

# PRUEBAS POST HOC

Una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc permiten determinar qué medias difieren. La prueba de rango post hoc identifica subconjuntos homogéneos de medias que no se diferencian entre sí.

## PRUEBA DE TUKEY

La prueba de Tukey, nombrado después Juan Tukey, es una prueba estadística utilizada general y conjuntamente con ANOVA, La prueba Tukey se usa en experimentos que implican un número elevado de comparaciones.

Es de fácil cálculo puesto que se define un solo comparador, resultante del producto del error estándar de la media por el valor tabular en la tabla de Tukey usando como numerador el número de tratamientos y como denominador los grados de libertad del error

Se conoce como Tukey-Kramer cuando las muestras no tienen el mismo número de datos

- Dado que el análisis de varianza acusa un efecto significativo, la prueba de Tukey provee un nivel de significancia global de  $\alpha$  cuando los tamaños de las muestras son iguales y de  $\alpha$  a lo sumo a cuando no son iguales.
- Se basa en la construcción de intervalos de confianza de las diferencias por pares. Si estos intervalos incluyen al 0, entonces no se rechaza la hipótesis nula

## PRUEBA DE SCHEFFÉ

La prueba de Scheffé es una prueba que se aplica para hacer comparaciones múltiples de las medias de grupos. Su uso está relacionado con la prueba del análisis de la varianza, y se incluye dentro de las llamadas pruebas de comparaciones múltiples.

La prueba del análisis de la varianza contrasta la hipótesis de igualdad de medias de dos o más grupos. Si el resultado se considera estadísticamente significativo, lo que se puede afirmar es que al menos la media de uno de los grupos es distinta a las restantes, o bien que hay otras medias diferentes entre sí.

El siguiente paso consiste en identificar qué grupos son los que tienen medias diferentes entre sí. Una solución es comparar las medias por pares, usando una prueba estadística como la t de Student. Pero al hacerlo así se produce un aumento del error tipo I que se quiere admitir. Las pruebas de comparaciones múltiples corrigen el error para conseguir que no sobrepase el nivel establecido, por ejemplo del 5%.

La prueba de Scheffé se realiza comparando todos los posibles pares de medias, pero usando como error típico el valor de la varianza residual o intragrupos obtenida en el análisis de la varianza.

## **MÉTODO DE LA DIFERENCIA MÍNIMA DE FISHER**

El método de la diferencia mínima de Fisher utiliza el estadístico F:

- Calcula una diferencia crítica a la que se comparan cada par de promedios
- Se pueden construir también intervalos de confianza para las diferencias por pares
- Sin embargo, el riesgo global  $\alpha$ , tiende a crecer con este método (este método utiliza tasas de error individuales).

## **PRUEBA DEL RANGO MÚLTIPLE (DUNCAN)**

La prueba del rango múltiple prueba las diferencias entre las medias empezando con la media más grande contra la segunda más grande, y así sucesivamente, comparando en cada caso con un valor crítico obtenido por tablas.

- Muy eficaz y popular debido a su poder de discriminación.
- Si el nivel de protección es  $\alpha$ , entonces las pruebas de las medias tienen un nivel de significación igual o mayor que  $\alpha$ .

## **PRUEBA DE BONFERRONI**

La prueba de Bonferroni permite comparar las medias de los  $t$  niveles de un factor después de haber rechazado la hipótesis nula ( $H_0$ ) de igualdad de medias mediante la técnica ANOVA.

Es un método que se utiliza para controlar el nivel de confianza simultáneo para un conjunto completo de intervalos de confianza. Es importante considerar el nivel de confianza simultáneo cuando se examinan múltiples intervalos de confianza porque las probabilidades de que al menos uno de los intervalos de confianza no contenga el parámetro de población es mayor para un conjunto de intervalos que para cualquier intervalo individual. Para contrarrestar esta tasa de error más elevada, el método de Bonferroni ajusta el nivel de confianza para cada intervalo individual, de manera que el nivel de confianza simultáneo resultante sea igual al valor que ha especificado.

Ajusta el nivel de significación en relación al número de pruebas estadísticas realizadas simultáneamente sobre un conjunto de datos.

Es un test de comparaciones múltiples.

En este procedimiento se fija un nivel de significación  $\alpha$  que se reparte entre cada una de las comparaciones consideradas y se utiliza la desigualdad de Bonferroni.

El Test de Bonferroni se basa en la creación de un umbral por encima del cual la diferencia entre las dos medias será significativa y por debajo del cual esa diferencia no lo será de estadísticamente significativa.

Suele ser bastante conservador y se utiliza más que todo cuando no son muchas las comparaciones a realizar y además, los grupos son homogéneos en varianzas. Una aproximación muy buena para su cálculo consiste en multiplicar el valor original de  $p$ , por el número de comparaciones posibles a realizar.

El método de Bonferroni ajusta el nivel de confianza para cada intervalo individual, de manera que el nivel de confianza simultáneo resultante sea igual al valor que ha especificado.

Los intervalos de confianza más amplios de Bonferroni proveen estimaciones menos precisas del parámetro de población, pero limitan a un máximo de 5 % la probabilidad de que uno o más de los intervalos de confianza no contenga el parámetro.