

Notes on the Normal Distribution, Moment Generating Functions, and Linear Transformations: A Guided Exploration

Synthesized from a Dialogue

April 17, 2025

Contents

1	Introduction: Charting a Course	2
2	The Normal Distribution	2
3	The Moment Generating Function (MGF)	2
3.1	Motivation and Definition	2
3.2	Deriving the MGF for the Normal Distribution	3
4	Linear Transformations of Normal Variables	5
4.1	Proof via PDF Transformation	5
4.2	Proof via MGF Method	6
5	Summary and Reflections	7
A	The Gaussian Integral	8

1 Introduction: Charting a Course

This document encapsulates a journey through fundamental concepts in probability theory, centered around the ubiquitous Normal (or Gaussian) distribution. Stemming from a dynamic conversation exploring definitions, derivations, and proofs, these notes aim to provide a clear, logical, and insightful pathway for understanding:

- The definition of the Normal distribution's probability density function (PDF).
- The concept and utility of the Moment Generating Function (MGF).
- The step-by-step derivation of the MGF for a Normal random variable, including tackling the famous Gaussian integral.
- The effect of linear transformations ($Y = aX + b$) on Normal random variables, explored through two distinct proof techniques: direct PDF manipulation and the MGF method.

Our goal is not just to present results, but to illuminate the reasoning behind them, revealing the elegance and interconnectedness of these ideas, much like uncovering the threads in a beautiful mathematical tapestry.

2 The Normal Distribution

Perhaps the most important continuous probability distribution, the Normal distribution, serves as a model for countless natural phenomena and is foundational in statistics.

Definition 2.1 (Normal Distribution PDF). A continuous random variable X follows a Normal (or Gaussian) distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}^+$ (i.e., $\sigma^2 > 0$), denoted as $X \sim \mathcal{N}(\mu, \sigma^2)$, if its probability density function (PDF) $f_X(x)$ is given by:

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

The term $\frac{1}{\sqrt{2\pi\sigma^2}}$ is the normalization constant, ensuring that $\int_{-\infty}^{\infty} f_X(x) dx = 1$. The term $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ is often called the Gaussian kernel.

Remark 2.2. The parameters μ and σ^2 represent the distribution's center (mean, expected value) and spread (variance), respectively. The standard deviation is $\sigma = \sqrt{\sigma^2}$.

3 The Moment Generating Function (MGF)

3.1 Motivation and Definition

While the PDF completely describes a distribution, other functions can provide valuable insights and tools. The Moment Generating Function (MGF) is one such tool.

Why MGFs?

- **Characterization:** For many common distributions, the MGF uniquely determines the distribution. If two random variables have the same MGF (within a region around $t = 0$), they follow the same distribution.

- **Moments:** As the name suggests, the MGF can be used to easily calculate the moments of a distribution (like the mean $E[X]$, variance $E[X^2] - (E[X])^2$, etc.) through differentiation. Specifically, $E[X^n] = M_X^{(n)}(0)$, the n -th derivative evaluated at $t = 0$.
- **Transformations & Sums:** MGFs often simplify the process of finding the distribution of sums of independent random variables or transformations of random variables.

Definition 3.1 (Moment Generating Function). The Moment Generating Function (MGF) of a random variable X , denoted $M_X(t)$, is defined as:

$$M_X(t) = \mathbb{E} \left[e^{tX} \right] \quad (2)$$

provided this expectation exists for t in some open interval containing 0. For a continuous random variable with PDF $f_X(x)$, this is calculated as:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \quad (3)$$

Remark 3.2. During our conversation, an initial expression considered $E[e^{-tX}]$. While this is a valid expectation ($M_X(-t)$), the standard definition uses $E[e^{tX}]$. We proceeded using the standard definition.

3.2 Deriving the MGF for the Normal Distribution

Let's derive the MGF for $X \sim \mathcal{N}(\mu, \sigma^2)$. This is a classic calculation that involves a crucial algebraic technique.

Step 1: Set up the integral Using the definition (3) and the Normal PDF (1):

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \end{aligned}$$

Step 2: Combine exponents Using the property $e^A e^B = e^{A+B}$:

$$\begin{aligned} M_X(t) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp \left(tx - \frac{(x-\mu)^2}{2\sigma^2} \right) dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp \left(\frac{2\sigma^2 tx - (x-\mu)^2}{2\sigma^2} \right) dx \end{aligned}$$

Step 3: Simplify the exponent - The core challenge The key to solving this integral lies in manipulating the exponent's numerator to resemble the structure found in a Normal PDF's exponent, namely $-(y - \tilde{\mu})^2 / (2\tilde{\sigma}^2)$. This requires "completing the square" with respect to x . Let's focus on the numerator, $N = 2\sigma^2 tx - (x - \mu)^2$:

$$\begin{aligned} N &= 2\sigma^2 tx - (x^2 - 2\mu x + \mu^2) \\ &= -x^2 + 2\mu x + 2\sigma^2 tx - \mu^2 \\ &= -[x^2 - 2(\mu + \sigma^2 t)x] - \mu^2 \end{aligned}$$

To complete the square for $x^2 - 2Ax$, we need to add and subtract A^2 . Here, $A = \mu + \sigma^2 t$.

$$\begin{aligned}
N &= -[x^2 - 2(\mu + \sigma^2 t)x + (\mu + \sigma^2 t)^2 - (\mu + \sigma^2 t)^2] - \mu^2 \\
&= -[(x - (\mu + \sigma^2 t))^2 - (\mu + \sigma^2 t)^2] - \mu^2 \\
&= -(x - (\mu + \sigma^2 t))^2 + (\mu + \sigma^2 t)^2 - \mu^2 \\
&= -(x - (\mu + \sigma^2 t))^2 + (\mu^2 + 2\mu\sigma^2 t + \sigma^4 t^2) - \mu^2 \\
&= -(x - (\mu + \sigma^2 t))^2 + 2\mu\sigma^2 t + \sigma^4 t^2
\end{aligned}$$

Now, substitute this back into the full exponent $\frac{N}{2\sigma^2}$:

$$\begin{aligned}
\text{Exponent} &= \frac{-(x - (\mu + \sigma^2 t))^2 + 2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2} \\
&= -\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2} + \frac{2\mu\sigma^2 t}{2\sigma^2} + \frac{\sigma^4 t^2}{2\sigma^2} \\
&= -\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2} + \mu t + \frac{1}{2}\sigma^2 t^2
\end{aligned}$$

Step 4: Substitute the simplified exponent back into the integral

$$M_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2} + \mu t + \frac{1}{2}\sigma^2 t^2\right) dx$$

Step 5: Separate the exponential and factor out constants Using $e^{A+B} = e^A e^B$. The term $\exp(\mu t + \frac{1}{2}\sigma^2 t^2)$ does not depend on x and can be pulled out of the integral.

$$\begin{aligned}
M_X(t) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2}\right) \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) dx \\
&= \frac{e^{\mu t + \frac{1}{2}\sigma^2 t^2}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2}\right) dx
\end{aligned}$$

Step 6: Evaluate the remaining integral The remaining integral looks very familiar. Consider the PDF of a Normal distribution with mean $\tilde{\mu} = \mu + \sigma^2 t$ and variance $\tilde{\sigma}^2 = \sigma^2$:

$$f_{N(\tilde{\mu}, \sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2}\right)$$

We know that any PDF must integrate to 1 over its domain:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2}\right) dx = 1$$

Multiplying both sides by $\sqrt{2\pi\sigma^2}$, we get:

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(x - (\mu + \sigma^2 t))^2}{2\sigma^2}\right) dx = \sqrt{2\pi\sigma^2}$$

Remark 3.3 (The Gaussian Integral). The fundamental reason this integral evaluates as it does relies on the value of the standard Gaussian integral $\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$. This result itself is typically derived using a clever trick involving polar coordinates (see Appendix A). The integral we needed is essentially a scaled and shifted version of this fundamental integral.

Step 7: Final Result Substitute the value of the integral back into the expression for $M_X(t)$:

$$\begin{aligned} M_X(t) &= \frac{e^{\mu t + \frac{1}{2}\sigma^2 t^2}}{\sqrt{2\pi\sigma^2}} \left(\sqrt{2\pi\sigma^2} \right) \\ &= e^{\mu t + \frac{1}{2}\sigma^2 t^2} \end{aligned}$$

Theorem 3.4 (MGF of Normal Distribution). *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Its Moment Generating Function (MGF) is given by:*

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \quad (4)$$

This elegant result is a cornerstone for working with Normal distributions.

4 Linear Transformations of Normal Variables

A fundamental question arises: if we take a Normal random variable X and apply a linear transformation, $Y = aX + b$, what is the distribution of Y ? Intuitively, scaling and shifting might preserve the "shape" of the distribution. Let's prove this rigorously.

Theorem 4.1 (Linear Transformation of a Normal Variable). *Let $X \sim \mathcal{N}(\mu, \sigma^2)$, and let $a, b \in \mathbb{R}$ with $a \neq 0$. Then the random variable $Y = aX + b$ also follows a Normal distribution:*

$$Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2) \quad (5)$$

We explored two methods to prove this important theorem during our conversation.

4.1 Proof via PDF Transformation

This method directly computes the PDF of Y using the change of variables formula.

Step 1: The Change of Variables Formula For a continuous random variable X with PDF $f_X(x)$, and a strictly monotonic (invertible) transformation $Y = g(X)$, the PDF of Y is given by:

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

where $g^{-1}(y)$ is the inverse transformation (X expressed in terms of Y) and the derivative term is the absolute value of the Jacobian of the transformation.

Step 2: Apply to $Y = aX + b$

- The inverse transformation is $X = g^{-1}(Y) = \frac{Y-b}{a}$.
- The derivative is $\frac{dX}{dY} = \frac{1}{a}$.
- The absolute value of the Jacobian is $\left| \frac{1}{a} \right| = \frac{1}{|a|}$.

Substituting into the formula:

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right) \frac{1}{|a|}$$

Step 3: Substitute the Normal PDF f_X Using $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, we replace x with $\frac{y-b}{a}$:

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\frac{y-b}{a} - \mu\right)^2}{2\sigma^2}\right) \cdot \frac{1}{|a|}$$

Step 4: Algebraic Simplification to Reveal the Target Form Our goal is to show this expression matches the PDF of $\mathcal{N}(a\mu+b, a^2\sigma^2)$. We need to simplify the constant and the exponent.

- **Exponent Simplification:**

$$\begin{aligned}\left(\frac{y-b}{a} - \mu\right)^2 &= \left(\frac{y-b-a\mu}{a}\right)^2 = \left(\frac{y-(a\mu+b)}{a}\right)^2 \\ &= \frac{(y-(a\mu+b))^2}{a^2}\end{aligned}$$

Substituting this into the exponent $-\frac{(\dots)^2}{2\sigma^2}$:

$$-\frac{\frac{(y-(a\mu+b))^2}{a^2}}{2\sigma^2} = -\frac{(y-(a\mu+b))^2}{2a^2\sigma^2}$$

This matches the form $-\frac{(y-\tilde{\mu})^2}{2\tilde{\sigma}^2}$ with $\tilde{\mu} = a\mu + b$ and $\tilde{\sigma}^2 = a^2\sigma^2$.

- **Constant Factor Simplification:** We have the factor $\frac{1}{|a|\sqrt{2\pi\sigma^2}}$. Since $|a| = \sqrt{a^2}$ for $a \neq 0$:

$$\frac{1}{|a|\sqrt{2\pi\sigma^2}} = \frac{1}{\sqrt{a^2}\sqrt{2\pi\sigma^2}} = \frac{1}{\sqrt{2\pi(a^2\sigma^2)}}$$

This matches the form $\frac{1}{\sqrt{2\pi\tilde{\sigma}^2}}$ with $\tilde{\sigma}^2 = a^2\sigma^2$.

Step 5: Combine Simplified Parts Putting the simplified constant and exponent together:

$$f_Y(y) = \frac{1}{\sqrt{2\pi(a^2\sigma^2)}} \exp\left(-\frac{(y-(a\mu+b))^2}{2(a^2\sigma^2)}\right)$$

This is precisely the PDF of a Normal distribution with mean $a\mu + b$ and variance $a^2\sigma^2$.

Conclusion (PDF Method): We have proven via direct transformation of the PDF that $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

4.2 Proof via MGF Method

This alternative proof leverages the MGF we derived earlier and its uniqueness property. It is often algebraically simpler for linear transformations.

Step 1: The MGF Property for Linear Transformations We first establish a general property relating the MGF of $Y = aX + b$ to the MGF of X .

Proposition 4.2 (MGF of a Linear Transformation). *Let X be a random variable with MGF $M_X(t)$, and let $Y = aX + b$ for constants a, b . Then the MGF of Y is given by:*

$$M_Y(t) = e^{bt} M_X(at)$$

Proof. By definition of the MGF and properties of expectation:

$$\begin{aligned}M_Y(t) &= \mathbb{E}[e^{tY}] \\ &= \mathbb{E}[e^{t(aX+b)}] \\ &= \mathbb{E}[e^{atX+bt}] \\ &= \mathbb{E}[e^{atX} e^{bt}] \quad (\text{exponent rule}) \\ &= e^{bt} \mathbb{E}[e^{(at)X}] \quad (\text{since } e^{bt} \text{ is constant w.r.t. } X) \\ &= e^{bt} M_X(at) \quad (\text{by definition of } M_X(\cdot) \text{ evaluated at } (at))\end{aligned}$$

□

Step 2: Apply the Property to Normal X We know $X \sim \mathcal{N}(\mu, \sigma^2)$ has $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$. Let $Y = aX + b$. Using the property above:

$$\begin{aligned} M_Y(t) &= e^{bt} M_X(at) \\ &= e^{bt} \exp\left(\mu(at) + \frac{1}{2}\sigma^2(at)^2\right) \\ &= e^{bt} \exp\left(a\mu t + \frac{1}{2}a^2\sigma^2 t^2\right) \end{aligned}$$

Step 3: Combine Exponents

$$\begin{aligned} M_Y(t) &= \exp\left(bt + a\mu t + \frac{1}{2}a^2\sigma^2 t^2\right) \\ &= \exp\left((a\mu + b)t + \frac{1}{2}(a^2\sigma^2)t^2\right) \end{aligned}$$

Step 4: Recognize the Resulting MGF Compare the expression $M_Y(t) = e^{(a\mu+b)t + \frac{1}{2}(a^2\sigma^2)t^2}$ to the general form of the Normal MGF, $e^{\tilde{\mu}t + \frac{1}{2}\tilde{\sigma}^2 t^2}$. We see that $M_Y(t)$ is precisely the MGF of a Normal distribution with:

- Mean $\tilde{\mu} = a\mu + b$
- Variance $\tilde{\sigma}^2 = a^2\sigma^2$

Step 5: Conclusion (MGF Method) By the uniqueness property of MGFs (which states that if the MGF exists, it uniquely determines the distribution), since $M_Y(t)$ is the MGF of $\mathcal{N}(a\mu + b, a^2\sigma^2)$, we conclude that Y must follow this distribution:

$$Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

Remark 4.3. Both proof methods yield the same correct result. The MGF method often bypasses more complex algebraic manipulation required by the PDF method, especially for linear transformations or sums of independent variables, highlighting the power of the MGF approach.

5 Summary and Reflections

Through this exploration, we have delved into the definition of the Normal distribution and rigorously derived its Moment Generating Function, $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$. This derivation required the technique of completing the square and implicitly relied on the evaluation of the Gaussian integral.

We then investigated the behavior of Normal variables under linear transformations, $Y = aX + b$. We proved, using two distinct methods (PDF transformation and MGF manipulation), the fundamental result that Y remains Normally distributed, with parameters adjusted according to the transformation: $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. This property underscores the stability and predictability of the Normal distribution under common transformations. Comparing the two proof techniques highlighted the elegance and efficiency the MGF approach can offer, leveraging the unique relationship between a distribution and its MGF.

These concepts form a vital part of the foundation for further study in probability, statistics, and related fields where the Normal distribution plays a central role.

A The Gaussian Integral

A key result underpinning the normalization of the Normal PDF and the derivation of its MGF is the value of the standard Gaussian integral.

Lemma A.1 (Gaussian Integral). *The definite integral of the standard Gaussian function is:*

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

Proof. Consider the square of the integral:

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \end{aligned}$$

This is an integral over the entire xy -plane. Convert to polar coordinates: $x = r \cos \theta$, $y = r \sin \theta$. Then $x^2 + y^2 = r^2$ and the area element $dx dy = r dr d\theta$. The limits become $r \in [0, \infty)$ and $\theta \in [0, 2\pi)$.

$$\begin{aligned} I^2 &= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} (r dr d\theta) \\ &= \left(\int_0^{2\pi} d\theta \right) \left(\int_0^{\infty} e^{-r^2} r dr \right) \end{aligned}$$

The first integral is $\int_0^{2\pi} d\theta = [\theta]_0^{2\pi} = 2\pi$. For the second integral, let $u = r^2$, so $du = 2r dr \implies r dr = \frac{1}{2} du$. The limits remain $u \in [0, \infty)$.

$$\begin{aligned} \int_0^{\infty} e^{-r^2} r dr &= \int_0^{\infty} e^{-u} \left(\frac{1}{2} du \right) \\ &= \frac{1}{2} \int_0^{\infty} e^{-u} du \\ &= \frac{1}{2} [-e^{-u}]_0^{\infty} \\ &= \frac{1}{2} \left(\lim_{u \rightarrow \infty} (-e^{-u}) - (-e^0) \right) \\ &= \frac{1}{2} (0 - (-1)) = \frac{1}{2} \end{aligned}$$

Combining the results:

$$I^2 = (2\pi) \times \left(\frac{1}{2} \right) = \pi$$

Since the integrand e^{-x^2} is always positive, I must be positive. Therefore, $I = \sqrt{\pi}$. □