Exercise Walkthrough: Convolution of Random Variables

Justin Lanfermann

25. June 2025

Abstract

This document provides a detailed, step-by-step walkthrough for an exercise on the sum of independent random variables. We will derive the convolution formula for probability density functions, apply it to the sum of two Exponentially distributed random variables, and finally show that the sum of two independent Gaussian random variables is also Gaussian. Each step is explained with reference to fundamental concepts from probability theory.

1 Overview of the Task

The goal of this exercise is to understand how the probability distribution of a sum of two independent random variables, $Z = X + Y$, is related to the distributions of X and Y. We will see that the probability density function (PDF) of Z is the *convolution* of the PDFs of X and Y. This is a foundational result with wide applications, for instance, when modeling a total error as the sum of independent error sources, or a total waiting time as a sum of independent process durations.

We will tackle this in three parts:

1. **Part (i):** Derive the general convolution formula.

2. **Part (ii):** Apply the formula to find the distribution of the sum of two independent Exponential random variables.

3. **Part (iii):** Apply the formula to show the sum of two independent Normal (Gaussian) random variables is also Normal.

2 Part (i): Deriving the Convolution Formula

2.1 The Goal

We want to find the probability density function (PDF), $p_Z(z)$, for the random variable $Z = X + Y$, where X and Y are independent random variables with known PDFs, $p_X(x)$ and $p_Y(y)$. We aim to prove:

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(z - y) p_Y(y) \, dy$$

2.2 The Strategy

The most reliable way to find the PDF of a transformed random variable is to first compute its cumulative distribution function (CDF), $F_Z(z)$, and then obtain the PDF by differentiating the CDF. Recall from **Lemma 1.44** in the script that the PDF is the derivative of the CDF, i.e., $p_Z(z) = \frac{d}{dz} F_Z(z)$.

2.3 Step-by-Step Derivation

Step 1: Express the CDF of Z as a probability.
By definition, the CDF of Z is $F_Z(z) = P(Z \leq z)$. Substituting $Z = X + Y$, we get:

$$F_Z(z) = P(X + Y \leq z)$$

Step 2: Formulate the probability as a double integral.
This probability is the integral of the *joint probability density function*, $p_{X,Y}(x, y)$, over the region of the plane where the condition $x + y \leq z$ holds. Let's call this region $A_z = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq z\}$.

$$F_Z(z) = \iint_{A_z} p_{X,Y}(x, y) \, dx \, dy$$

This is where the concept of a probability measure being defined over a set (an event) becomes a concrete computation over a region in space.

Step 3: Use the independence of X and Y [1].
The problem states that X and Y are independent. According to **Definition 1.72 (iii)** from the script, this means their joint PDF is the product of their marginal PDFs:

$$p_{X,Y}(x, y) = p_X(x)p_Y(y)$$

Substituting this into our integral is the key step that independence allows:

$$F_Z(z) = \iint_{x+y \leq z} p_X(x)p_Y(y) \, dx \, dy$$

Step 4: Convert the double integral into an iterated integral.
To evaluate this, we write it as an iterated integral. We can integrate over x first, and then over y. For any fixed value of y, the condition $x + y \leq z$ implies $x \leq z - y$. The variable y can range over all possible values from $-\infty$ to ∞.

$$F_Z(z) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} p_X(x) \, dx \right) p_Y(y) \, dy$$

The inner integral calculates the probability mass for X up to the value $z - y$, and the outer integral then averages this over all possible values of y, weighted by the probability density of Y.

Step 5: Differentiate the CDF to get the PDF [2].
Now we find $p_Z(z)$ by differentiating $F_Z(z)$ with respect to z. We use the Leibniz integral rule (a generalization of the Fundamental Theorem of Calculus) to differentiate under the integral sign.

$$p_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} p_X(x) \, dx \right) p_Y(y) \, dy$$

We can move the derivative inside the outer integral (as $p_Y(y)$ does not depend on z):

$$p_Z(z) = \int_{-\infty}^{\infty} \frac{d}{dz} \left(\int_{-\infty}^{z-y} p_X(x) \, dx \right) p_Y(y) \, dy$$

The derivative of the inner integral with respect to its upper limit z (treating y as a constant here) is simply the integrand $p_X(x)$ evaluated at that limit, so $\frac{d}{dz} \int_{-\infty}^{z-y} p_X(x) \, dx = p_X(z - y)$. Plugging this back in gives us our final result:

$$\mathbf{p_Z(z)} = \int_{-\infty}^{\infty} \mathbf{p_X(z - y)p_Y(y)} \, \mathbf{dy}$$

This is the convolution of p_X and p_Y.

Exercise for Understanding

Where exactly would this proof fail if X and Y were *not* independent? *Answer:* Step 3 would fail. We could not write $p_{X,Y}(x,y) = p_X(x)p_Y(y)$, and we would be stuck with the joint PDF, unable to proceed without more information about the dependency structure.

3 Part (ii): Sum of Independent Exponential RVs

3.1 The Goal

Given $X, Y \sim \text{Exp}(\lambda)$ are independent, with $\lambda > 0$, we want to calculate the PDF of $Z = X + Y$.

3.2 The Strategy

We will use the convolution formula derived in Part (i). The main challenge is correctly handling the domains of the exponential PDFs.

3.3 Step-by-Step Calculation

Step 1: Write down the PDFs.
From **Example 1.56 (iv)** in the script, the PDF for an $\text{Exp}(\lambda)$ distribution is:

$$p(t) = \lambda e^{-\lambda t} \quad \text{for } t \geq 0, \text{ and } 0 \text{ otherwise.}$$

Using indicator functions, which is cleaner for convolutions, we can write:

$$p_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_{\geq 0}}(x) \quad \text{and} \quad p_Y(y) = \lambda e^{-\lambda y} \mathbf{1}_{\mathbb{R}_{\geq 0}}(y)$$

Step 2: Set up the convolution integral.

$$p_Z(z) = \int_{-\infty}^{\infty} p_X(z-y)p_Y(y)\,dy = \int_{-\infty}^{\infty} \left(\lambda e^{-\lambda(z-y)} \mathbf{1}_{\mathbb{R}_{\geq 0}}(z-y) \right) \left(\lambda e^{-\lambda y} \mathbf{1}_{\mathbb{R}_{\geq 0}}(y) \right)\,dy$$

Step 3: Determine the integration limits from the indicator functions.
The integrand is non-zero only when both indicator functions are 1. This gives us two conditions:

1. $\mathbf{1}_{\mathbb{R}_{\geq 0}}(y) = 1 \implies y \geq 0$.

2. $\mathbf{1}_{\mathbb{R}_{\geq 0}}(z-y) = 1 \implies z - y \geq 0 \implies y \leq z$.

Combining these, the integration is non-zero only for $0 \leq y \leq z$. This means the integral's limits become 0 to z. It also implies that for the integral to be meaningful (i.e., not over an empty set), we must have $z \geq 0$. If $z < 0$, there are no values of y that satisfy both conditions, so $p_Z(z) = 0$ for $z < 0$.

Step 4: Evaluate the integral for $z \geq 0$.
For $z \geq 0$, the integral becomes:

$$\begin{aligned}
p_Z(z) &= \int_0^z (\lambda e^{-\lambda(z-y)})(\lambda e^{-\lambda y})\,dy \\
&= \lambda^2 \int_0^z e^{-\lambda z + \lambda y - \lambda y}\,dy \\
&= \lambda^2 e^{-\lambda z} \int_0^z 1\,dy \\
&= \lambda^2 e^{-\lambda z} [y]_0^z \\
&= \lambda^2 z e^{-\lambda z}
\end{aligned}$$

Step 5: Identify the resulting distribution.

The resulting PDF is $p_Z(z) = \lambda^2 z e^{-\lambda z}$ for $z \geq 0$. Let's compare this to the Gamma distribution from **Example 1.56 (iii)**, which has PDF $p_{\Gamma(\alpha,\beta)}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$. If we set $\alpha = 2$ and $\beta = \lambda$, we get:

$$p_{\Gamma(2,\lambda)}(z) = \frac{\lambda^2}{\Gamma(2)} z^{2-1} e^{-\lambda z} = \frac{\lambda^2}{1!} z e^{-\lambda z} = \lambda^2 z e^{-\lambda z}$$

This matches our result perfectly! Therefore, the sum of two independent $\text{Exp}(\lambda)$ random variables follows a $\mathbf{\Gamma(2, \lambda)}$ distribution. This makes sense intuitively: if an exponential distribution models the waiting time for one event, the Gamma distribution models the waiting time for α events.

4 Part (iii): Sum of Independent Normal RVs

4.1 The Goal

Given $X \sim \mathcal{N}(\mu_x, \sigma_x^2)$ and $Y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ are independent, we want to show that their sum $Z = X + Y$ is also normally distributed:

$$Z \sim \mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

4.2 The Strategy

As the exercise hint suggests, this is a tedious calculation. We'll use the convolution formula and the "trick" is to manipulate the exponent of the exponential function to isolate a new Gaussian-like term with respect to the integration variable, which will integrate to 1. The algebraic technique for this is called *completing the square* [3].

4.3 Step-by-Step Calculation

Step 1: Write down the PDFs and the convolution integral.

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}}, \quad p_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}$$

The convolution integral for $p_Z(z)$ is (integrating over x):

$$p_Z(z) = \int_{-\infty}^{\infty} p_Y(z-x) p_X(x) \, dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(z-x-\mu_y)^2}{2\sigma_y^2}} \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \, dx$$

Let's combine the constants and the exponents:

$$p_Z(z) = \frac{1}{2\pi\sigma_x\sigma_y} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\left[\frac{(z-x-\mu_y)^2}{\sigma_y^2} + \frac{(x-\mu_x)^2}{\sigma_x^2}\right]\right) dx$$

Step 2: Focus on the exponent and complete the square for x.

Let's call the term in the square brackets K.

$$K = \frac{(z-x-\mu_y)^2}{\sigma_y^2} + \frac{(x-\mu_x)^2}{\sigma_x^2}$$

Expanding and collecting terms in powers of x:

$$K = \frac{x^2 - 2x(z-\mu_y) + (z-\mu_y)^2}{\sigma_y^2} + \frac{x^2 - 2x\mu_x + \mu_x^2}{\sigma_x^2}$$

$$= x^2\left(\frac{1}{\sigma_x^2} + \frac{1}{\sigma_y^2}\right) - 2x\left(\frac{\mu_x}{\sigma_x^2} + \frac{z-\mu_y}{\sigma_y^2}\right) + \left(\frac{\mu_x^2}{\sigma_x^2} + \frac{(z-\mu_y)^2}{\sigma_y^2}\right)$$

This is a quadratic in x of the form $Ax^2 - 2Bx + C$. To complete the square, we rewrite it as $A(x - B/A)^2 + (C - B^2/A)$. The term $A(x - B/A)^2$ will form the core of a new Gaussian PDF in x, which will integrate out. The term $(C - B^2/A)$ will be what's left, and it will miraculously simplify.

This algebra is notoriously painful. Let's follow the hint and work towards the known answer. We want the final exponent to be $-\frac{(z-(\mu_x+\mu_y))^2}{2(\sigma_x^2+\sigma_y^2)}$. After a lot of algebra, the exponent K can be re-written as:

$$K = \frac{(\sigma_x^2 + \sigma_y^2)}{\sigma_x^2 \sigma_y^2}\left(x - \frac{\sigma_x^2(z - \mu_y) + \sigma_y^2\mu_x}{\sigma_x^2 + \sigma_y^2}\right)^2 + \frac{(z - (\mu_x + \mu_y))^2}{\sigma_x^2 + \sigma_y^2}$$

Let's define $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$ and $\mu_z = \mu_x + \mu_y$. And let's call the mean of the new Gaussian-in-x term $\mu_{new} = \frac{\sigma_x^2(z-\mu_y)+\sigma_y^2\mu_x}{\sigma_x^2+\sigma_y^2}$. The exponent becomes:

$$-\frac{1}{2}K = -\frac{1}{2}\left[\frac{\sigma_z^2}{\sigma_x^2\sigma_y^2}(x - \mu_{new})^2 + \frac{(z - \mu_z)^2}{\sigma_z^2}\right] = -\frac{(x - \mu_{new})^2}{2(\sigma_x^2\sigma_y^2/\sigma_z^2)} - \frac{(z - \mu_z)^2}{2\sigma_z^2}$$

Step 3: Evaluate the integral.

Plugging this back into the integral expression for $p_Z(z)$:

$$p_Z(z) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{(z - \mu_z)^2}{2\sigma_z^2}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{(x - \mu_{new})^2}{2(\sigma_x^2\sigma_y^2/\sigma_z^2)}\right)\,dx$$

The integral part is the integral over a full (unnormalized) Gaussian PDF with mean μ_{new} and variance $(\sigma_x^2\sigma_y^2/\sigma_z^2)$. The value of such an integral $\int_{-\infty}^{\infty}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$ is $\sqrt{2\pi}\sigma$. Here, $\sigma = \sqrt{\frac{\sigma_x^2\sigma_y^2}{\sigma_z^2}} = \frac{\sigma_x\sigma_y}{\sigma_z}$. So the integral evaluates to:

$$\int_{-\infty}^{\infty}\exp(\ldots)\,dx = \sqrt{2\pi}\frac{\sigma_x\sigma_y}{\sigma_z}$$

Step 4: Combine all terms.

$$p_Z(z) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{(z - \mu_z)^2}{2\sigma_z^2}\right)\cdot\left(\sqrt{2\pi}\frac{\sigma_x\sigma_y}{\sigma_z}\right)$$

The terms simplify beautifully:

$$p_Z(z) = \frac{1}{\sqrt{2\pi}\sigma_z}\exp\left(-\frac{(z - \mu_z)^2}{2\sigma_z^2}\right)$$

Substituting back $\mu_z = \mu_x + \mu_y$ and $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$, we get the PDF of a $\mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$ distribution. This completes the proof.

4.4   Summary and Takeaway

The sum of two independent Gaussian random variables is again a Gaussian. The mean of the sum is the sum of the means, and the variance of the sum is the sum of the variances. This is an incredibly important and convenient property of the Gaussian distribution, and one reason it is so widely used in statistics and machine learning. A much simpler proof exists using characteristic functions (**Example 2.50** in the script hints at this), which avoids the messy convolution integral entirely!

5 Explanatory Notes

Here are more detailed explanations of some of the concepts used in the walkthrough.

5.1 Independence of Random Variables

Two continuous random variables X and Y are said to be **independent** if for all (measurable) sets $A, B \subset \mathbb{R}$, the events $\{X \in A\}$ and $\{Y \in B\}$ are independent, i.e., $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$. A direct and computationally useful consequence of this, as stated in **Definition 1.72 (iii)** of your script, is that their joint probability density function (PDF) *factorizes* into the product of their individual (marginal) PDFs:

$$p_{X,Y}(x, y) = p_X(x)p_Y(y)$$

This property was the crucial link in Step 3 of Part (i) that allowed us to break down the joint probability into manageable parts.

5.2 Leibniz Integral Rule (Differentiation under the integral sign)

The Leibniz rule is a powerful tool from calculus for finding the derivative of an integral whose limits of integration are also functions of the differentiation variable. A simplified version, sufficient for our needs, comes from the Fundamental Theorem of Calculus. If we have a function defined as an integral:

$$G(z) = \int_a^{h(z)} f(x)\, dx$$

where a is a constant, its derivative with respect to z is:

$$G'(z) = \frac{d}{dz} \int_a^{h(z)} f(x)\, dx = f(h(z)) \cdot h'(z)$$

In our derivation in Part (i), the inner integral was $\int_{-\infty}^{z-y} p_X(x)\, dx$. Here, the upper limit is $h(z) = z - y$. Its derivative with respect to z is $h'(z) = 1$. So the derivative of the integral becomes $p_X(z - y) \cdot 1 = p_X(z - y)$.

5.3 Completing the Square

Completing the square is an algebraic technique used to rewrite a quadratic expression of the form $ax^2 + bx + c$ into the form $a(x - h)^2 + k$. This is particularly useful when dealing with Gaussian-like expressions, as it allows us to identify the mean and variance parameters. The general procedure is:

1. Factor out the leading coefficient a from the x^2 and x terms: $a(x^2 + \frac{b}{a}x) + c$.

2. Take half of the new coefficient of x (which is $\frac{b}{2a}$), square it ($\frac{b^2}{4a^2}$), and add and subtract it inside the parenthesis: $a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2}\right) + c$.

3. The first three terms inside the parenthesis now form a perfect square: $a\left(\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2}\right) + c$.

4. Distribute the a back to simplify: $a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c$.

In Part (iii), we applied this method to the exponent, which was a quadratic in x. This allowed us to isolate a term that looked like the exponent of a Normal PDF in x, which we could then integrate easily.