



FACULTAD DE CIENCIAS
AGRONÓMICAS
UNIVERSIDAD DE CHILE

Practico Análisis de la Varianza

Erika Kania Kuhl
Ing. Agr. Dr.



Modelo de ANAVA a un criterio de clasificación

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

La **media general** (μ) puede ser entendida como la media de todas las medias poblacionales asociadas a cada uno de los grupos en estudio. Su valor es constante a través de todas las observaciones.

Los **efectos de tratamientos** (τ_i) representan la diferencia o corrimiento entre la media poblacional del tratamiento considerado y la media general común a todos los tratamientos

$$\tau_i = \mu_i - \mu$$

Así el **efecto del tratamiento 1** (τ_1) es la diferencia que hay entre la media del tratamiento 1 y la media general.

$$\tau_1 = \mu_1 - \mu$$

Ejemplo

Caso	Tratamiento	Repetición	Variable respuesta
1	1	1	5
2	1	2	6
3	1	3	4
4	2	1	8
5	2	2	10
6	2	3	6
7	3	1	11
8	3	2	12
9	3	3	10
10			

$i = 1, 2, 3$ (tratamientos)
 $j = 1, 2, 3$ (repeticiones)

Media general común a todos los tratamientos (μ)
 Media por tratamiento (μ_1, μ_2, μ_3)
 Efecto del tratamiento (τ_1, τ_2, τ_3) $\tau_i = \mu_i - \mu$



Ejemplo



Media general común a todos los tratamientos (μ)

Medidas resumen

Variable	Media
Variable respuesta	8.00

Media por tratamiento (μ_1, μ_2, μ_3)

Medidas resumen

Tratamiento	Variable	Media
1	Variable respuesta	5.00
2	Variable respuesta	8.00
3	Variable respuesta	11.00

Efecto del tratamiento (τ_1, τ_2, τ_3)

$$\tau_1 = 5 - 8 = -3$$

$$\tau_2 = 8 - 8 = 0$$

$$\tau_3 = 11 - 8 = 3$$

Ejemplo

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

e_{ij} = VALOR OBSERVADO – VALOR ESTIMADO

Se calcula como la diferencia entre el valor observado de la variable respuesta y el valor estimado (predicho) bajo el modelo lineal especificado.

(e_{ij} = residuo asociado a una unidad experimental)

Ejemplo



Caso	Tratamiento	Repetición	Variable respuesta	PRED_Variante respuesta	RDUO_Variante respuesta
1	1	1	5	5.00	e_{ij}
2	1	2	6	5.00	
3	1	3	4	5.00	
4	2	1	8	8.00	
5	2	2	10	8.00	
6	2	3	6	8.00	
7	3	1	11	11.00	
8	3	2	12	11.00	
9	3	3	10	11.00	
10					

e_{ij} = VALOR OBSERVADO – VALOR ESTIMADO

Modelo Lineal

Modelo que explica a cada observación:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

Recordemos que para comparar las medias de dos poblaciones, las hipótesis para una prueba bilateral, son:

$$H_0 : \mu_1 = \mu_2$$

$$H_A : \mu_1 \neq \mu_2$$

Para el caso de la comparación de medias de varias poblaciones, las hipótesis que se plantean son:

$$H_0 : u_1 = u_2 = u_3 = \dots = u_t (= u)$$

$$H_A : \text{existe } u_i \neq u$$

➔ Igualdad de medias poblacionales de todos los tratamientos comparados

Prueba de hipótesis para comparar varias poblaciones

El objetivo del Análisis de Varianza es contrastar la hipótesis de que los efectos de los tratamientos son nulos versus que al menos uno no lo es.

En términos estadísticos:

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \dots = \tau_t = 0$$

$$H_A : \text{algún } \tau_i \neq 0$$

Si la conclusión es rechazar H₀:

Al menos un tratamiento tiene efecto distinto de cero

τ_i : EFFECTO TRATAMIENTO

Hipótesis

$$H_0 : u_1 = u_2 = u_3 = \dots = ut (= u)$$

H_A : existe $ui \neq u$

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \dots = \pi = 0$$

H_A : algún τ_i distinto de cero.



Hipótesis
equivalentes

Análisis de Varianza (ANDEVA)

El ANDEVA es un procedimiento que permite descomponer la variabilidad total de las observaciones en la variación **entre tratamientos** y la variación **dentro de tratamientos** (variabilidad entre las unidades experimentales que reciben el mismo tratamiento- Error Experimental).

Análisis de Varianza (ANDEVA)

Esta técnica permite descomponer la **variación total** de una variable respuesta en **distintas fuentes de variación**, algunas de ellas **conocidas** y otras totalmente **desconocidas**, atribuibles al error experimental, con el fin de determinar si las variaciones conocidas pueden atribuirse al efecto de ciertos factores (tratamientos) o por el contrario a variaciones meramente muestrales.

Análisis de Varianza (ANDEVA)

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
Largo brote 1	9	0,82	0,76	17,68

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	54,00	2	27,00	13,50	0,0060
Tratamiento	54,00	2	27,00	13,50	0,0060
Error	12,00	6	2,00		
Total	66,00	8			

Prueba F

Todo ANDEVA concluye en una **Prueba F** que expresa cuantas veces es mayor la varianza explicada por los tratamientos que la no explicada (error)

Prueba F

$$F = \frac{\text{CM Tratamiento}}{\text{CM Error}}$$



F: Prueba que compara dos varianzas (por medio de sus cuocientes)

ANDEVA

Se plantearon las siguientes hipótesis:

$$H_0 : u_1 = u_2 = u_3 = \dots = ut(= u)$$

$$H_A : \text{existe } ui \neq u$$

ó

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \dots = \pi = 0$$

$$H_A : \text{algún } \tau_i \text{ distinto de cero.}$$

Estas hipótesis equivalen a decir:

HO : Varianza trat \leq Varianza error

HA : Varianza trat $>$ Varianza error

Nivel de significancia = α

En la práctica, es frecuente conducir ensayos con un **NIVEL DE SIGNIFICANCIA** de 0,05 ó 5%.

(Seleccionado a priori por el investigador)

$$\alpha = P(\text{Rechazar } H_0 / H_0 \text{ es verdadera}),$$

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \dots = \pi = 0$$

$$H_A : \text{algún } \tau_i \text{ distinto de cero.}$$

$$H_0 : u_1 = u_2 = u_3 = \dots = ut(= u)$$

$$H_A : \text{existe } ui \neq u$$

Si por ejemplo se escoge un nivel de significancia del 5%, entonces hay 5 oportunidades entre 100 de rechazar la hipótesis nula cuando debiera haberse aceptado.

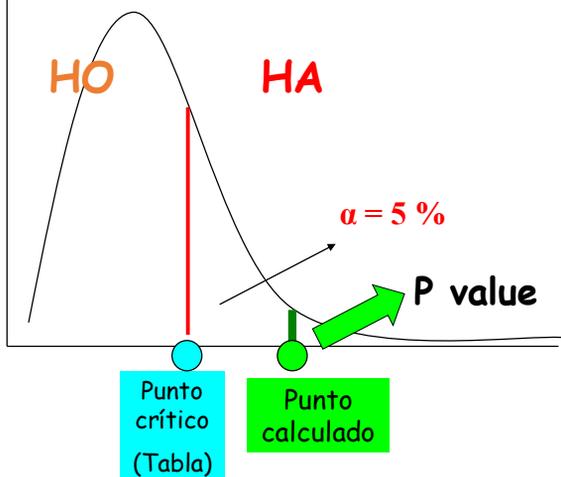
P - valor

Los software estadísticos presentan como resultado de la prueba una cantidad llamada "p-valor"

(o en inglés **p-value**)

P valor

Es el valor de probabilidad correspondiente al punto calculado

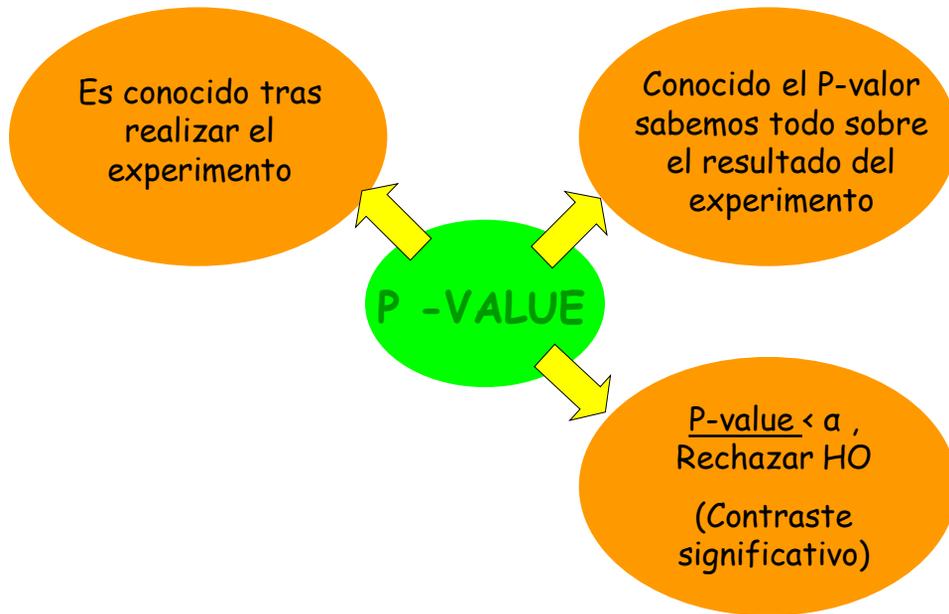


Es el valor exacto, a **posteriori**, de:

P (rechazar H_0 / H_0 verdadera)

Por lo tanto, el p-value es el nivel de significancia **OBSERVADO**

P valor



P - valor

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
Largo brote 1	9	0,82	0,76	17,68

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	54,00	2	27,00	13,50	0,0060
Tratamiento	54,00	2	27,00	13,50	0,0060
Error	12,00	6	2,00		
Total	66,00	8			

Conclusión: Los resultados sugieren que existen diferencias significativas entre los tratamientos (p valor = 0,0060 < 0,05), es decir se rechaza la hipótesis nula de igualdad de medias de los tratamientos.

ESTA CONCLUSION ES VALIDA SOLAMENTE SI

Se cumplen los **SUPUESTOS DEL ANAVA**, que deben cumplirse para que el valor **P reportado** sea válido.



FACULTAD DE CIENCIAS
AGRONÓMICAS
UNIVERSIDAD DE CHILE

Practico

Análisis de la Varianza

Erika Kania Kuhl
Ing. Agr. Dr.

