

## Ejercicio 3.

Keyner Martinez Miranda, Jairo Enrique Alba.

6/3/2023

### PUNTO 3. USO DEL VALOR-P COMO CRITERIO DE DECISIÓN

Con base en el artículo: **“Statistical Errors: P values, the gold standard of statistical validity, are not as reliable as many scientists assume”** Escriba un resumen (máximo 2 páginas) sobre el artículo e incluya en este sus opiniones en cuanto al uso del valor p como criterio de decisión en inferencia estadística.

En muchos campos de investigación y áreas del conocimiento se juzgan los resultados de comparación, predicción o correlación a través de los valores de p, el cual sirve para probar o rechazar la hipótesis nula, la cual postula la igualdad, la correlación o el ajuste de un modelo de tal forma que para valores de p pequeños se estima que sea más menos probable que los resultados sean debido a los errores muestrales o debidos al azar.

Los resultados del contraste de hipótesis se consideran “significativos estadísticamente,” cuando el valor de p es inferior a 0.05 (se rechaza  $H_0$ ), siempre que se trabaje con una significancia del 5%(confianza del 95%), es decir, la probabilidad de rechazar la hipótesis nula siendo verdadera, en el respectivo contraste de hipótesis que se este realizando.

En el artículo se plante al caso de Motyl, estudiante de doctorado en psicología de la Universidad de Virginia en Charlottesville. Quien con un estudio de casi 2000 personas parecían mostrar que los políticos moderados veían los tonos grises con mayor precisión que los extremistas de derecha o de izquierda. Su investigación inicial arrojó un P value de 0.01, generalmente interpretado como ‘muy significativo.’ Sin embargo, ante las críticas, decidieron repetir la medición con datos adicionales encontrando un P value de 0.59, ni siquiera cerca del nivel de significación convencional, 0.05. en ese momento todo se desvirtuó. Pero el problema no estaba en los datos sino en la naturaleza sorprendentemente resbaladiza del valor P, que no es ni tan fiable ni tan objetivo como suponen la mayoría de los científicos [1].

Un segundo caso se da en 2005, el epidemiólogo John Ioannidis de la Universidad de Stanford en California sugirió que la mayoría de los hallazgos publicados son falsos<sup>2</sup>; desde entonces, una serie de problemas de replicación de alto perfil ha obligado a los científicos a repensar cómo evalúan los resultados. Los valores de p siempre han tenido críticas. En sus casi nueve décadas de existencia. Existe amplia literatura alrededor de P, avalándola y una significativa advirtiendo de su inestabilidad. El Valor P es bastante susceptible los datos, por ejemplo, su tamaño, la existencia de datos atípicos, la aleatoriedad de la muestra, etc. por lo que estos factores podrían tener tal vez una incidencia en su resultado, es importante no solo los datos per se, sino contar también con el contexto de la información que se analiza [1].

Por lo tanto para garantizar la confiabilidad y validez de un trabajo de investigación además del p-valor debe estar apoyado estadísticamente con otras técnicas de investigación como los intervalos de confianza y el análisis de expertos en el tema, ya que se pueden presentar casos en los cuales, se puedan lograr establecer algún tipo de correlación (asociación) estadísticamente significativas, sin embargo pueden ser espurias o falsas debido a causas como el azar, presentación de sesgos provenientes de errores en la información, selección en las unidades observadas, falta de control en las variables extrañas, lo cual conduce a obtener asociaciones estadísticas cuando en realidad no existen.

Por lo tanto, el contraste de hipótesis y en particular el valor de  $p$ , debe ser una herramienta más y no la única para llegar a obtener el objetivo de investigación y su uso se debe hacer de manera responsable siempre apoyado en el uso de otras herramientas para la toma de decisiones y no sólo en los  $p$ -valores encontrados. Este enfoque y estas aclaraciones se deben tener muy en cuenta en el momento de enseñar estadística, para cambiar la cultura arraigada del  $p$ -valor como única medida a tener en cuenta en la toma de decisiones.

## Bibliografía

- [1] R. Nuzzo, “Statistical errors: P values, the ‘gold standard’ of statistical validity, are not as reliable as many scientists assume,” *Nature*, vol. 506, no. 7487, pp. 150–153, 2014.