mu_1'^2. \tag{9}$$

In addition, it is possible to define a set of moments about an arbitrary point, a, in which case you would replace μ in Equations 7 and 8 with the point a.

1.2. **Moment-Generating Functions.** Instead of calculating the moments of a probability distribution on an ad hoc basis, we can identify a function from which the moments can be immediately derived. This so-called *moment-generating* function exploits the series expansion of e^{tx} to create a function that, when successively differentiated at zero, produces the raw moment of a probability distribution corresponding to the level of differentiation. For example, the first derivative at zero produces the first raw moment, the second derivative the second raw moment, and so on. This works by applying Equation 4 to find the expected value of e^{tX}, using $g(X) = e^{tX}$ and integrating to yield a moment-generating function,

$$M(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} P(x) \, dx. \tag{10}$$

Expanding e^{tx} to its infinite series gives us

$$
\begin{aligned}
M(t) &= \int_{-\infty}^{\infty} \left(\sum_{n=0}^{\infty} \frac{t^n x^n}{n!} \right) P(x)\, \mathrm{d}x; \\
&= \int_{-\infty}^{\infty} \left(1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \cdots \right) P(x)\, \mathrm{d}x; \\
&= \int_{-\infty}^{\infty} P(x)\, \mathrm{d}x + t \int_{-\infty}^{\infty} x P(x)\, \mathrm{d}x + \frac{t^2}{2!} \int_{-\infty}^{\infty} x^2 P(x)\, \mathrm{d}x + \cdots; \\
&= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots .
\end{aligned}
\tag{11}
$$

We arrive at Equation 11 by substituting terms using Equation 6.

If we differentiate $M(t)$ with respect to t we get

$$
M'(t) = \frac{\mathrm{d}M}{\mathrm{d}t} = \mu_1' + \frac{2t}{2!}\mu_2' + \frac{3t^2}{3!}\mu_3' + \cdots .
$$

Evaluating $M'(t)$ at $t = 0$ eliminates all but the first term, yielding

$$
M'(0) = \left. \frac{\mathrm{d}M}{\mathrm{d}t} \right|_{t=0} = \mu_1'.
$$

Taking the second derivative produces

$$
M''(t) = \frac{\mathrm{d}^2 M}{\mathrm{d}t^2} = \frac{2}{2!}\mu_2' + \frac{6t}{3!}\mu_3' + \frac{12t^2}{4!} + \cdots;
$$

and evaluating it at $t = 0$ results in

$$
M''(0) = \left. \frac{\mathrm{d}^2 M}{\mathrm{d}t^2} \right|_{t=0} = \mu_2'.
$$

It should be apparent that the kth derivative evaluated at $t = 0$ produces the kth raw moment according to

$$
\left. \frac{\mathrm{d}^k}{\mathrm{d}t^k} M(t) \right|_{t=0} = \mu_k' \quad \text{where} \quad k > 0.
\tag{12}
$$

As with Equation 4, if X is a continuous random variable with probability density function $P(x)$ and $g(X)$ is a function of X, then

$$
M(t) = E(e^{tg(X)}) = \int_{-\infty}^{\infty} e^{tg(x)} P(x)\, \mathrm{d}x.
\tag{13}
$$

We can use this formulation to find the moment-generating function of $X - \mu$ by setting $g(X) = X - \mu$. This can be useful to derive the variance directly from the second moment instead of using Equation 9. However, we will only be using raw moments.

1.3. **Uniform Distribution Moments.** So that we can understand how to derive the moment-generating function of the Gaussian distribution in Section 4, we will now work through the example of deriving the moment-generating function of the uniform probability distribution,

$$P(x) = \begin{cases} \frac{1}{b-a} & \text{for} \quad a \leq x \leq b, \\ 0 & \text{for} \quad x < a \quad \text{or} \quad x > b. \end{cases} \tag{14}$$

Applying Equation 10 to the unifrom probability density function from Equation 14, we have

$$M(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{b-a} \, dx = \frac{1}{b-a} \int_{a}^{b} e^{tx} \, dx.$$

Integrating gives us

$$M(t) = \frac{1}{b-a} \cdot \frac{e^{tx}}{t} \Big|_{a}^{b} = \frac{1}{b-a} \left(\frac{e^{bt} - e^{at}}{t} \right),$$

which produces the indeterminate form $\frac{0}{0}$ at $t = 0$, requiring us to apply l'Hôpital's rule as follows:

$$\lim_{t \to 0} \frac{1}{b-a} \left(\frac{e^{bt} - e^{at}}{t} \right) = \lim_{t \to 0} \frac{be^{bt} - ae^{at}}{b-a} = \frac{b-a}{b-a} = 1.$$

This gives us the complete moment-generating function for the uniform probability distribution,

$$M(t) = \begin{cases} \frac{e^{bt} - e^{at}}{(b-a)t} & \text{for} \quad t \neq 0, \\ 1 & \text{for} \quad t = 0. \end{cases} \tag{15}$$

We can now find the first raw moment of the uniform probability distribution by finding the first derivative with respect to t of its moment-generating function. Using the sum and product rules, we get

$$M'(t) = \frac{d}{dt} M(t) = \frac{1}{b-a} \left(\frac{d}{dt} \frac{e^{bt}}{t} - \frac{d}{dt} \frac{e^{at}}{t} \right);$$

$$= \frac{1}{b-a} \left(\frac{be^{bt}}{t} - \frac{e^{bt}}{t^2} - \frac{ae^{at}}{t} + \frac{e^{at}}{t^2} \right);$$

$$= \frac{1}{b-a} \left(\frac{be^{bt} - ae^{at}}{t} + \frac{e^{at} - e^{bt}}{t^2} \right).$$

The resulting function is indeterminate at $t = 0$, requiring us to rewrite it to produce a result of the form $\frac{0}{0}$, allowing us to apply l'Hôpital's rule as

follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}M(t)\bigg|_{t=0} = \lim_{t\to 0}\frac{1}{b-a}\left(\frac{bte^{bt} - ate^{at} + e^{at} - e^{bt}}{t^2}\right);$$

$$= \frac{1}{b-a}\lim_{t\to 0}\frac{b^2te^{bt} + be^{bt} - a^2te^{at} - ae^{at} + ae^{at} - be^{bt}}{2t};$$

$$= \frac{1}{b-a}\lim_{t\to 0}\frac{b^2te^{bt} - a^2te^{at}}{2t};$$

$$= \frac{1}{b-a}\lim_{t\to 0}\frac{b^2e^{bt} - a^2e^{at}}{2};$$

$$= \frac{1}{b-a}\cdot\frac{b^2 - a^2}{2} = \frac{b+a}{2}.$$

Swapping the terms from the last result, the first raw moment of the uniform probability distribution is

$$M'(0) = \mu'_1 = \frac{a+b}{2}. \tag{16}$$

We can verify the first raw moment is correct by using Equation 2 to find the mean,

$$\mu = \int_{-\infty}^{\infty} x\cdot\frac{1}{b-a}\,\mathrm{d}x;$$

$$= \frac{1}{b-a}\int_a^b x\,\mathrm{d}x;$$

$$= \frac{1}{b-a}\cdot\frac{x^2}{2}\bigg|_a^b;$$

$$= \frac{1}{b-a}\cdot\frac{b^2 - a^2}{2} = \frac{a+b}{2}.$$

As a final exercise, let's derive the second raw moment of the uniform probability distribution. Differentiating the first derivative from before, we find the second derivative is

$$M''(t) = \frac{1}{b-a}\left(\frac{b^2e^{bt} - a^2e^{at}}{t} - \frac{be^{bt} - ae^{at}}{t^2} + \frac{ae^{at} - be^{bt}}{t^2} - 2\left(\frac{e^{at} - e^{bt}}{t^3}\right)\right).$$

Again, we have a function that cannot be readily evaluated at $t = 0$. As before, we restructure it to allow the application of l'Hôpital's rule, giving

us

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}M(t)\bigg|_{t=0} = \frac{1}{b-a}\lim_{t\to 0}\left(\frac{t^2(b^2\mathrm{e}^{bt} - a^2\mathrm{e}^{at})}{t^3} - \frac{t(b\mathrm{e}^{bt} - a\mathrm{e}^{at})}{t^3}\right.$$

$$\left. + \frac{t(a\mathrm{e}^{at} - b\mathrm{e}^{bt})}{t^3} - \frac{2(\mathrm{e}^{at} - \mathrm{e}^{bt})}{t^3}\right);$$

$$= \frac{1}{b-a}\lim_{t\to 0}\frac{t^2(b^2\mathrm{e}^{bt} - a^2\mathrm{e}^{at}) + 2t(a\mathrm{e}^{at} - b\mathrm{e}^{bt}) + 2(\mathrm{e}^{bt} - \mathrm{e}^{at})}{t^3};$$

$$= \frac{1}{b-a}\lim_{t\to 0}\left(\frac{t^2(b^3\mathrm{e}^{bt} - a^3\mathrm{e}^{at})}{3t^2}\right.$$

$$+ \frac{2t(b^2\mathrm{e}^{bt} - a^2\mathrm{e}^{at}) + 2t(a^2\mathrm{e}^{at} - b^2\mathrm{e}^{bt})}{3t^2}$$

$$\left. + \frac{2(a\mathrm{e}^{at} - b\mathrm{e}^{bt}) + 2(b\mathrm{e}^{bt} - a\mathrm{e}^{at})}{3t^2}\right);$$

$$= \frac{1}{b-a}\lim_{t\to 0}\frac{(b^3\mathrm{e}^{bt} - a^3\mathrm{e}^{at})}{3};$$

$$= \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}.$$

Using Equation 9, we can formulate the variance of the uniform probability distribution as

$$\sigma^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2;$$

$$= \frac{4a^2 + 4ab + 4b^2 - 3a^2 - 6ab - 3b^2}{12};$$

$$= \frac{a^2 - 2ab + b^2}{12} = \frac{b^2 - 2ab + a^2}{12};$$

$$= \frac{(b-a)^2}{12}. \tag{17}$$

2. Gaussian Functions

Mathematics literature uses the term *Gaussian function* either narrowly—using it exclusively to refer to the Gaussian probability density function—or broadly—using it to refer to a class of exponential functions. We will use the term in its broad sense, starting with the simplest instance, $f(x) = \mathrm{e}^{-x^2}$, before examining more complex forms. Unfortunately, the author has failed to ascertain the historical origin of Gaussian functions and how their earliest applications eventually led to the identification of the Gaussian probability

density function. Therefore, we will present a pedagogical analysis of how starting from $f(x) = e^{-x^2}$ eventually leads to the Gaussian probability density function.

Generally speaking, a Gaussian function is a function whose natural logarithm is a concave quadratic function. A quadratic function is concave if its second derivative is negative. In essence, it is a downward-growing parabola as in Figure 1. Therefore, we can define a Gaussian function as
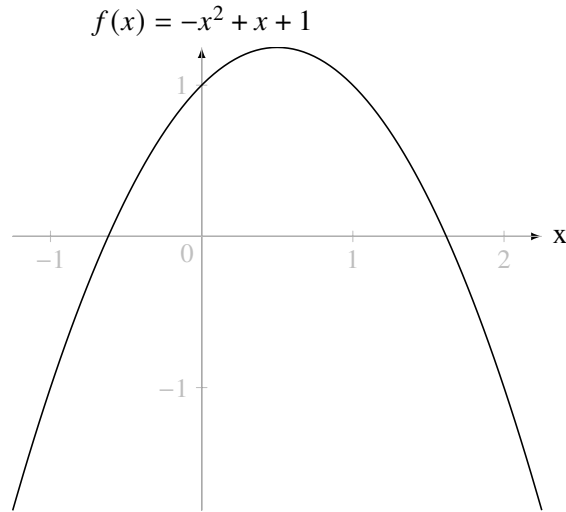
$$f(x) = -x^2 + x + 1$$

FIGURE 1. A concave quadratic function.

having the form $f(x) = ae^{-bx^2+cx+d}$, where b is positive in order to ensure the second derivative of the quadratic function is negative. To simplify our discussion, we restrict our scrutiny to quadratic functions of the form $-a(x-b)^2$, giving $f(x) = ae^{-b(x-c)^2}$ as our model Gaussian function. The simplest Gaussian function we can construct sets $a = 1$, $b = 1$, and $c = 0$, leaving us with $f(x) = e^{-x^2}$. We plot this function in Figure 2 and show how it varies by changing the function parameters.

Starting with $f(x) = e^{-x^2}$, we can visually note it is an *even function*, meaning that $f(x) = f(-x)$. This property means the function is symmetric about the y-axis. More generally, all Gaussian functions are symmetric about their midpoints, with the area under the curve to the left of the midpoint being equal to the area under the curve to the right of the midpoint. We will exploit this property in Section 3 to compute half of the Gaussian integral.

The midpoint of the function can be shifted by changing the value of c, as done in the plot of $f(x) = e^{-(x-2)^2}$ in Figure 2. Notice that the midpoint is shifted by an amount equal to c, shifting in the positive x direction when
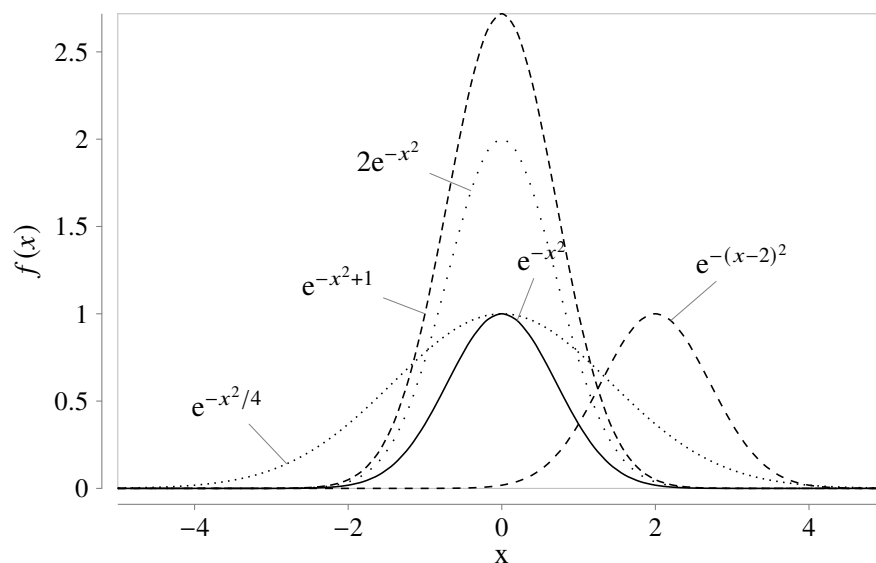
FIGURE 2. The Gaussian function, $f(x) = e^{-x^2}$, can be stretched, widened or narrowed, and shifted by changing a, b, and c in $f(x) = ae^{-b(x-c)^2}$.

c is positive and in the negative x direction when c is negative. Changing the value of b expands or contracts the width of the curve. A value greater than one narrows the curve and a value less than one and greater than zero widens the curve, as in the plot of $f(x) = e^{-x^2/4}$. Notice how the base of the curve expands by an amount equal to $\frac{1}{b}$. The function can be stretched upward or downward by multiplying it by a constant, as in the plot of $f(x) = 2e^{-x^2}$ where we see that the maximum of the function changes by a factor of a. The same effect can be achieved by adding a constant k to the quadratic term, as in the plot of $f(x) = e^{-x^2+1}$, where $k = 1$. But that is equivalent to multiplication by e^k, which is one reason we chose not to use a general quadratic form for the exponent (another is that it simplifies integration by substitution as we will see in Section 4). It should be clear that $e \cdot e^{-x^2} = e^{-x^2+1}$ and $2e^{-x^2} = e^{\ln 2} \cdot e^{-x^2} = e^{-x^2+\ln 2}$, making an additive term redundant when parameterizing the function and studying the effects of changing parameter values. In Section 4, we will see that parameters a, b, and c posses special meanings with respect to the Gaussian probability distribution.

3. The Gaussian Integral

Recognizing that $f(x) = e^{-x^2}$ ranges in value from 0 to 1, suppose we wanted to use a Gaussian function as a probability density function. Our first step would be to ensure that the integral over $(-\infty, \infty)$ is equal to 1. We would have to integrate the function to determine the area, A, under the curve. If $A \neq 1$, we would normalize the function to $\frac{1}{A} f(x)$ so that its integral over $(-\infty, \infty)$ equaled 1. We will, in fact, do all of this while deriving the Gaussian probability density function. However, embarking on this journey requires evaluating the Gaussian integral, for which we must first take a brief detour.

Solving problems in mathematics often requires what could be called *tricks*; like when we restructured the derivatives in Section 1.3 so we could apply l'Hôpital's rule. The Gaussian integral can be solved in various ways, all of which require some trickery. We will be using two tricks, the first of which is based on the seemingly trivial identity $x = \sqrt{x^2}$, allowing us to observe that

$$\int_a^b f'(x)\, dx = \left(\int_a^b f'(x)\, dx \cdot \int_a^b f'(x)\, dx \right)^{\frac{1}{2}} \tag{18}$$

implies

$$f(b) - f(a) = \sqrt{(f(b) - f(a))^2}. \tag{19}$$

The second trick we will use is to convert from Cartesian to polar coordinates, for which the reader will have to refer to a calculus textbook if the process is unclear.

We wish to evaluate the improper integral

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx$$

to see if the integrand can be adapted for use as a probability density function. We cannot evaluate the integral based on an existing indefinite integral because the integrand has no elementary indefinite integral. To evaluate the improper integral will depend exactly on its improper nature, integrating from negative to positive infinity. First, we apply Equation 18 to rewrite the integral as

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx \cdot \int_{-\infty}^{\infty} e^{-x^2}\, dx \right)^{\frac{1}{2}}.$$

Next, we replace one of the x dummy variables with another dummy variable, y, allowing us to rewrite the integral as the following double integral:

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

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

For clarity, we'll set aside the square root until the end and evaluate the double integral by converting to polar coordinates as follows:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} \, dx \, dy = \int_{0}^{2\pi} \int_{0}^{\infty} e^{-r^2} \, dr \, d\theta;$$

$$= \int_{0}^{2\pi} -\frac{1}{2} e^{-r^2} \Big|_{0}^{\infty} \, d\theta;$$

$$= \int_{0}^{2\pi} \frac{1}{2} \, d\theta;$$

$$= \frac{1}{2} \theta \Big|_{0}^{2\pi};$$

$$= \pi.$$

Now we can apply the square root to arrive at our final result,

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi}. \tag{20}$$

The various transformations we performed were possible because $f(x) = e^{-x^2}$ is a continuous function. As noted in Section 2, $f(x) = f(-x)$ holds true for a Gaussian function. That allows us to infer that

$$\int_{0}^{\infty} e^{-x^2} \, dx = \frac{\sqrt{\pi}}{2}. \tag{21}$$

More formally, if $f(x)$ is continuous on the interval $(-\infty, \infty)$ containing point a then

$$\int_{-\infty}^{\infty} f(x) \, dx = \int_{-\infty}^{a} f(x) \, dx + \int_{a}^{\infty} f(x) \, dx.$$

This property of improper integrals allows us to write

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx = \int_{-\infty}^{0} e^{-x^2} \, dx + \int_{0}^{\infty} e^{-x^2} \, dx.$$

Observing that $f(x) = f(-x)$ gives us

$$\int_{-\infty}^{0} e^{-x^2} \, dx = \int_{0}^{\infty} e^{-x^2} \, dx;$$

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx = 2 \int_{0}^{\infty} e^{-x^2} \, dx;$$

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

$$\int_{0}^{\infty} e^{-x^2} \, dx = \frac{\sqrt{\pi}}{2}.$$

4. The Gaussian Probability Density Function

Having evaluated the most basic Gaussian integral as $\sqrt{\pi}$, we can start to build a probability density function derived from the simplest Gaussian function, $f(x) = e^{-x^2}$, normalizing it to $f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$ so that its integral over $(-\infty, \infty)$ evaluates to 1. To investigate whether this is a useful construction, we can derive its moment-generating function and first two raw moments. Using Equation 10, we have

$$M(t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2} \, dx;$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x^2 - tx)} \, dx;$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x^2 - tx)} \, dx \cdot \frac{e^{-t^2/4}}{e^{-t^2/4}};$$

$$= \frac{e^{t^2/4}}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x^2 - tx + t^2/4)} \, dx;$$

$$= \frac{e^{t^2/4}}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(x - t/2)^2} \, dx.$$

Using $u = x - \frac{t}{2}$ and $du = 1 \cdot dx$, we find

$$M(t) = \frac{e^{t^2/4}}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2} \, du;$$

$$= \frac{e^{t^2/4}}{\sqrt{\pi}} \cdot \sqrt{\pi};$$

$$= e^{t^2/4}.$$

Note that we used Equation 20 to evaluate $\int_{-\infty}^{\infty} e^{-(x-t/2)^2} \, dx$ via substitution.

We now use Equation 12 to find the first two raw moments. The first raw moment is

$$\frac{\mathrm{d}}{\mathrm{d}t}M(t)\bigg|_{t=0} = \mu_1' = \lim_{t \to 0} \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{t^2/4} = \lim_{t \to 0} \frac{t}{2}\mathrm{e}^{t^2/4} = 0.$$

A mean of 0 is exactly what we would expect for a probability density function centered about the origin. The second raw moment is

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}M(t)\bigg|_{t=0} = \mu_2' = \lim_{t \to 0} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{e}^{t^2/4} = \lim_{t \to 0} \left(\frac{1}{2}\mathrm{e}^{t^2/4} + \frac{t^2}{4}\mathrm{e}^{t^2/4}\right) = \frac{1}{2},$$

giving us a variance of

$$\sigma^2 = \mu_2' - \mu_1'^2 = \frac{1}{2} - 0 = \frac{1}{2}$$

after using Equation 9.

The values of the raw moments we found are constants that don't really help us toward formulating a general Gaussian probability density function. But they do give us clues about how such a function might be formulated. In Section 2, we saw how the center of a Gaussian function could be shifted by changing the value of the parameter c in $f(x) = a\mathrm{e}^{-b(x-c)^2}$. We now have a clue that c is the mean of the distribution because it lies at the midpoint about which the function has symmetric areas. The parameter b may be related in some way to the variance, given how it affects the width of the curve. The parameter a must be a normalizing constant. To test our theory, we can hold these parameters constant, with b being positive, and try to find a moment-generating function parameterized in terms of a, b, and c.

We already know that c does not affect the area under the curve. Therefore we will ignore it and evaluate the integral of $f(x) = a\mathrm{e}^{-bx^2}$ over $(-\infty, \infty)$ to find a value of a dependent on b that normalizes the function. That will enable us to find a moment-generating function for the function including c using substitution as we did earlier. Using our earlier work, we can readily evaluate the integral as follows:

$$\int_{-\infty}^{\infty} a\mathrm{e}^{-bx^2}\,\mathrm{d}x = a\int_{-\infty}^{\infty} \frac{\mathrm{e}^{-u^2}}{\sqrt{b}}\,\mathrm{d}u$$

where $u = \sqrt{b}x$ and $\mathrm{d}u = \sqrt{b}\,\mathrm{d}x$. Substituting Equation 20, we get

$$\frac{a}{\sqrt{b}}\int_{-\infty}^{\infty} \mathrm{e}^{-u^2}\,\mathrm{d}u = \frac{a}{\sqrt{b}}\cdot\sqrt{\pi} = a\sqrt{\frac{\pi}{b}}.$$

Solving for a in

$$a\sqrt{\frac{\pi}{b}} = 1$$

to normalize the function leaves us with

$$a = \sqrt{\frac{b}{\pi}}. \tag{22}$$

Now we can use $f(x) = \sqrt{\frac{b}{\pi}}e^{-b(x-c)^2}$ as a probability density function and find its moment-generating function. Starting with

$$M(t) = \int_{-\infty}^{\infty} e^{tx} \cdot \sqrt{\frac{b}{\pi}}e^{-b(x-c)^2} \, dx;$$

$$= \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-b(x-c)^2} \, dx;$$

$$= \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} e^{-(bx^2-(2bc+t)x+bc^2)} \, dx;$$

$$= \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} e^{-(bx^2-(2bc+t)x+bc^2)} \, dx \cdot \frac{e^{-(ct+t^2/4b)}}{e^{-(ct+t^2/4b)}};$$

$$= e^{ct+t^2/4b} \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} e^{-(bx^2-(2bc+t)x+bc^2+ct+t^2/4b)} \, dx;$$

$$= e^{ct+t^2/4b} \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} e^{-\left(bx^2-(2bc+t)x+\left(\frac{bc+t/2}{\sqrt{b}}\right)^2\right)} \, dx;$$

$$= e^{ct+t^2/4b} \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} e^{-\left(\sqrt{b}x-\left(\frac{bc+t/2}{\sqrt{b}}\right)\right)^2} \, dx;$$

we can substitute $u = \sqrt{b}x - \left(\frac{bc+t/2}{\sqrt{b}}\right)$ and $du = \sqrt{b}\, dx$ to get

$$M(t) = e^{ct+t^2/4b} \sqrt{\frac{b}{\pi}} \int_{-\infty}^{\infty} \frac{e^{-u^2}}{\sqrt{b}} \, du;$$

$$= e^{ct+t^2/4b} \sqrt{\frac{b}{\pi}} \cdot \sqrt{\frac{\pi}{b}} = e^{ct+t^2/4b}.$$

The first moment is

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}M(t)\right|_{t=0} = \mu = \lim_{t\to 0}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{ct+t^2/4b};$$

$$= \lim_{t\to 0}\left((c + \frac{t}{2b})\cdot \mathrm{e}^{ct+t^2/4b}\right);$$

$$= c.$$

As we suspected, c is equal to the mean, μ. The second moment is

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}t^2}M(t)\right|_{t=0} = \lim_{t\to 0}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{e}^{ct+t^2/4b};$$

$$= \lim_{t\to 0}\frac{\mathrm{d}}{\mathrm{d}t}\left((c + \frac{t}{2b})\cdot \mathrm{e}^{ct+t^2/4b}\right);$$

$$= \lim_{t\to 0}\frac{\mathrm{d}}{\mathrm{d}t}(c\mathrm{e}^{ct+t^2/4b}) + \lim_{t\to 0}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t}{2b}\mathrm{e}^{ct+t^2/4b}\right);$$

$$= \lim_{t\to 0}\left(c(c + \frac{t}{2b})\cdot \mathrm{e}^{ct+t^2/4b}\right) + \lim_{t\to 0}\left(\frac{\mathrm{e}^{ct+t^2/4b}}{2b} + \left(\frac{t}{2b}(c + \frac{t}{2b})\cdot \mathrm{e}^{ct+t^2/4b}\right)\right);$$

$$= c^2 + \frac{1}{2b}.$$

We can now find the variance using Equation 9 as

$$\sigma^2 = \mu_2' - \mu^2 = \mu_2' - c^2;$$

$$= c^2 + \frac{1}{2b} - c^2;$$

$$= \frac{1}{2b}.$$

Solving for b in terms of σ produces

$$b = \frac{1}{2\sigma^2}.$$

Substituting this value into Equation 22 gives us

$$a = \frac{1}{\sigma\sqrt{2\pi}}.$$

Replacing a, b, and c in $f(x) = a\mathrm{e}^{-b(x-c)^2}$ with their parameterized values produces the probability density function for what is known as the Gaussian or normal distribution,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{23}$$

Notice how Equation 23 is dependent on the mean, μ, and the standard deviation, σ. As mentioned previously, changing μ just moves the center of the distribution left or right. But σ controls both the width and the height of the distribution. When the distribution widens, it gets shorter. When the distribution narrows, it gets taller. The role of σ is to constrain the area under the curve to remain constant independent of the value of σ. Therefore, the Gaussian probability density function represents a family of functions of unit area.

When $\mu = 0$ and $\sigma = 1$, the distribution corresponding to the probability density function is called the *standard normal distribution*,

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{24}$$

The standard normal distribution's probability density function is used primarily as a reference function for numerically integrating probabilities that can be scaled to calculate probabilities for other normal distributions based on the value of σ. The probability within n standard deviations of the mean is the same for all normal distributions regardless of the value of σ. Therefore, calculating probabilities for the standard normal distribution allows you to determine them for any normal distribution as long as you express your ranges in relative terms of standard deviations from the mean instead of using absolute numbers.

5. Commentary

The Gaussian probability density function is usually presented as a formula to be used, but not ncessarily understood. Although we attempted to show a step-by-step process from which one can get from $f(x) = e^{-x^2}$ to Equation 23, we did not explain the origin of $f(x) = e^{-x^2}$. Also, we cheated and chose $-b(x - c)^2$ as an exponent instead of $-bx^2 + cx + d$ with little explanation. At this point, it should be apparent that we needed the exponent to be in the form of a square paralleling $-x^2$ so that we could integrate by substitution.

Ultimately, our goal was to show that the Gaussian probability density function did not sprout out of thin air fully formed. Instead, it evolved from a series of observations about the family of functions derived from $f(x) = e^{-x^2}$. The Gaussian integral formed the core of our exercise because we only needed to evaluate it once and were subsequently able to use it multiple times via substitution, demonstrating its utility.

The Gaussian probability distribution occurs frequently in many contexts as a result of the central limit theorem.

The Central Limit Theorem. *Given a random sample $X_1, X_2, \ldots, X_N$ of independent and identically distributed random variables from a distribution with finite mean μ and finite variance σ^2, let*

$$D_N = \frac{\sum_{n=1}^{N} X_n - N\mu}{\sigma \sqrt{N}} = \frac{\sqrt{N}}{\sigma}(\bar{X} - \mu) \quad where \quad \bar{X} = \frac{1}{N}\sum_{n=1}^{N} X_n.$$

Then

$$\lim_{N \to \infty} P(D_N \leq d) = \int_{-\infty}^{d} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \quad for\ all\ d.$$

In other words, the distribution function D_N converges to the standard normal distribution from Equation 24 as $N \to \infty$.

For large N—usually greater than 30—one can pretend that $\bar{X}$ is distributed according to the normal distribution when calculating probabilities such as $P(a \leq \bar{X} \leq b)$. This should not be done recklessly, but does work in many situations. We will not prove the central limit theorem, but familiarity with it goes a long way to understanding why it is so common for textbooks to assume a Gaussian distribution in various contexts.

As a final comment, we explored Gaussian functions and the Gaussian probability distribution function in one dimension. The functions can can be extended to multiple dimensions, where analysis becomes considerably more complicated than what we have explored.

6. Addendum

For completeness, we have decided to demonstrate how to evaluate the integral of $f(x) = a e^{-bx^2 + cx + d}$ over the real numbers and summarize the integrals we covered in this paper for easy reference.

6.1. **General Gaussian Integral.** After evaluating the integral of the exponential function of a general concave quadratic function, it should be possible to evaluate the integral of any Gaussian function by simple parameter substitution. As before, the integral will depend on already having arrived at the result of Equation 20 and completing the square to allow integration by

substitution as follows:

$$\int_{-\infty}^{\infty} ae^{-bx^2+cx+d} \, dx = a \int_{-\infty}^{\infty} e^{-(bx^2-cx-d)} \, dx;$$

$$= a \int_{-\infty}^{\infty} e^{-(bx^2-cx-d)} \, dx \cdot \frac{e^{c^2/4b}}{e^{c^2/4b}};$$

$$= a \int_{-\infty}^{\infty} e^{-(bx^2-cx+c^2/4b)} \, dx \cdot e^{\frac{c^2}{4b}+d};$$

$$= ae^{\frac{c^2}{4b}+d} \int_{-\infty}^{\infty} e^{-(\sqrt{b}x-\frac{c}{2\sqrt{b}})^2} \, dx;$$

substituting $u = \sqrt{b}x - \frac{c}{2\sqrt{b}}$ and $du = \sqrt{b} \, dx$ to get

$$= ae^{\frac{c^2}{4b}+d} \int_{-\infty}^{\infty} \frac{e^{-u^2}}{\sqrt{b}} \, du;$$

$$= a\sqrt{\frac{\pi}{b}} e^{\frac{c^2}{4b}+d}.$$

6.2. **Summary of Gaussian Integrals.** We now summarize the results of evaluating the Gaussian integrals in this paper:

$$\int_{-\infty}^{\infty} e^{-x^2} \, dx = \sqrt{\pi};$$

$$\int_{0}^{\infty} e^{-x^2} \, dx = \frac{\sqrt{\pi}}{2};$$

$$\int_{-\infty}^{\infty} ae^{-bx^2} \, dx = a\sqrt{\frac{\pi}{b}};$$

$$\int_{-\infty}^{\infty} ae^{-b(x-c)^2} \, dx = a\sqrt{\frac{\pi}{b}};$$

$$\int_{-\infty}^{\infty} ae^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx = a\sigma\sqrt{2\pi};$$

and

$$\int_{-\infty}^{\infty} ae^{-bx^2+cx+d} \, dx = a\sqrt{\frac{\pi}{b}} e^{\frac{c^2}{4b}+d}.$$

All you really need is the final identity, from which all the others can be derived by substituting the appropriate values for the equation's parameters. Note, however, that we needed to evaluate the first integral in order to evaluate all of the others.