

Biometría



Clase 13

Análisis de la varianza (ANOVA) de un factor

Objetivo

- ▣ El análisis de la varianza (ANOVA) permite comparar **dos o más** medias poblacionales asociadas a tratamientos
- ▣ Se utiliza cuando existe una variable dependiente o respuesta (cuantitativa) y una o más variables explicatorias o independientes o factores cualitativas que afectan a la variable respuesta

¿Qué son los tratamientos?

- En un experimento de un factor, los tratamientos son los niveles del factor (los distintos valores que toma)

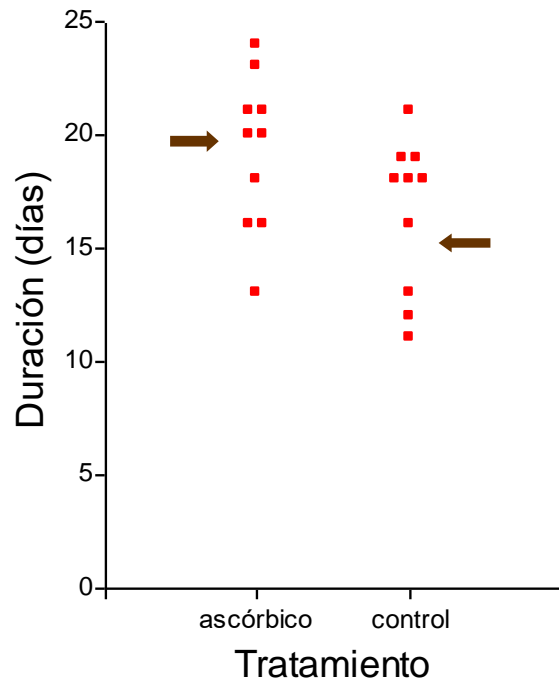
Ej: Se prueban 2 dietas y un control

- Existe un solo factor (tratamiento con dietas)
- con tres niveles o tratamientos

¿En qué consiste el Anova?

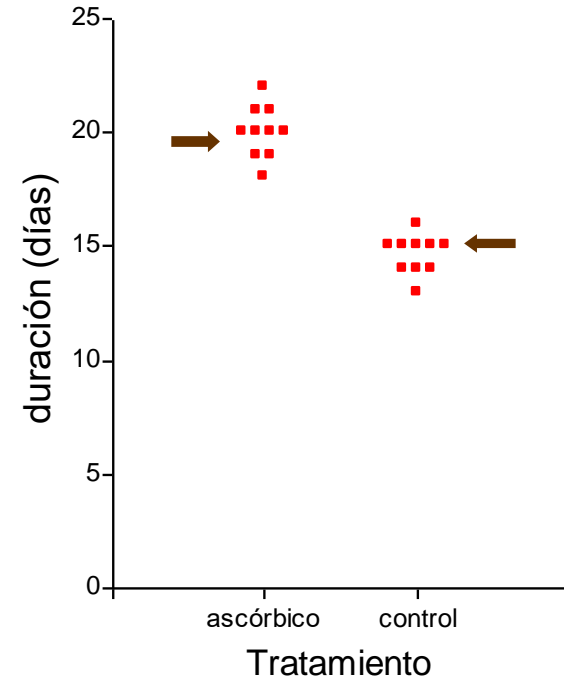
- ❑ La manera para determinar si existen diferencias entre dos o más medias poblacionales es examinar la diferencia **entre** las medias muestrales y compararla con una medida de variabilidad **dentro** de las muestras
- ❑ La separación entre las medias es una parte de la historia; la otra parte es la variabilidad dentro de cada grupo
- ❑ El ANOVA usa esta estrategia para comparar dos o más medias

¿En qué consiste el Anova?



La diferencia entre las medias muestrales es relativamente **pequeña** cuando se la compara con la variabilidad dentro de cada grupo

$H_0: \mu_1 = \mu_2$ No se rechaza H_0



La diferencia entre las medias muestrales es relativamente **grande** cuando se la compara con la variabilidad dentro de cada grupo

Se rechaza H_0

Anova

Consiste en descomponer la variabilidad de la variable respuesta en:

- ▣ Variabilidad explicada o debida a los tratamientos (variabilidad **entre** tratamientos)
- ▣ Variabilidad no explicada por los tratamientos o variabilidad aleatoria o residual (variabilidad **dentro** de los tratamientos)

Ejemplo



Efecto de la suplementación durante la lactancia sobre la ganancia en peso de las crías de cabras al destete

15 cabritos recién nacidos fueron asignados al azar a alguno de los siguientes tratamientos:

- ▣ Sin suplementación (control)
- ▣ Suplementación con 500 g de alfalfa / día
- ▣ Suplementación con 500 g de alfalfa y 125 g de maíz partido / día

Los cabritos fueron pesados al nacimiento y al momento del destete y se calculó la ganancia en peso (y)

Resultados (en kg)

	Control	Suplem alfalfa	Suplem alfalfa y maiz	
a=3	1.93	2.54	2.95	
n _i =5	2.72	2.65	3.24	
n=15	2.44	2.57	2.55	
	2.2	2.94	2.84	
	1.97	2.93	3.05	
Media	2.252	2.726	2.926	2.6347
varianza	0.1101	0.038	0.0658	
desvío std	0.3318	0.1950	0.2566	

El modelo estadístico es:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

- ▣ donde μ es la media general o media de la población
- ▣ α_i es el efecto del tratamiento i y que es común a todos los individuos que recibieron ese tratamiento
- ▣ ε_{ij} es el residuo o error aleatorio que existe dentro de cada tratamiento, entre los individuos

Hipótesis en ANOVA

- $H_0: \mu_1 = \mu_2 = \mu_3 = \mu$ es decir:
 - la respuesta (ganancia en peso) promedio de los tratamientos es la misma
 - $\alpha_i = 0$
 - No existe efecto debido a la suplementación sobre la ganancia en peso
- $H_1: \text{alguna } \mu_i \neq \mu$ es decir:
 - al menos un tratamiento difiere del resto en cuanto a la ganancia en peso promedio
 - Algún $\alpha_i \neq 0$
 - Existe efecto debido a la suplementación sobre la ganancia en peso

Tabla de ANOVA

Se descompone la variabilidad de los datos en:

- ▣ Variabilidad entre tratamientos
- ▣ Variabilidad dentro de los tratamientos

La variabilidad se cuantifica mediante varianzas

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

$$\text{Varianza o cuadrados medios} = \frac{\text{Suma de cuadrados}}{\text{Grados de libertad}}$$

Tabla de ANOVA

Fuente de Variación	SC	GL	CM	F
Entre Tratamientos	$\sum n_i (\bar{y}_i - \bar{y})^2$	a-1	$\frac{SCe}{GLe}$	$\frac{CMe}{CMd}$
Dentro de tratamientos	$\sum (y_{ij} - \bar{y}_i)^2$	(ni-1)a = n-a	$\frac{SCd}{GLd}$	
Total	$\sum (y_{ij} - \bar{y})^2$	n-1		

	Control	Suplem alfalfa	Sup alfalfa y maiz	
a=3	1.93	2.54	2.95	
n _i =5	2.72	2.65	3.24	
n=15	2.44	2.57	2.55	
	2.2	2.94	2.84	
	1.97	2.93	3.05	
Media	2.252	2.726	2.926	2.6347

Tabla de ANOVA

- La pregunta es: hay más variabilidad **entre** las medias de los tratamientos que **dentro** de los tratamientos?
- Para contestarla se hace un **cociente** entre la varianza entre tratamientos y la varianza dentro de los tratamientos (F_m)
- Si H_0 es verdadera (no hay diferencias entre tratamientos) el cociente debería ser de 1 ya que ambas varianzas (entre y dentro) están estimando la misma variabilidad
- Por ser un cociente de varianzas, ese estadístico F sigue una **distribución F de Fisher** con GL del numerador y GL del denominador

Volviendo al ejemplo:

- ▣ En Infostat
 - Estadísticas >
 - Medidas resumen

Estadística descriptiva

Trat	Variable	n	Media	D.E.	Var (n-1)	Mín	Máx
1	Ganancia en peso	5	2.252	0.332	0.110	1.930	2.720
2	Ganancia en peso	5	2.726	0.195	0.038	2.540	2.940
3	Ganancia en peso	5	2.926	0.257	0.066	2.550	3.240

Resultados:

Análisis de la varianza

Variable	N	R ²	R ² Aj	CV
Ganancia en peso	15	0.58	0.51	10.14

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	1.20	2	0.60	8.40	0.0052
Trat	1.20	2	0.60	8.40	0.0052
Error	0.86	12	0.07		
Total	2.05	14			

Observese que $SC_{\text{entre}} + SC_{\text{dentro}} = SC_{\text{total}}$
 $GL_{\text{entre}} + GL_{\text{dentro}} = GL_{\text{total}}$

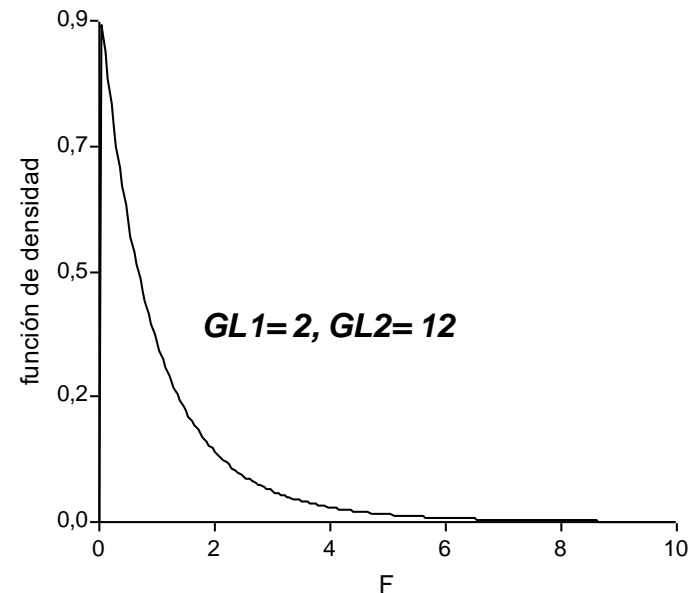
¿Cómo se decide si se rechaza o no H_0 ?

- ▣ Asumiendo que H_0 es verdadera, calcular la probabilidad de obtener un valor tan o más extremo que el obtenido en la muestra (**p de la prueba o p valor**). Si p es menor al nivel de significación α , se rechaza H_0 y se concluye que al menos un promedio difiere del resto (existen diferencias significativas). Caso contrario, se concluye que no existen evidencias muestrales suficientes para rechazar H_0

Volviendo al ejemplo:

Cuadro de Análisis de la Varianza

F.V.	SC	gl	CM	F	p-valor
Modelo	1.20	2	0.60	8.40	0.0052
Trat	1.20	2	0.60	8.40	0.0052
Error	0.86	12	0.07		
Total	2.05	14			



$$p < \alpha$$
$$0.0052 < 0.05$$

Rechazo H_0

Al menos un tratamiento difiere del resto
(con una probabilidad de error del 5%)

Supuestos del Anova



Supuestos del modelo

Para que las conclusiones del Anova sean válidas, se deben verificar una serie de supuestos:

- ❑ Las muestras deben ser **aleatorias** y las observaciones **independientes** entre sí.
- ❑ Las observaciones de cada tratamiento deben proceder de poblaciones donde la variable presenta distribución **normal**
- ❑ Los tratamientos deben tener la **misma variabilidad** (homocedasticidad) \Rightarrow

Muy importante!



Los diseños balanceados son menos sensibles al incumplimiento de los supuestos

Supuestos del modelo

- ▣ Estos supuestos se pueden resumir en:

ε_{ij} deben ser independientes y estar normalmente distribuidos, con una media de 0 y un desvío σ , es decir:

$$\varepsilon_{ij} \sim \text{NID}(0, \sigma)$$

Los ε_{ij} se estiman con los residuos e_{ij}

¿Qué son los residuos?

- El residuo e_{ij} es la diferencia entre el valor observado en una unidad experimental y el valor esperado si no existiesen factores no controlados

	Control	Suplem alfalfa	Suplem alfalfa y maiz
	1.93	2.54	2.95
	2.72	2.65	3.24
	2.44	2.57	2.55
	2.2	2.94	2.84
	1.97	2.93	3.05
\bar{x}	2.252	2.726	2.926

	Control	Suplem alfalfa	Suplem alfalfa y maiz
	-0,322	-0,186	0,024
	0,468	-0,076	0,314
	0,188	-0,156	-0,376
	-0,052	0,214	-0,086
	-0,282	0,204	0,124
\bar{x}			

¿Cómo probar normalidad?

Para cada tratamiento la **variable** debe seguir una distribución normal

O alternatively, los **residuos** deben ajustar a una distribución normal

- ♦ **Gráficamente:** QQ plot

- ♦ **Analíticamente:**

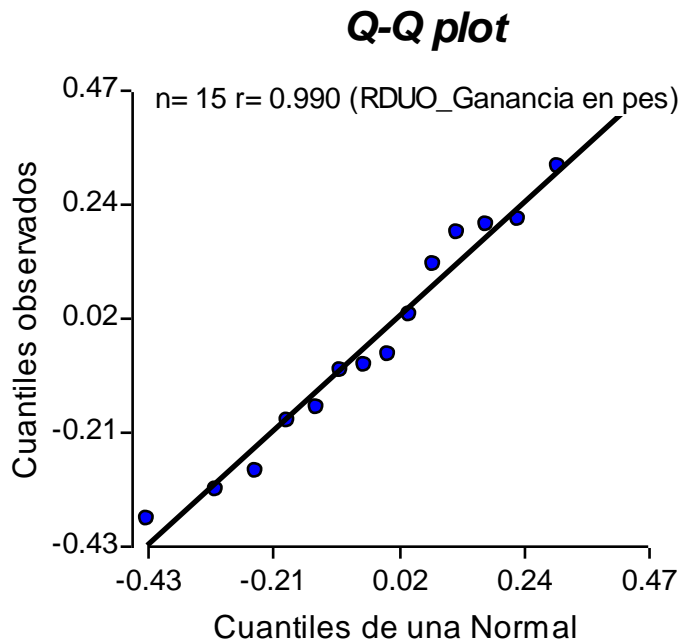
 - Prueba de Shapiro-Wilks

Se trata de una prueba estadística no paramétricas

Ho: **la variable tiene buen ajuste a una distribución normal**

En el ejemplo:

- Supuesto: Para cada tratamiento la variable sigue una distribución normal
 - Pero como solo son 5 datos por tratamiento, se prueba normalidad con los 15 residuos



Shapiro-Wilks (modificado)

Variable	n	Media	D.E.	W*	p (una cola)
RDUO ganancia	15	0,00	0,25	0,95	0,6784

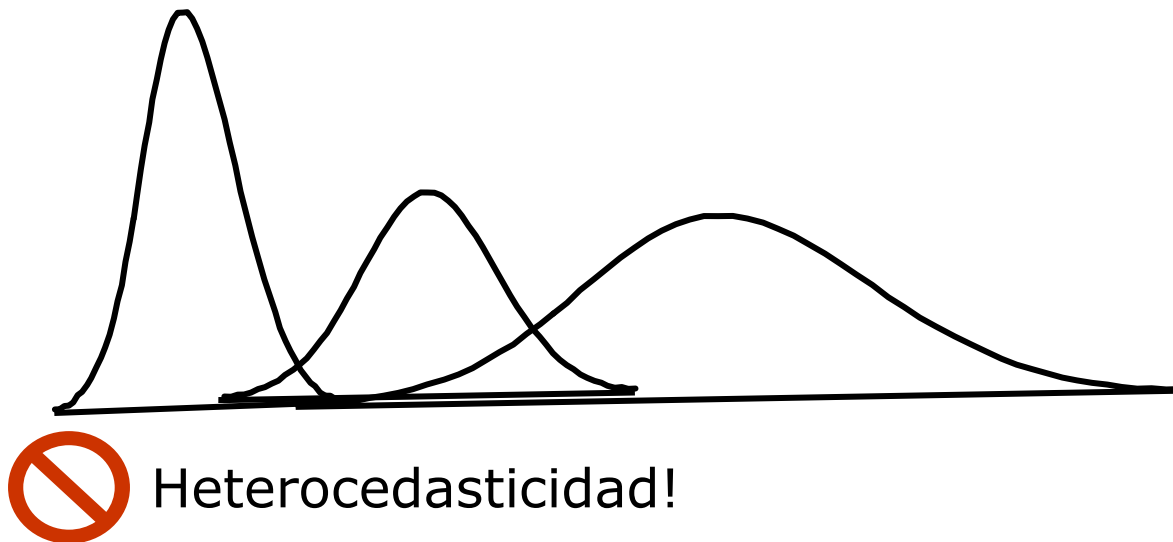
Normalidad

Algunas consideraciones

- ❑ No es crítica (el Anova es bastante robusto a la falta de normalidad)
- ❑ Si la cantidad de réplicas **por tratamiento** es grande (por ej. ≥ 30) no hace falta (Teorema Central del límite)
- ❑ Si la variable se aparta mucho de la normalidad puede ser necesario:
 - Ajustar modelos lineales generalizados o **Transformar** los datos (todos).

Homocedasticidad

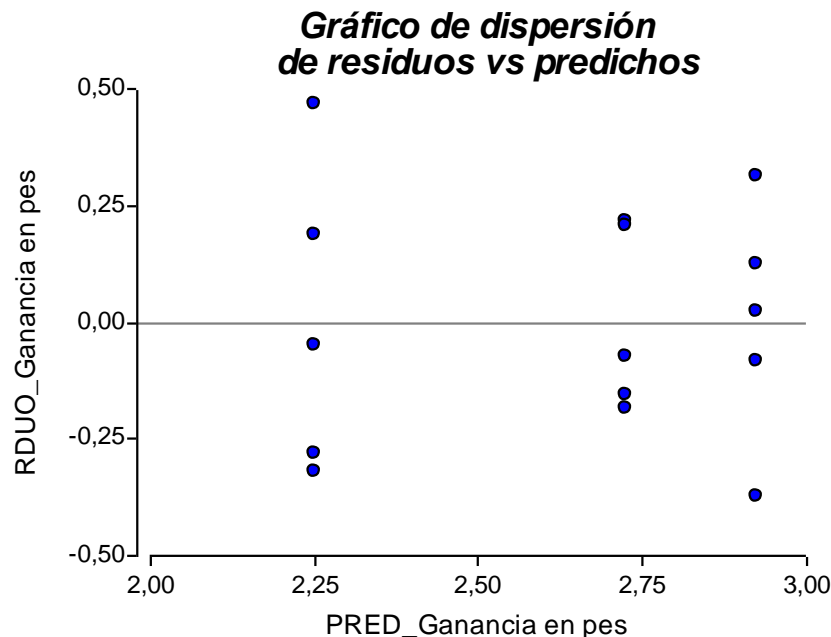
- Consiste en suponer que todos los tratamientos tienen la misma variabilidad



¿Cómo detectamos si hay heterocedasticidad?

□ Gráficamente:

- Se calculan los residuos y se grafican vs los valores esperados o predichos
- Se espera encontrar una distribución al azar y con variabilidad constante



¿Cómo detectamos si hay heterocedasticidad?

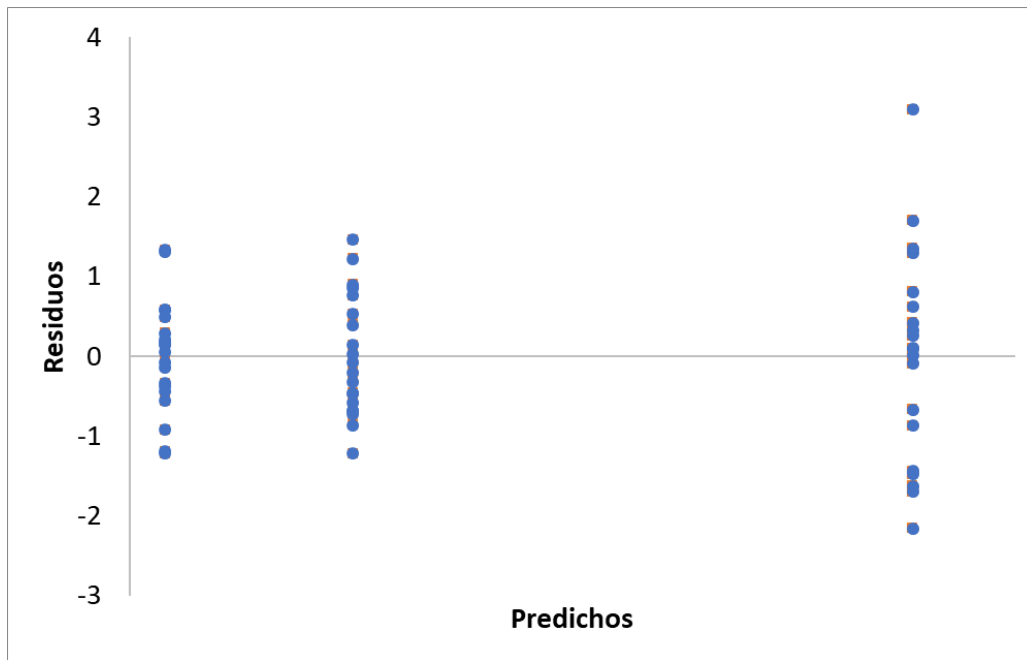
▣ Prueba de Levene

$$H_0: \sigma^2_1 = \sigma^2_2 = \sigma^2_3 = \sigma^2$$

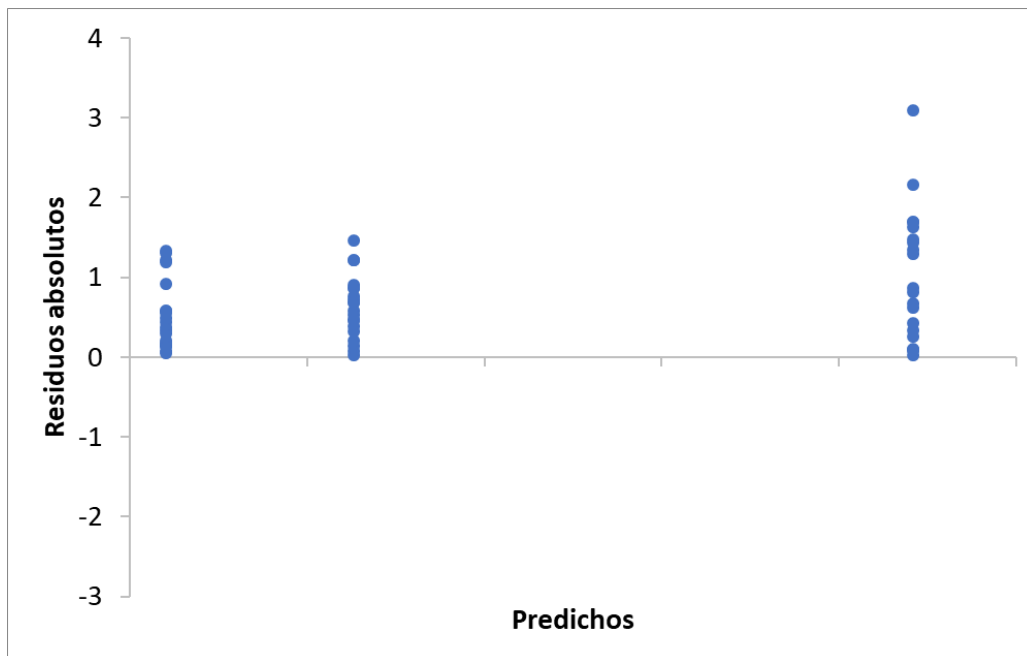
$$H_1: \text{alguna } \sigma^2_i \neq \sigma^2$$

La prueba consiste en realizar un ANOVA con los residuos absolutos

Nota: La falta de homocedasticidad es más grave si el diseño es desbalanceado



Residuos vs predichos



Residuos absolutos vs predichos

Resumiendo:

- ▣ Los supuestos se cumplen, por lo tanto las conclusiones del análisis son válidas. De todas maneras los supuestos deben probarse antes de la PH
- ▣ $H_0: \mu_1 = \mu_2 = \mu_3 = \mu$ $H_1: \text{alguna } \mu_i \neq \mu$

Pruebas de hipótesis tipo III - prueba

	Source	numDF	denDF	F-value	p-value
1	Dieta	2	12	8.40	0.0052

Esta hipótesis es muy general. Para decidir cuáles tratamientos difieren se deben efectuar comparaciones múltiples

Comparaciones entre tratamientos



Comparaciones múltiples

- ❑ Sirven para detectar diferencias entre las medias de los tratamientos
- ❑ ¿Por qué no múltiples pruebas t para comparar los grupos entre sí?
 - El error global sería muy alto (cada prueba tendría asociado un $\alpha\%$ de error)
 - Por ejemplo, si se comparan 4 tratamientos, hay pruebas t posibles. Si cada una tiene asociado un nivel de significación del 5%, la probabilidad de cometer al menos 1 error de tipo I (**error global o tasa de error por experimento**) es

$$P_{binomial}(x \geq 1 / n = 6, p = 0,05) = 0,26$$

- ❑ Existen distintos métodos de comparación. Cada método controla de alguna manera el error global, es decir la probabilidad de cometer al menos un error tipo I.

Clasificación de los métodos

□ A priori: planeados

- Las hipótesis se plantean antes del muestreo
- Se basan en información independiente del experimento
- pueden detectar diferencias aunque el anova no lo haya hecho, ya que incorporan información que ni la hipótesis global ni las comparaciones a posteriori poseen
- Mayor sensibilidad, mayor potencia
 - Contrastes ortogonales

□ A posteriori: no planeados

- se aplican sólo si el Anova dio significativo
- “búsqueda de significación”, exploratorio
- Más conservativos, menos potentes
 - Método de Tukey

Contrastes ortogonales

- ❑ **Contraste** f es una combinación lineal de los parámetros del modelo
- ❑ En cada contraste, $\sum c_i = 0$
- ❑ Los contrastes **ortogonales** deben ser independientes entre sí
- ❑ Consisten en una descomposición de la SC de los tratamientos
- ❑ Como máximo #trat-1 contrastes
- ❑ Solo para diseños balanceados
- ❑ Para cada contraste se calcula una F

$$f = \sum c_i \mu_i \qquad \hat{f} = \sum c_i \bar{y}_i$$

$$\sum c_i c_j = 0 \text{ para todos } (i, j)$$

Contrastes ortogonales

Contraste 1: f_1 $H_0: \mu_1 = \frac{\mu_2 + \mu_3}{2}$ o lo que es lo mismo $\mu_1 - \frac{\mu_2 + \mu_3}{2} = 0$

$H_a: \mu_1 \neq \frac{\mu_2 + \mu_3}{2}$ o lo que es lo mismo $\mu_1 - \frac{\mu_2 + \mu_3}{2} \neq 0$

Contraste 2: f_2 $H_0: \mu_2 = \mu_3$ o lo que es lo mismo $\mu_2 - \mu_3 = 0$

$H_a: \mu_2 \neq \mu_3$ o lo que es lo mismo $\mu_2 - \mu_3 \neq 0$

Comprobar ortogonalidad: $\sum c_i c_j = 0$ para todos (i, j)

Contrastes ortogonales

$$H_{o1} : 2\mu_1 - 1\mu_2 - 1\mu_3 = 0$$

$$H_{o2} : 0\mu_1 + 1\mu_2 - 1\mu_3 = 0$$

Contraste	Tratamiento	n_i	Media	c_i
f1	Control	5	2,252	2
	+ alfalfa	5	2,726	-1
	+ alfalfa +granos	5	2,926	-1
f2	Control	5	2,252	0
	+ alfalfa	5	2,726	1
	+ alfalfa +granos	5	2,926	-1

$$F_{(\hat{f})} = \frac{CM(\hat{f}_i)}{CM_{error}}$$

$$CM(\hat{f}_i) = \frac{SC(\hat{f}_i)}{GL(\hat{f}_i)}$$

$$SC(\hat{f}_i) = \frac{\hat{f}_i^2}{\sum c_i^2 / n_i}$$

$$SC(\hat{f}_1) = \frac{-1,148^2}{6/5} = \frac{1,32}{1,20} = 1,098$$

$$SC(\hat{f}_2) = \frac{-0,20^2}{2/5} = \frac{0,04}{0,40} = 0,1$$

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	1,198	2	0,599	8,402	0,0052
Dieta	1,198	2	0,599	8,402	0,0052
Error	0,856	12	0,071		
Total	2,054	14			

Contrastes

Dieta	SC	gl	CM	F	p-valor
Contraste1	1,098	1	1,098	15,401	0,0020
Contraste2	0,100	1	0,100	1,402	0,2593
Total	1,198	2	0,599	8,402	0,0052

Prueba de Tukey

- ▣ Compara todos los pares de medias posibles
 - $H_0: \mu_i = \mu_j$ (o lo que es lo mismo, $\mu_i - \mu_j = 0$)
 - $H_1: \mu_i \neq \mu_j$ (o lo que es lo mismo, $\mu_i - \mu_j \neq 0$)

- ▣ El estadístico es

$$\hat{f}_{ij} = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{\frac{CM_{dentro}}{n_1} + \frac{CM_{dentro}}{n_2}}} \sim q_{GLdentro, \alpha}$$

Utilizando Infostat:

\$lsmeans

Dieta	lsmean	SE	df	lower.CL	upper.CL	.group
Control	2.252	0.119	12	1.992	2.512	1
Alfalfa	2.726	0.119	12	2.466	2.986	2
Alfalfa y maiz	2.926	0.119	12	2.666	3.186	2

Confidence level used: 0.95

\$contrasts

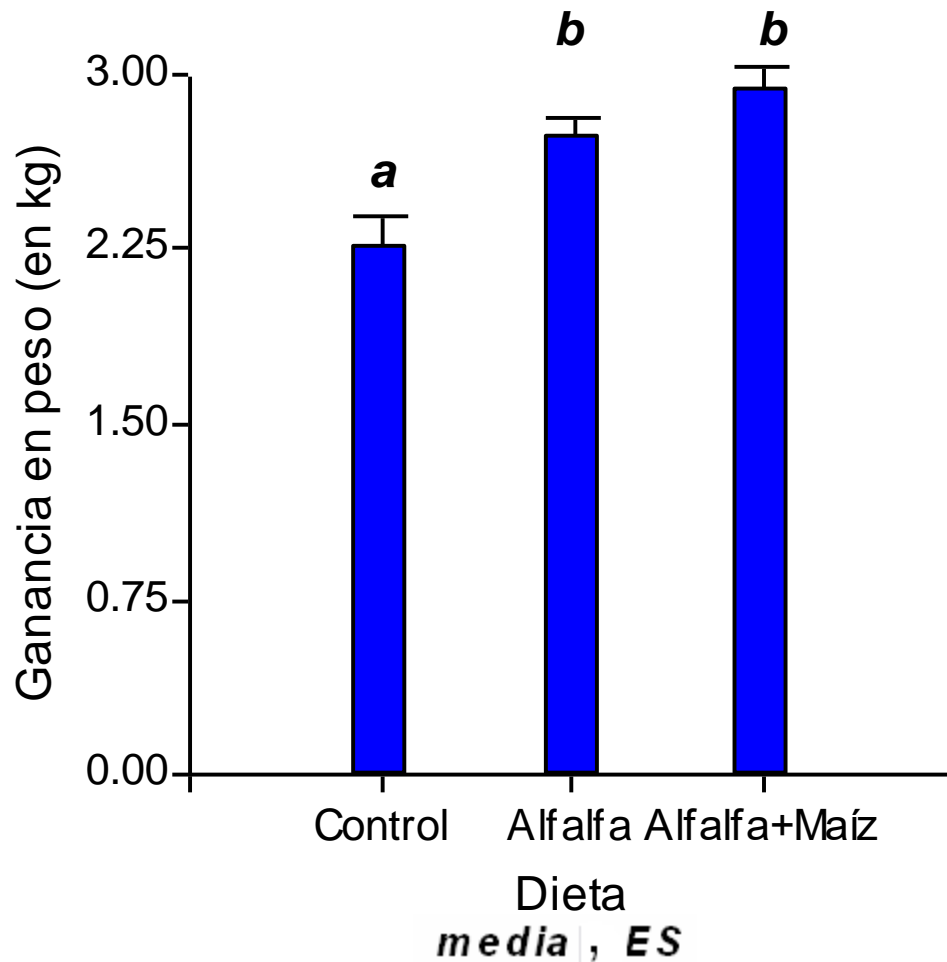
contrast	estimate	SE	df	t.ratio	p.value
Control - Alfalfa	-0.474	0.169	12	-2.807	0.039
Control - Alfalfa y maiz	-0.674	0.169	12	-3.991	0.005
Alfalfa - Alfalfa y maiz	-0.200	0.169	12	-1.184	0.484

P value adjustment: tukey method for comparing a family of 3 estimates

Conclusión:

Resumiendo:

Efecto de la suplementación materna en cabritos



**Gráfico
de barras**

Conclusiones:

- ▣ La suplementación produjo un aumento significativo de la ganancia en peso de los cabritos ($p < 0.05$)
- ▣ El agregado de maíz a la alfalfa no produjo un efecto significativo

Magnitud del efecto d: mide la diferencia promedio entre dos o más tratamientos en valores absolutos y/o relativos.

- ▣ El incremento en la ganancia en peso por suplementación con alfalfa fue de 474g o del 21% en promedio