ESCUELA NACIONAL DE ESTUDIOS SUPERIORES
UNIDAD
UNAM MÉRIDA

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
ENES MÉRIDA
LICENCIATURA EN ECOLOGÍA

ESTADÍSTICA APLICADA I
ANOVA
(Parte 3 – Efectos aleatorios)

Prof. Edlin J. Guerra Castro

FACTORES

1) FACTOR: Variable categórica que identifica varios grupos o niveles que se desean comparar

2) FACTOR: puede ser fijo o aleatorio

3) Importante definir porque influye:
    1) Las suposiciones del modelo lineal
    2) Los cálculos de los cuadrados medios (MS)
    3) La interpretación de las hipótesis que se están sometiendo a prueba
    4) La validez de las inferencias estadísticas

FACTOR FIJO

1) Un conjunto finito de niveles

2) Todos ellos ocurren en el experimento

3) Hay hipótesis específicas asociadas a esos niveles

4) Si se repite el experimento, los mismos niveles se eligen

5) Los efectos de estos factores se consideran constantes en las variables respuestas.

FACTOR ALEATORIO

1) Los niveles se escogen aleatoriamente de un conjunto de posibles niveles

2) Es una fuente de variación dentro del modelo y no en los residuales

3) Los efectos se miden como tamaño de la varianza

4) Repetición del experimento no se traduce necesariamente en la selección de los mismos niveles

Naturaleza de la hipótesis

Efectos fijos
3 o más fuentes de variación

Efectos Aleatorios

Se aplica el

ANOVA

$H_0: \mu_1 = \mu_2 = \ldots = \mu_i$

$H_0: \sigma_A^2 = 0$

Próximo paso

Toca precisar la hipótesis alternativa con **otras pruebas** para interpretar el modelo

Se interpreta el modelo. Hay solo una hipótesis alternativa

FIJO O ALEATORIO
(Consideraciones)

1) El contexto de la hipótesis ¿cuál es la pregunta?

2) ¿Cómo fueron elegidos los niveles de ese factor’

3) ¿Tienen sentido las comparaciones entre niveles?

4) El alcance de las inferencias: ¿los niveles representan una población más amplia de ellos?
   1) Si = Aleatorio: inferencias se extienden a la población
   2) NO = Fijo: inferencias se limitan a los niveles escogidos

•

$$F = \frac{MS_A}{MS_{RES}}$$

FIJO

$$E(MS_{RES}) = \sigma_\varepsilon^2$$

$$E(MS_A) = \sigma_\varepsilon^2 + n\sum_{i=1}^{a}\alpha_i^2 /(a-1)$$

$$\frac{E(MS_A)}{E(MS_{RES})} = \frac{\sigma_\varepsilon^2 + n\sum_{i=1}^{a}\alpha_i^2 /(a-1)}{\sigma_\varepsilon^2}$$

$$\mathrm{H}_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$$

ALEATORIO

$$E(MS_{RES}) = \sigma_\varepsilon^2$$

$$E(MS_A) = \sigma_\varepsilon^2 + n\sigma_\alpha^2$$

$$\frac{E(MS_A)}{E(MS_{RES})} = \frac{\sigma_\varepsilon^2 + n\sigma_\alpha^2}{\sigma_\varepsilon^2}$$

$$\mathrm{H}_0 : \sigma_\alpha^2 = 0$$

ANOVA ORTOGONAL DOS FACTORES

Table 9.11 *F*-ratios used for testing main effects and interactions in a two factor ANOVA model for different combinations of fixed and random factors

Source	A and B fixed	A and B random	A fixed, B random Restricted version	A fixed, B random Unrestricted version
A	$\frac{MS_A}{MS_{Residual}}$	$\frac{MS_A}{MS_{AB}}$	$\frac{MS_A}{MS_{AB}}$	$\frac{MS_A}{MS_{AB}}$
B	$\frac{MS_B}{MS_{Residual}}$	$\frac{MS_B}{MS_{AB}}$	$\frac{MS_B}{MS_{Residual}}$	$\frac{MS_B}{MS_{AB}}$
AB	$\frac{MS_{AB}}{MS_{Residual}}$	$\frac{MS_{AB}}{MS_{Residual}}$	$\frac{MS_{AB}}{MS_{Residual}}$	$\frac{MS_{AB}}{MS_{Residual}}$

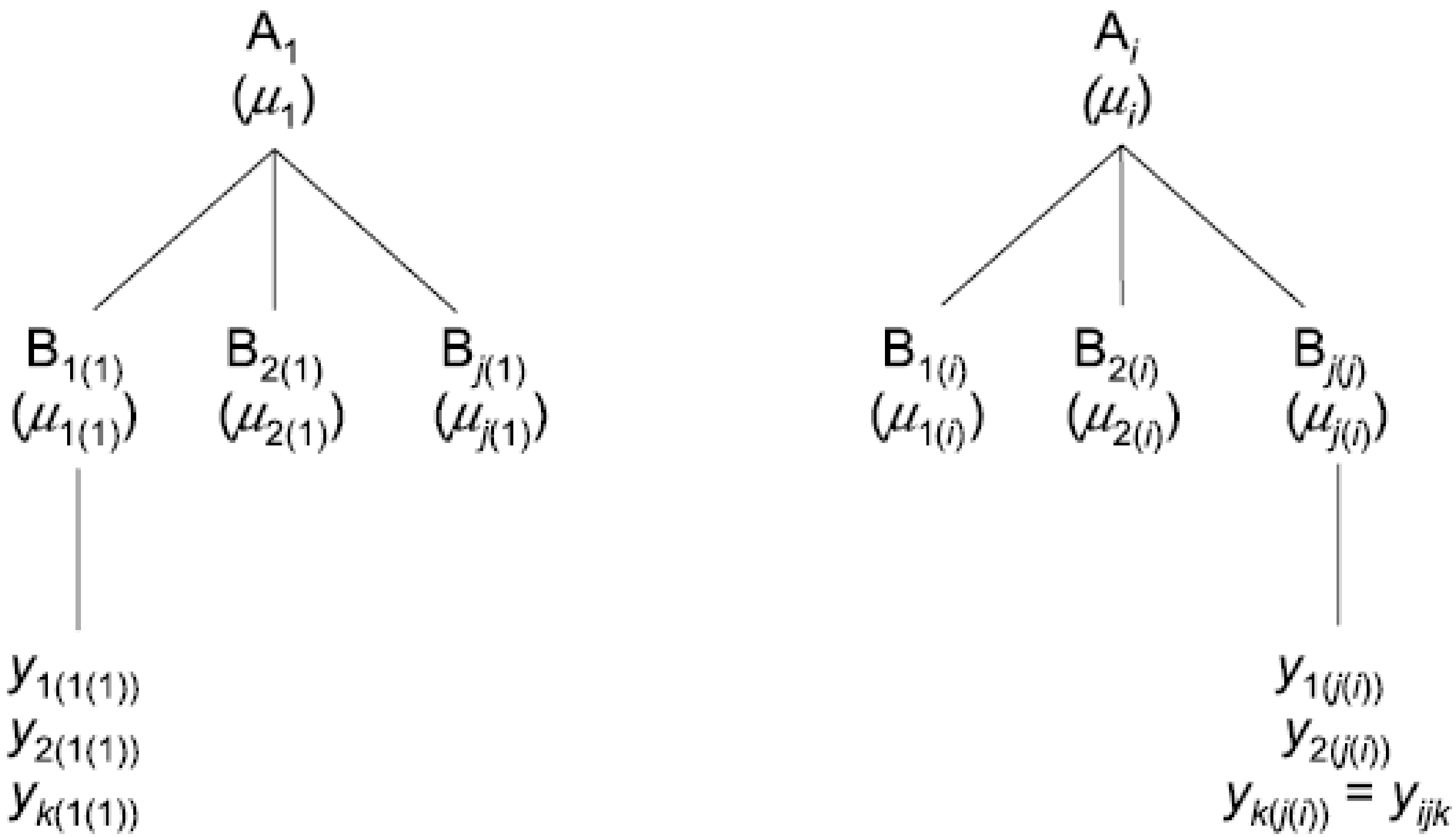
	A, B fixed	A, B random
MS_A	$\sigma_\varepsilon^2 + nq\frac{\sum_{i=1}^{p}\alpha_i^2}{p-1}$	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + nq\sigma_\alpha^2$
MS_B	$\sigma_\varepsilon^2 + np\frac{\sum_{j=1}^{q}\beta_j^2}{q-1}$	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2 + np\sigma_\beta^2$
MS_{AB}	$\sigma_\varepsilon^2 + n\frac{\sum_{i=1}^{p}\sum_{j=1}^{q}(\alpha\beta)_{ij}^2}{(p-1)(q-1)}$	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta}^2$
$MS_{Residual}$	σ_ε^2	σ_ε^2

ANOVA ANIDADO

FACTORES ANIDADOS

Un factor está anidado en otro factor si sus niveles pertenecen únicamente a un solo nivel del otro factor

Evaluar la DBO en 4 puntos a lo largo del Río luego de la instalación de la planta


A_1
(μ_1)
$B_{1(1)}$
$(\mu_{1(1)})$
$B_{2(1)}$
$(\mu_{2(1)})$
$B_{j(1)}$
$(\mu_{j(1)})$
$y_{1(1(1))}$
$y_{2(1(1))}$
$y_{k(1(1))}$
A_i
(μ_i)
$B_{1(i)}$
$(\mu_{1(i)})$
$B_{2(i)}$
$(\mu_{2(i)})$
$B_{j(i)}$
$(\mu_{j(i)})$
$y_{1(j(i))}$
$y_{2(j(i))}$
$y_{k(j(i))} = y_{ijk}$

ANALISIS DE VARIANZA (anidado)

Poco común, y debería ser obligatorio, en estudios espaciales

Factor 1 $$H_0 : \mu_1 = \mu_2 = \mu_3 \ldots = \mu_i$$

Factor 2 (1) $$H_0 : \sigma_B^2 = 0$$

Modelo líneal

Datos = Modelo o Señal + error (ruido)

$$H_0 : x_{ij} = \mu + e_{ij}$$

$$H_i : x_{ij} = \mu + A + e_{ij}$$

$$H_i : x_{ij} = \mu + A + B\left(A\right) + e_{ij}$$

Table 9.3 ANOVA table for two factor nested linear model with factor A (p levels), factor B (q levels) nested within A, and n replicates within each combination of A and B

Source	SS	df	MS
A	$nq\sum_{i=1}^{p}(\bar{y}_i - \bar{y})^2$	$p-1$	$\frac{SS_A}{p-1}$
B(A)	$n\sum_{i=1}^{p}\sum_{j=1}^{q}(\bar{y}_{j(i)} - \bar{y}_i)^2$	$p(q-1)$	$\frac{SS_{B(A)}}{p(q-1)}$
Residual	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{j(i)})^2$	$pq(n-1)$	$\frac{SS_{Residual}}{pq(n-1)}$
Total	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{n}(y_{ijk} - \bar{y})^2$	$pqn-1$	

Table 9.4 Expected mean squares and F-ratios for tests of null hypotheses for two factor nested ANOVA model

	A fixed, B random	
Source	Expected mean square	F-ratio
A	$\sigma_\varepsilon^2 + n\sigma_\beta^2 + nq\frac{\sum_{i=1}^{p}\alpha_i^2}{p-1}$	$\frac{MS_A}{MS_{B(A)}}$
B(A)	$\sigma_\varepsilon^2 + n\sigma_\beta^2$	$\frac{MS_{B(A)}}{MS_{Residual}}$
Residual	σ_ε^2	σ_ε^2

Luego del ANOVA, lo siguiente es estimar el Componente de Variación de cada fuente

Table 9.5 Estimates of variance components (using ANOVA approach) for two factor nested design with B(A) random

Source	Estimated variance component
A	$\frac{MS_A - MS_{B(A)}}{nq}$*
B(A)	$\frac{MS_{B(A)} - MS_{Residual}}{n}$
Residual	$MS_{Residual}$

Note:
*This represents variance between population means of specific levels of A if factor A is fixed and a true added variance component if A is random.

RESUMEN

- El ANOVA puede ser tan complejo como lo requiera la hipótesis y modelo.
- El diseño del muestreo o experimento es fundamental para que el análisis estadístico sea robusto.
- Los factores se definen según la naturaleza de la hipótesis en Fijos o Aleatorios
- Los ANOVA pueden generar interacciones (ortogonales) o anidarse en otro, siempre que sean aleatorios.
- La mecánica del análisis es **dependiente** de cómo definamos cada factor y sus interrelaciones