

Convolution in Probability

Owen McClure

Spring 2026

1 Convolution in Probability

To gain a better foundation for understanding convolution in the context of image processing and computer vision, it is helpful to first discuss convolution in probability theory.

Given two independent random variables X and Y with means μ_X, μ_Y and standard deviations σ_X, σ_Y , their sum $S = X + Y$ is also a random variable with mean equal to the sum of the means and variance equal to the sum of the variances. Additionally, if the input variables are normal, their result is also normal.

The distribution of S is obtained by convolving the distributions of X and Y . For discrete random variables with probability mass functions p_X and p_Y , the distribution of the sum is given by the convolution of the two distributions [1]:

$$p_{X+Y}(s) = [p_X * p_Y](s) = \sum_x p_X(x) p_Y(s - x). \quad (1)$$

The continuous case is similar, but uses probability density functions and integrals instead of sums. If f and g are the probability density functions of two independent random variables X and Y , the density of their sum $S = X + Y$ is the convolution of f and g :

$$[f * g](s) = \int_{-\infty}^{\infty} f(x) g(s - x) dx. \quad (2)$$

Here $f(x)$ is the density function of X and g is the density of Y . The integral runs over all possible pairs of values x and y constrained to the given sum s . Letting x vary over all real numbers and setting the argument of g to $y = s - x$ ensures that the sum is constrained to s . The integrand therefore evaluates the joint contribution of the pair $(x, s - x)$ to the density at s [1].

This is the convolution theorem in probability: the distribution of the sum of two independent random variables is the convolution of their individual distributions [2]. The random variables must be independent; if they are not, the convolution operation does not apply in the same way.

A convolution can be interpreted as a moving average. One function is flipped and shifted across the other, and at each position the overlapping values are multiplied and summed. In the discrete case this operation appears as a finite sum, while in the continuous case it becomes an integral. Despite this difference, both formulations express the same underlying idea.

To illustrate the discrete case, consider the arrays

$$a = (1, 2, 3), \quad b = (4, 5, 6). \quad (3)$$

Reversing and sliding the second array across the first gives

$$(1, 2, 3) * (4, 5, 6) = (4, 13, 28, 27, 18). \quad (4)$$

More generally, the convolution of sequences a and b is

$$(a * b)_k = \sum_i a_i b_{k-i}. \quad (5)$$

References

- [1] Sanderson G. But what is a convolution? ; 2022. Available from: <https://www.3blue1brown.com/lessons/convolutions>.
- [2] Murphy KP. Probabilistic machine learning: An introduction. The MIT Press; 2022.