

Conditional probability reasoning: A replication study

Alan Emmanuel Reyes-López ¹ 

Ana Luisa Gómez-Blancarte ² 

What problem is addressed?

The study of conditional probability is fundamental in probability and statistics as it allows the probability of an event to be updated based on new information (Díaz & de la Fuente, 2005). However, for more than thirty years, the specialized literature has documented a persistent variety of misconceptions in the understanding of this concept. Pioneering studies, such as that by Falk (1986), already addressed fundamental misconceptions: interpreting conditionality as causality, defining the conditioning event, and confusion of the inverse. Not only are these misconceptions still reported in recent studies, but others have been added that hinder the correct understanding of conditional probability.

The prevalence of misconceptions affects both students and teachers. The study by Díaz et al. (2012), conducted with future teachers, revealed a high incidence of misconceptions about conditional probability. These results confirm that future teachers have difficulties with these concepts. These findings suggest the need to strengthen their training, and raise awareness of these difficulties to prevent them from being passed on directly to their students.

¹ areyesl2400@alumno.ipn.mx

Instituto Politécnico Nacional, Ciudad de México, México

² algomezbl@ipn.mx

Instituto Politécnico Nacional, Ciudad de México, México

Identifying misconceptions provides a solid foundation for understanding the challenges in teaching and learning conditional probability. Therefore, this study focuses on replicating an instrument that helps identify misconceptions held by university students. The results from this instrument will allow teachers to better understand these students, enabling them to provide more conceptual explanations of conditional probability and to support their learning.

What is replicated?

This study is a conceptual and external replication because it applies an existing questionnaire to a different student population in order to identify misconceptions about conditional probability. The questionnaire on *Conditional Probability Reasoning* (CPR), originally developed by Díaz (2007), was adapted and administered. This instrument identifies misconceptions about conditional probability, as widely documented in the literature, through a questionnaire comprising 18 items (8 multiple-choice and 10 open-ended).

The instrument for the present replication study was sourced from Díaz et al. (2012), who administered the questionnaire to a sample of 196 students in Spain, comprising 95 last-year mathematics students and 101 future secondary education teachers, both groups with prior training in conditional probability. The current analysis is based on a subset of five multiple-choice items (1, 3, 4, 5, and 6). It is important to note that Díaz et al. (2012) employed modified versions of item 7 from the original questionnaire of Díaz (2007), and their item 2 also differs from the one used in this study. Consequently, a direct comparison for those specific items was not feasible.

The application was carried out electronically using the LimeSurvey platform, which facilitated access to a wider population and enabled efficient data collection. The process was carried out in August 2024, during which two teachers were trained to apply the instrument in their respective groups. The participants were third-semester engineering students (Mechatronics, Mechanical, Industrial, Renewable Energy, and Computer Science) with no prior university training in conditional probability, thus ensuring that their responses reflected initial knowledge not influenced by formal instruction. The questionnaire was completed by 123 students.

Results

Table 1 presents the percentage of misconceptions in each engineering student group in the present study, compared with those of Díaz et al. (2012). The data indicate that the time axis fallacy is the most common,

particularly among engineering students (79.7%), followed by the transposed conditional fallacy (70.7%). In item 1, engineering students show lower percentages than the other groups in the confusion between conditional probability and intersection (39.0%) and in the base rate fallacy (28.5%).

Table 1
Percentage of misconceptions by group

Misconception	Item number and options	Engineering students (123)	Mathematics students (95)	Future secondary education teachers (101)
Time axis fallacy	6 (a and d)	79.7%	67.4%	69.3%
Fallacy of the transposed conditional	5 (a and c)	70.7%	53.6%	58.4%
Confuse conditional probability and intersection	1 (c)	39.0%	58.9%	43.6%
Base rate fallacy	1 (a and b)	28.5%	31.6%	47.5%
Consider simple probability and intersection as equiprobable	4 (c)	45.5%	42.1%	17.8%
Confuse conditional probability and intersection	3 (c)	38.2%	26.3%	36.6%
Conjunction fallacy	4 (b)	12.2%	13.7%	9.9%
Confuse the conditioned event	6 (b)	12.2%	9.5%	5.9%

These results indicate a higher incidence of misconceptions among engineering students. The difference can be attributed to the fact that they had received no prior training in conditional probability. Despite this difference, it is noteworthy that the percentage of students who exhibited fundamental misconceptions remains remarkably consistent across all three groups. This observed similarity suggests that prior training in conditional probability may not be a significant factor in mitigating the persistence of misconceptions.

What is the importance and potential significance?

The relevance of the CPR questionnaire lies in its potential to detect misconceptions in the concept of conditional probability. The value of this instrument lies in its ability to generate evidence that enables the development of knowledge about the misconceptions held by university students in Mexico, knowledge necessary for effective teaching strategies and the design of specialized educational resources. For instance, this evidence could inform a targeted teacher training course, designed as an

implementation study. In this course, teachers would learn to identify misconceptions and design lessons that promote correct reasoning about conditional probability. Such a study would contribute to the field of probability and mathematics education.

References

- Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología* [Tesis doctoral, Universidad de Granada]. Repositorio Institucional de la Universidad de Granada. <https://digibug.ugr.es/handle/10481/1486>
- Díaz, C., Contreras, J. M., Batanero, C., & Roa, R. (2012). Evaluación de sesgos en el razonamiento sobre probabilidad condicional en futuros profesores de educación secundaria. *Bolema: Boletim de Educação Matemática*, 26(44), 1207–1225. <https://doi.org/10.1590/S0103-636X2012000400006>
- Díaz, C., & de la Fuente, I. (2005). Razonamiento sobre probabilidad condicional e implicaciones para la enseñanza de la estadística. *Epsilon*, 59, 245–260. https://www.ugr.es/~batanero/pages/ARTICULOS/epsilon_condicional.pdf
- Falk, R. (1986). Conditional probabilities: Insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 292–297). University of Victoria. <https://www.fpce.uc.pt/iase-web/documents/papers/icots2/Falk.pdf>