

Comparaciones múltiples y riesgo de falsos positivos

Diana María Mérida.

¹Fundación Teófilo Hernando, 28290, Las Rozas de Madrid, Madrid, España.

²Departamento de Medicina Preventiva y Salud Pública, Facultad de Medicina, Universidad Autónoma de Madrid, 28029, Madrid, España.

Introducción

Cuando realizamos múltiples comparaciones en un mismo conjunto de datos, aumenta la probabilidad de obtener un resultado estadísticamente significativo por azar, es decir, de cometer un error tipo I (obtener falsos positivos). El error tipo I representa el riesgo de detectar una diferencia significativa cuando en realidad no la hay^{1,2}.

En estadística, una comparación múltiple se refiere a la realización de varias pruebas de hipótesis dentro de un mismo análisis, como la comparación de más de dos grupos, la evaluación de diferentes momentos dentro del mismo estudio, o el análisis de múltiples variables al mismo tiempo. Cada comparación adicional aumenta la probabilidad de obtener al menos un resultado significativo por “pura casualidad”³.

Para reducir este riesgo en comparaciones múltiples, se utilizan métodos de ajuste estadístico, como la corrección de Bonferroni. La corrección de Bonferroni ajusta el nivel de significación dividiendo α (comúnmente establecido en 0,05) por el número de comparaciones realizadas. Esto se hace con el objetivo de limitar la probabilidad de obtener falsos positivos. Sin embargo, este ajuste puede ser demasiado estricto, lo que aumenta el riesgo de error tipo II (falsos negativos), ya que dificulta la detección de diferencias o efectos. Además, su aplicación es más adecuada cuando las pruebas son independientes, ya que, en presencia de correlaciones entre comparaciones, puede resultar en un ajuste excesivo⁴.

El uso de la corrección de Bonferroni sigue siendo objeto de debate. Generalmente, se prefiere en estudios confirmatorios, donde es crucial minimizar los falsos positivos. Sin embargo, en estudios exploratorios, donde el objetivo es identificar posibles tendencias sin descartar efectos reales, muchos investigadores optan por métodos más flexibles, como el método Holm-Bonferroni, Benjamini-Hochberg y el control de la tasa de descubrimiento falso (FDR), que permiten reducir la probabilidad de error tipo 1 sin perder poder estadístico⁵.

En definitiva, la elección del método de corrección depende del contexto del estudio y del número de comparaciones. La corrección de Bonferroni puede ser útil cuando se realizan pocas comparaciones y un falso positivo supone un error grave, pero en otros escenarios, métodos menos conservadores pueden ser más apropiados.

Bibliografía

1. Kim HY. Statistical notes for clinical researchers: Type I and type II errors in statistical decision. *Restor Dent Endod.* 2015;40(3):249. doi:10.5395/RDE.2015.40.3.249
2. Revista de Pediatría de Atención Primaria - El problema de las comparaciones múltiples. Accessed February 10, 2025. <https://pap.es/articulo/12072/el-problema-de-las-comparaciones-multiples>
3. Lee S, Lee DK. What is the proper way to apply the multiple comparison test? *Korean J Anesthesiol.* 2018;71(5):353. doi:10.4097/KJA.D.18.00242
4. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt.* 2014;34(5):502-508. doi:10.1111/OPO.12131
5. 6.1: Multiple Comparisons - Statistics LibreTexts. Accessed February 10, 2025. [https://stats.libretexts.org/Bookshelves/Applied_Statistics/Biological_Statistics_\(McDonald\)/06%3A_Multiple_Tests/6.01%3A_Multiple_Comparisons](https://stats.libretexts.org/Bookshelves/Applied_Statistics/Biological_Statistics_(McDonald)/06%3A_Multiple_Tests/6.01%3A_Multiple_Comparisons)