

Subject: Probability and Statistics

Class X: Variance Analysis



UdeA

Bioengineering

Francisco José Campuzano Cardona

Bioengineerer, MSc in Engineering

El análisis de varianza es una prueba estadística que permite conocer si los factores de un experimento dado, tienen efecto sobre la variable respuesta. Estas pruebas se basan en modelos lineales de la siguiente forma:

$$y_{ijk\dots} = \mu + \tau_i + \beta_j + \dots + \epsilon_{ijk\dots}$$

Diagram illustrating the components of the ANOVA model equation:

- $y_{ijk\dots}$: Observación $ijk\dots$ -esima estimada
- μ : Media poblacional
- τ_i : Efecto del factor τ , en el nivel i -ésimo.
- β_j : Efecto del factor β , en el nivel j -ésimo.
- $\epsilon_{ijk\dots}$: Error aleatorio en de la observación $ijk\dots$ -esima

El Análisis de Varianza por ANOVA es una prueba paramétrica que tiene varios supuestos:

- 1. Los errores o residuos son normales con media 0**
- 2. Los errores o residuos son independientes y aleatorios**
- 3. Las Observaciones son normales (variable respuesta)**
- 4. Hay homocedasticidad entre factores.**
- 5. Hay independencia entre puntos experimentales.**

Prueba de Runs. (WALD-WOLFOWITZ RunsTest)

Determina si una secuencia de datos están distribuidos de forma aleatoria, o si hay rachas o patrones.

H₀: La secuencia es aleatoria

H₁: La secuencia no es aleatoria.

Es de utilidad para corroborar que los residuos de un ANOVA son aleatorios.

Pasos en el Análisis de Varianza:

1. Estructurar los datos.

Previamente se debió establecer un diseño experimental, donde, por ejemplo, si se trataba de un diseño factorial, entonces se tendría una tabla con las diferentes corridas experimentales, que son las combinaciones posibles de los diferentes niveles de los diferentes factores, con sus réplicas. Entonces a cada corrida experimental, se le asigna su respectiva observación de la variable respuesta.

Pasos en el Análisis de Varianza:

1. Estructurar los datos.

Factor 1	Factor 2	Variable respuesta
1	1	3.5
1	1	3.6
1	0	2.5
1	0	2.4
0	1	2.3
0	1	2.4
0	0	6.1
0	0	5.8

Pasos en el Análisis de Varianza:

2. Corroborar los supuesto de Normalidad y Homocedasticidad entre niveles de los factores.
3. Determinar si hay interacciones (gráfico)
4. Realizar el ANOVA.
5. Corroborar los supuestos de Normalidad e independencia de los residuos o errores. Normalidad se puede analizar con un gráfico de probabilidad normal. ***

Para cada efecto, el ANOVA tiene la siguientes hipótesis

***H₀:** El efecto es igual a cero, es decir no es significativo*

***H₁:** El efecto es diferente de cero, es decir es significativo.*

Ejemplo:

En el proceso de limpieza de unas piezas de metal, se realiza un estudio para determinar si 3 factores influyen en la conductividad eléctrica y sus efectos. Los factores que se cree afectan la conductividad son:

1. Concentración de ácido sulfúrico: se usan 4 niveles 0%, 6%, 12% y 18%
2. Concentración de sal: se usan 3 niveles: 0%, 10% y 20%
3. Temperatura del baño: se usan 2 niveles 80°F y 100°F.

Además de realizan dos réplicas

Comparaciones Múltiples

Cuando se hace un ANOVA y en el experimento se tienen 3 o más niveles en algún factor que resulte tener efectos significativos, resulta útil realizar una prueba de comparaciones múltiples, para determinar si hay diferencias entre los diferentes niveles de ese factor.

Prueba HSD de Tukey ***

Prueba LSD de Fisher (t_test)

$$H_0: \mu_i = \mu_j$$

$$H_1: \mu_i \neq \mu_j$$

Prueba Kruskal Wallis

Cuando no se logran cumplir la totalidad de los supuestos de un ANOVA se debe recurrir a una prueba no paramétrica como esta. Tiene el supuesto de independencia.

Esta prueba no permite analizar las interacciones, de modo que se deben analizar los factores por separado. Y determinar si hay diferencias para cada nivel.

***H₀**: La media de las n muestras es igual.*

***H₁**: La media de al menos 1 muestra es diferente de las demás.*

Subject: Probability and Statistics



UdeA

Bioengineering

¡Thanks!

Francisco José Campuzano Cardona

Bioengineering. MSc in Engineering