

2-2017

## Why Mixture of Probability Distributions

Andrzej Pownuk

*The University of Texas at El Paso*, [ampownuk@utep.edu](mailto:ampownuk@utep.edu)

Vladik Kreinovich

*The University of Texas at El Paso*, [vladik@utep.edu](mailto:vladik@utep.edu)

Follow this and additional works at: [https://scholarworks.utep.edu/cs\\_techrep](https://scholarworks.utep.edu/cs_techrep)



Part of the [Computer Sciences Commons](#), and the [Mathematics Commons](#)

Comments:

Technical Report: UTEP-CS-17-15

To appear in *International Journal of Intelligent Technologies and Applied Statistics IJITAS*

---

### Recommended Citation

Pownuk, Andrzej and Kreinovich, Vladik, "Why Mixture of Probability Distributions" (2017). *Departmental Technical Reports (CS)*. 1103.

[https://scholarworks.utep.edu/cs\\_techrep/1103](https://scholarworks.utep.edu/cs_techrep/1103)

This Article is brought to you for free and open access by the Computer Science at ScholarWorks@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of ScholarWorks@UTEP. For more information, please contact [lweber@utep.edu](mailto:lweber@utep.edu).

# Why Mixture of Probability Distributions?

Andrzej Pownuk and Vladik Kreinovich  
Computational Science Program  
University of Texas at El Paso  
El Paso, Texas 79968, USA  
ampownuk@utep.edu, vladik@utep.edu

## Abstract

If we have two random variables  $\xi_1$  and  $\xi_2$ , then we can form their *mixture* if we take  $\xi_1$  with some probability  $w$  and  $\xi_2$  with the remaining probability  $1 - w$ . The probability density function (pdf)  $\rho(x)$  of the mixture is a convex combination of the pdfs of the original variables:  $\rho(x) = w \cdot \rho_1(x) + (1 - w) \cdot \rho_2(x)$ . A natural question is: can we use other functions  $f(\rho_1, \rho_2)$  to combine the pdfs, i.e., to produce a new pdf  $\rho(x) = f(\rho_1(x), \rho_2(x))$ ? In this paper, we prove that the only combination operations that always lead to a pdf are the operations

$$f(\rho_1, \rho_2) = w \cdot \rho_1 + (1 - w) \cdot \rho_2$$

corresponding to mixture.

## 1 Formulation of the Problem

**What is mixture.** If we have two random variables  $\xi_1$  and  $\xi_2$ , then, for each probability  $w \in [0, 1]$ , we can form a *mixture*  $\xi$  of these variables by selecting  $\xi_1$  with probability  $w$  and  $\xi_2$  with the remaining probability  $1 - w$ ; see, e.g., [1].

In particular, if we know the probability density function (pdf)  $\rho_1(x)$  corresponding to the first random variable and the probability density function  $\rho_2(x)$  corresponding to the second random variable, then the probability density function  $\rho(x)$  corresponding to their mixture has the usual form

$$\rho(x) = w \cdot \rho_1(x) + (1 - w) \cdot \rho_2(x). \quad (1)$$

**A natural question.** A natural question is: are there other combination operations  $f(\rho_1, \rho_2)$  that always transform two probability distributions  $\rho_1(x)$  and  $\rho_2(x)$  into a new probability distribution

$$\rho(x) = f(\rho_1(x), \rho_2(x)). \quad (2)$$

**Our result.** Our result is that the only possible transformation (2) that always generates a probability distribution is the mixture (1), for which

$$f(\rho_1, \rho_2) = w \cdot \rho_1 + (1 - w) \cdot \rho_2 \quad (3)$$

for some  $w \in [0, 1]$ .

## 2 Main Result

**Definition 1.** We say that a function  $f(\rho_1, \rho_2)$  that maps pairs of non-negative real numbers into a non-negative real number is a probability combination operation if for every two probability density functions  $\rho_1(x)$  and  $\rho_2(x)$  defined on the same set  $X$ , the function  $\rho(x) = f(\rho_1(x), \rho_2(x))$  is also a probability density function, i.e.,  $\int \rho(x) dx = 1$ .

**Proposition.** A function  $f(\rho_1, \rho_2)$  is a probability combination operation if and only if it has the form  $f(\rho_1, \rho_2) = w \cdot \rho_1 + (1 - w) \cdot \rho_2$  for some  $w \in [0, 1]$ .

**Proof.**

1°. Let us first prove that  $f(0, 0) = 1$ .

Indeed, let us take  $X = \mathbb{R}$ , and the following pdfs:

- $\rho_1(x) = \rho_2(x) = 1$  for  $x \in [0, 1]$  and
- $\rho_1(x) = \rho_2(x) = 0$  for all other values  $x$ .

Then, the combined function  $\rho(x) = f(\rho_1(x), \rho_2(x))$  has the following form:

- $\rho(x) = f(1, 1)$  when  $x \in [0, 1]$  and
- $\rho(x) = f(0, 0)$  for  $x \notin [0, 1]$ .

Let us use the condition  $\int \rho(x) dx = 1$  to prove that  $f(0, 0) = 0$ .

We can prove it by contradiction. If we had  $f(0, 0) \neq 0$ , i.e., if we had  $f(0, 0) > 0$ , then we would have

$$\int \rho(x) dx = f(1, 1) \cdot 1 + f(0, 0) \cdot \infty = \infty \neq 1.$$

Thus, we should have  $f(0, 0) = 0$ .

2°. Let us now prove that  $f(0, \rho_2) = k_2 \cdot \rho_2$  for some  $k_2 \geq 0$ .

Let us take the following function  $\rho_1(x)$ :

- $\rho_1(x) = 1$  for  $x \in [-1, 0]$  and
- $\rho_1(x) = 0$  for all other  $x$ .

Let us now pick any number  $\rho_2 > 0$  and define the following pdf  $\rho_2(x)$ :

- $\rho_2(x) = \rho_2$  for  $x \in [0, 1/\rho_2]$  and
- $\rho_2(x) = 0$  for all other  $x$ .

In this case, the combined function  $\rho(x)$  has the following form:

- $\rho(x) = f(1, 0)$  for  $x \in [-1, 0]$ ;
- $\rho(x) = f(0, \rho_2)$  for  $x \in [0, 1/\rho_2]$ , and
- $\rho(x) = f(0, 0) = 0$  for all other  $x$ .

Thus, the condition  $\int \rho(x) dx = 1$  takes the form

$$f(1, 0) + f(0, \rho_2) \cdot (1/\rho_2) = 1,$$

hence  $f(0, \rho_2) \cdot (1/\rho_2) = 1 - f(1, 0)$  and therefore,  $f(0, \rho_2) = k_2 \cdot \rho_2$ , where we denoted  $k_2 \stackrel{\text{def}}{=} 1 - f(1, 0)$ .

3°. Similarly, we can prove that  $f(\rho_1, 0) = k_1 \cdot \rho_1$  for some  $k_1 \geq 0$ .

4°. Let us now prove that for all  $\rho_1$  and  $\rho_2$ , we have  $f(\rho_1, \rho_2) = k_1 \cdot \rho_1 + k_2 \cdot \rho_2$ .

We already know, from Parts 1, 2 and 3 of this proof, that the desired equality holds when one of the values  $\rho_i$  is equal to 0.

Let us now take any values  $\rho_1 > 0$  and  $\rho_2 > 0$ . Let us then pick a positive value  $\Delta \leq 1/\max(\rho_1, \rho_2)$  and define the following pdfs. The first pdf  $\rho_1(x)$  is defined by the following formulas:

- $\rho_1(x) = \rho_1$  for  $x \in [0, \Delta]$ ,
- $\rho_1(x) = 1$  for  $x \in [-(1 - \Delta \cdot \rho_1), 0]$ , and
- $\rho_1(x) = 0$  for all other  $x$ .

The second pdf  $\rho_2(x)$  is defined by the following formula:

- $\rho_2(x) = \rho_2$  for  $x \in [0, \Delta]$ ,
- $\rho_2(x) = 1$  for  $x \in [\Delta, \Delta + (1 - \Delta \cdot \rho_2)]$ , and
- $\rho_2(x) = 0$  for all other  $x$ .

Then, the combined function  $\rho(x) = f(\rho_1(x), \rho_2(x))$  has the following form

- $\rho(x) = f(1, 0) = k_1$  for  $x \in [-(1 - \Delta \cdot \rho_1), 0]$ ,
- $\rho(x) = f(\rho_1, \rho_2)$  for  $x \in [0, \Delta]$ ,
- $\rho(x) = f(0, 1) = k_2$  for  $x \in [\Delta, \Delta + (1 - \Delta \cdot \rho_2)]$ , and

- $\rho(x) = f(0,0) = 0$  for all other  $x$ .

For this combined function  $\rho(x)$ , the condition that  $\int \rho(x) dx = 1$  takes the form

$$k_1 \cdot (1 - \Delta \cdot \rho_1) + f(\rho_1, \rho_2) \cdot \Delta + k_2 \cdot (1 - \Delta \cdot \rho_2) = 1. \quad (4)$$

Let us now consider a different pair of pdfs,  $\rho'_1(x)$  and  $\rho'_2(x)$ . The first pdf  $\rho'_1(x)$  is defined by the following formulas:

- $\rho'_1(x) = 2\rho_1$  for  $x \in [0, \Delta/2]$ ,
- $\rho'_1(x) = 1$  for  $x \in [-(1 - \Delta \cdot \rho_1), 0]$ , and
- $\rho'_1(x) = 0$  for all other  $x$ .

The second pdf  $\rho'_2(x)$  is defined by the following formula:

- $\rho'_2(x) = 2\rho_2$  for  $x \in [\Delta/2, \Delta]$ ,
- $\rho'_2(x) = 1$  for  $x \in [\Delta, \Delta + (1 - \Delta \cdot \rho_2)]$ , and
- $\rho'_2(x) = 0$  for all other  $x$ .

Then, the combined function  $\rho'(x) = f(\rho'_1(x), \rho'_2(x))$  has the following form

- $\rho'(x) = f(1, 0) = k_1$  for  $x \in [-(1 - \Delta \cdot \rho_1), 0]$ ,
- $\rho'(x) = f(2\rho_1, 0) = k_1 \cdot (2\rho_1)$  for  $x \in [0, \Delta/2]$ ,
- $\rho'(x) = f(0, 2\rho_2) = k_2 \cdot (2\rho_2)$  for  $x \in [\Delta/2, \Delta]$ ,
- $\rho'(x) = f(0, 1) = k_2$  for  $x \in [\Delta, \Delta + (1 - \Delta \cdot \rho_2)]$ , and
- $\rho'(x) = f(0, 0) = 0$  for all other  $x$ .

For this combined function  $\rho'(x)$ , the condition that  $\int \rho'(x) dx = 1$  takes the form

$$k_1 \cdot (1 - \Delta \cdot \rho_1) + k_1 \cdot (2\rho_1) \cdot (\Delta/2) + k_2 \cdot (2\rho_2) \cdot (\Delta/2) + k_2 \cdot (1 - \Delta \cdot \rho_2) = 1. \quad (5)$$

If we subtract (5) from (4) and divide the difference by  $\Delta > 0$ , then we conclude that  $f(\rho_1, \rho_2) - k_1 \cdot \rho_1 - k_2 \cdot \rho_2 = 0$ , i.e., exactly what we want to prove in this section.

5°. To complete the proof, we need to show that  $k_2 = 1 - k_1$ , i.e., that  $k_1 + k_2 = 1$ .

Indeed, let us take:

- $\rho_1(x) = \rho_2(x) = 1$  when  $x \in [0, 1]$  and
- $\rho_1(x) = \rho_2(x)$  for all other  $x$ .

Then, for the combined pdf, we have:

- $\rho(x) = f(\rho_1(x), \rho_2(x)) = k_1 + k_2$  for  $x \in [0, 1]$  and
- $\rho(x) = 0$  for all other  $x$ .

For this combined function  $\rho(x)$ , the condition  $\int \rho(x) dx = 1$  implies that

$$k_1 + k_2 = 1.$$

The proposition is proven.

## Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award “UTEP and Prudential Actuarial Science Academy and Pipeline Initiative” from Prudential Foundation.

## References

- [1] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.