**Interpretando correctamente Valores P**

El uso del *valor P* es ampliamente extendido en la ciencia. Investigadores los usan en todas partes, encontrándose valores P en *pruebas t*, *ANOVAs*, *análisis de regresión*, entre otras. Éste nos permite inferir si los resultados de una prueba de hipótesis son estadísticamente significativos. Esta probabilidad, altamente controversial, se ha vuelto tan importante en la ciencia moderna que ha tomado vida propia, determinado desde qué estudios se publican en revistas especializadas hasta que proyectos reciben financiación.

Pese a su amplio uso, los *valores P* son frecuentemente mal interpretados. Creyéndose que los resultados de estas pruebas permiten concluir de las siguientes formas:

* La probabilidad que la hipótesis nula sea cierta.
* La probabilidad de que la hipótesis alternativa sea cierta.
* La probabilidad de que los hallazgos sean reproducibles.
* Si los resultados son importantes o verdaderos.
* Si el experimento fue bien conducido.

La prevalencia de estas malas interpretaciones demuestra con claridad que el concepto del valor P es resbaladizo. Pero tranquilos, en esta entrada intentaré explicar de manera intuitiva y simple los valores P de manera que se eviten malas interpretaciones y problemas serios a la hora de inferir sobre los resultados de un experimento. Empecemos.

Bajo el marco lógico que se desarrolla el *valor P* uno de los conceptos centrales es el de **prueba de hipótesis**. Éste evalúa dos afirmaciones mutuamente excluyente acerca de una población en dirección a determinar cuál de ellas es mejor soportada por los datos muéstrales. Las dos afirmaciones en cuestión son llamadas **hipótesis nula** e **hipótesis alternativa**. En el caso de la hipótesis nula, el nombre  **nula**  es usado para recordarnos que estamos actuando como escépticos: damos crédito a la posibilidad de que no hayan diferencias entre los grupos o tratamientos.

Para comprender esta idea, imaginemos un estudio hipotético para un nuevo antibiótico el cual de partida sabemos que totalmente inefectivo, es decir **la hipótesis nula es verdadera**. No hay diferencias a nivel de poblacional entre los sujetos que tomaron el antibiótico y aquellos que no.

A pesar que la hipótesis nula es cierta, nosotros podríamos observar, con cierta probabilidad, algunas diferencias entre los sujetos que tomaron el antibiótico y aquellos que no. Este efecto estaría dado, principalmente, por error aleatorio en el muestreo más que un efecto real dado por la condición a la que fueron sometidos los sujetos. Por lo tanto, la conclusión lógica sería que el efecto observado es producido por el error de muestreo y no por el tratamiento.

**Entonces ¿Qué es el *valor P*?**

En breve, estos valores evalúan la consistencia de sus datos o estadísticas muéstrales con la hipótesis nula. Específicamente, si la hipótesis nula es correcta, ¿Cuál es la probabilidad de obtener un efecto al menos tan grande como el de su muestra?

* *Valores P* **altos**: Los datos muéstrales son consistente con una hipótesis nula que es verdadera.
* *Valores P* **bajos**: Los datos muéstrales **NO** son consistente con una hipótesis nula que es verdadera.

Si el valor P obtenido es lo sufrientemente pequeño, se puede concluir que los datos son incompatibles con la hipótesis nula, por la tanto se puede **rechazar** la hipótesis nula.

**Entonces.. ¿Cómo lo interpretamos correctamente?**

Volvamos al experimento hipotético del antibiótico. Supongamos que en una prueba de hipótesis obtuvimos un valor P de 0.01. La interpretación correcta de esta resultado sería:

*Si tomar el antibiótico no tiene efecto en la población, 1% de los estudios obtendrían el efecto observado en el experimento, o uno mayor, solo por error aleatorio en el muestreo.*

Así, la única pregunta que podemos resolver con el *valor P* es ¿Qué tan probable es ver sus resultados suponiendo que la hipótesis nula es **verdadera**?

**¿Qué son los niveles de significancia (**\alpha*α***)?**

Finalmente, para completar la interpretación hay que incluir el nivel de significancia dentro este marco lógico. Este valor es un estándar  probatorio que los investigadores fijan antes del estudio (típicamente se usa \alpha = 0.05*α*=0.05), indicándonos que tan fuerte debe ser la evidencia para contradecir la hipótesis nula antes de rechazarla para toda la población. En otras palabras, es la probabilidad de inferir que hay efecto cuando realmente no lo hay.

Con todo lo anterior en mente, los valores P y los niveles significancia se usan en conjunto de la siguiente manera:

* *Valores P* **altos** (P>0.05*P*>0.05): Indica evidencia débil contra la hipótesis nula, por lo tanto no se puede rechazar la hipótesis nula.
* *Valores P* **bajos** (P \leq 0.05*P*≤0.05): Indica una fuerte evidencia contra la hipótesis nula, por lo que rechazar la hipótesis nula es lo más apropiado.