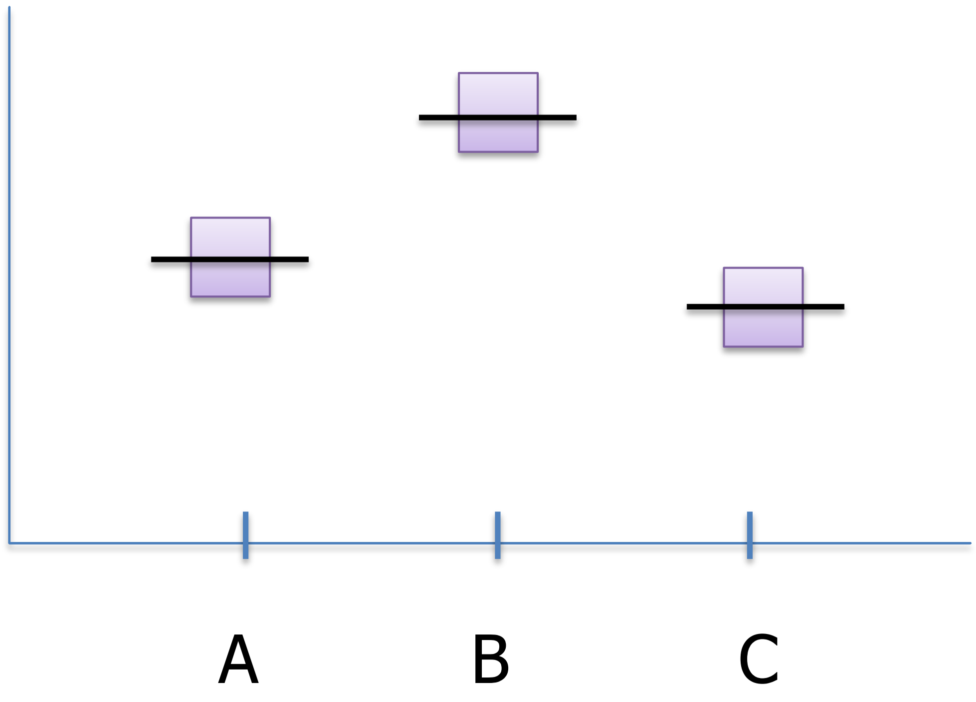
ANOVA

Introducción, conceptos fundamentales
y
“como funciona por dentro”

Universitat d’Alacant
Universidad de Alicante
Departamento de Ciencias del Mar y Biología Aplicada
Prof. Jose Jacobo Zubcoff Vallejo

ANOVA

¿Qué es y para que sirve?

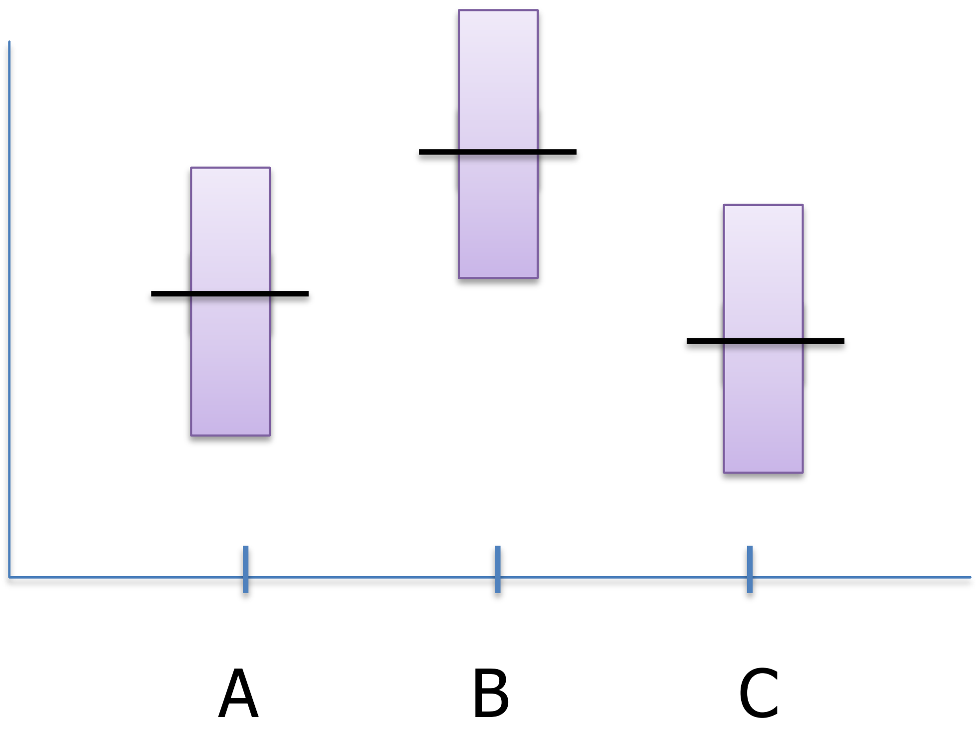


https://josezubcoff.shinyapps.io/Understanding_ANOVA

Para entender el comportamiento y sensibilidad de ANOVA

ANOVA

¿Qué es y para que sirve?



$$H_0: \mu_A = \mu_B = \mu_C = \ldots$$

H_1: *Al menos una igualdad no es cierta*

ANOVA

Modelo lineal

Efecto del

$$X_{ij} = \mu_{..} + \boldsymbol{Factor_i} + \boldsymbol{e_{ij}}$$

Contraste de Hipótesis:

H_0 : No hay efecto del Factor → Las medias son iguales

H_1 : Hipótesis nula no cierta → Alguna media es distinta

ANOVA

Supongamos un universo de notas de 9 alumnos y alumnas de 3 grupos distintos

Grupo 1	Grupo 2	Grupo 3
5	5	5
5	5	5
5	5	5

No hay diferencia ENTRE grupos
Ni DENTRO de los grupos

$$X_{i,j} = \mu$$

ANOVA

Supongamos que aplicamos un método de enseñanza (*factor*) que afecta:

Grupo 1	Grupo 2	Grupo 3
5+1=6	5+2=7	5+0=5
5+1=6	5+2=7	5+0=5
5+1=6	5+2=7	5+0=5

$$X_{i,j} = \mu + \boldsymbol{\alpha_i}$$

Donde $\boldsymbol{\alpha_i} = \{1,2,0\}$ *es el efecto del factor*

El factor influye en establecer diferencias ENTRE grupos
Estas diferencias se pueden cuantificar

ANOVA

Por razones ALEATORIAS algunos alumnos o alumnas rinden mas que otras

Grupo 1	Grupo 2	Grupo 3
5+1**-1** = 5	5+2**+2** = 9	5+0**+3** = 8
5+1**-2** = 4	5+2**+0** = 7	5+0**+4** = 9
5+1**+0** = 6	5+2**+1** = 8	5+0**+0** = 5

$$X_{i,j} = \mu + \alpha_i + \varepsilon_{i,j}$$

Donde $\varepsilon_{i,j}$= {-1,-2,0,2,0,1,3,4,0} *efecto de la aleatoriedad*

La ALEATORIEDAD influye en la variabilidad **DENTRO** de los grupos

https://josezubcoff.shinyapps.io/Understanding_ANOVA

Para entender el comportamiento y sensibilidad de ANOVA

ANOVA

Grupo 1	Grupo 2	Grupo 3	
5+1-1 = 5	5+2+2 = 9	5+0+3 = 8	
5+1-2 = 4	5+2+0 = 7	5+0+4 = 9	
5+1+0 = 6	5+2+1 = 8	5+0+0 = 5	
$X_{1.} = 5$	$X_{2.} = 8$	$X_{3.} = 7.33$	$X_{..} = 6.78$

Calculamos las medias por grupo y la media global

La variabilidad **TOTAL** la podemos desdoblar: **ENTRE** y **DENTRO** de los grupos

ANOVA

Grupo 1	Grupo 2	Grupo 3	
5	9	8	
4	7	9	
6	8	5	
$X_{1.} = 5$	$X_{2.} = 8$	$X_{3.} = 7.33$	$X_{..} = 6.78$

Para calcular el efecto **aleatorio**: medimos las diferencias **DENTRO de cada grupo**

$$\sum \sum \left(X_{ij} - \overline{X}_{..}\right)^2 = \boxed{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - \overline{X}_{i.}\right)^2} + \sum_{i=1}^{k} n_i \left(\overline{X}_{i.} - \overline{X}_{..}\right)^2$$

Variabilidad **TOTAL** Variabilidad **DENTRO** Variabilidad **ENTRE**

ANOVA

Grupo 1	Grupo 2	Grupo 3	
5	9	8	
4	7	9	
6	8	5	
$X_{1.} = 5$	$X_{2.} = 8$	$X_{3.} = 7.33$	$X_{..} = 6.78$

Para calcular el **efecto del factor**: medimos las diferencias **ENTRE grupos**

$$\sum\sum\left(X_{ij} - \overline{X}_{..}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{i.}\right)^2 + \boxed{\sum_{i=1}^{k} n_i\left(\overline{X}_{i.} - \overline{X}_{..}\right)^2}$$

Variabilidad **TOTAL** Variabilidad **DENTRO** Variabilidad **ENTRE**

ANOVA

Tenemos dos tipos de ***variabilidades***:

- **ENTRE grupos (debida al factor)**
- **DENTRO grupos (debida a la aleatoriedad)**

Para poder afirmar que el factor produce efectos si:

La variabilidad **ENTRE** grupos es significativamente grande respecto a la variabilidad **DENTRO** grupos

Ratio de variabilidades

$$\boxed{F_{\exp} = \frac{ENTRE}{DENTRO}}$$

Comparación deshonesta

$$F_{\exp} = \frac{\sum_{i=1}^{k} n_i \left(\overline{X}_{i.} - \overline{X}_{..}\right)^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - \overline{X}_{i.}\right)^2}$$

Comparación mas honesta

$$\boxed{F_{\exp} = \frac{\dfrac{SC_{FACTOR}}{g.l._{FACTOR}}}{\dfrac{SC_{error}}{g.l._{error}}}}$$

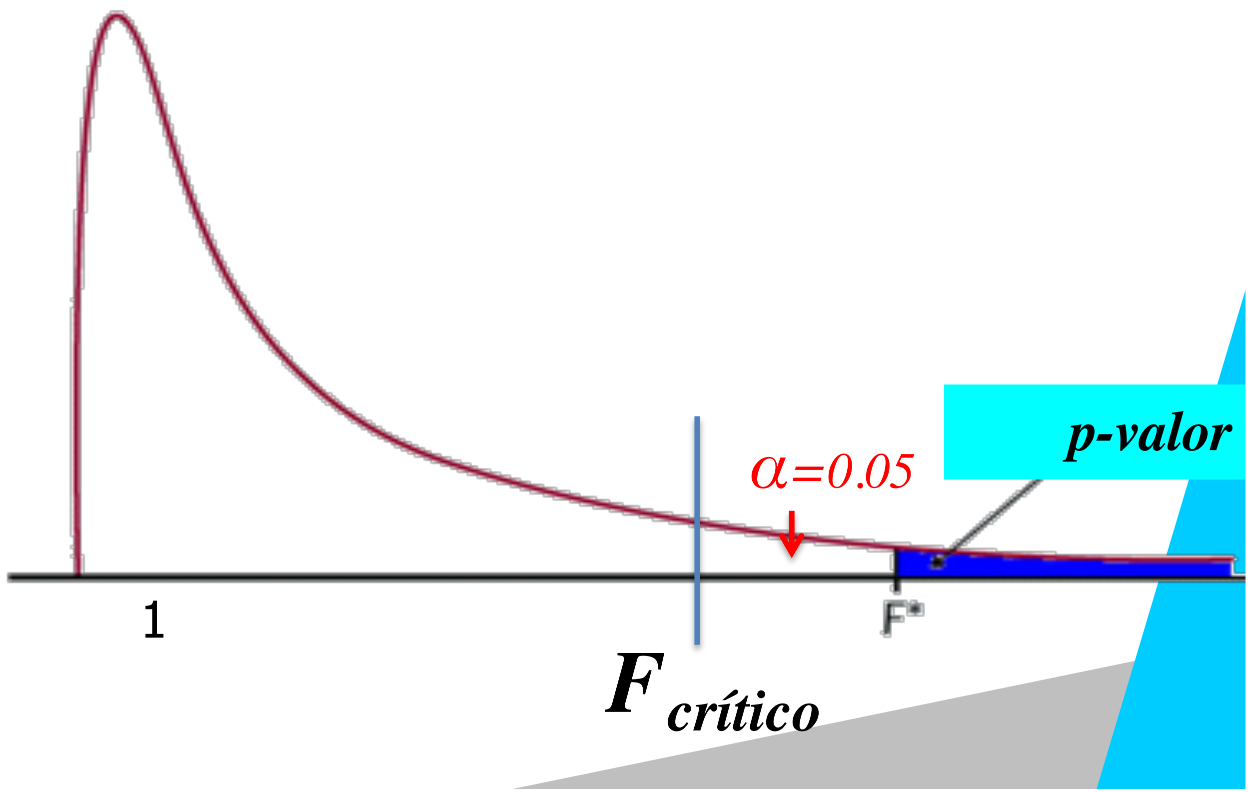
ANOVA

$$\left.\begin{array}{l} H_0: \mu_1 = \mu_2 = \ldots = \mu_k \\ H_1: \textit{Al menos una igualdad no es cierta} \end{array}\right\}$$

$$\left.\begin{array}{l} H_0 : \text{No hay efecto del Factor} \\ H_1 : \text{Hipótesis nula no cierta} \end{array}\right\} \quad \alpha=0.05$$

Estadístico para el contraste

$$F_{\exp} = \frac{\dfrac{SC_{FACTOR}}{g.l._{FACTOR}}}{\dfrac{SC_{error}}{g.l._{error}}} \leq F_{critico}$$



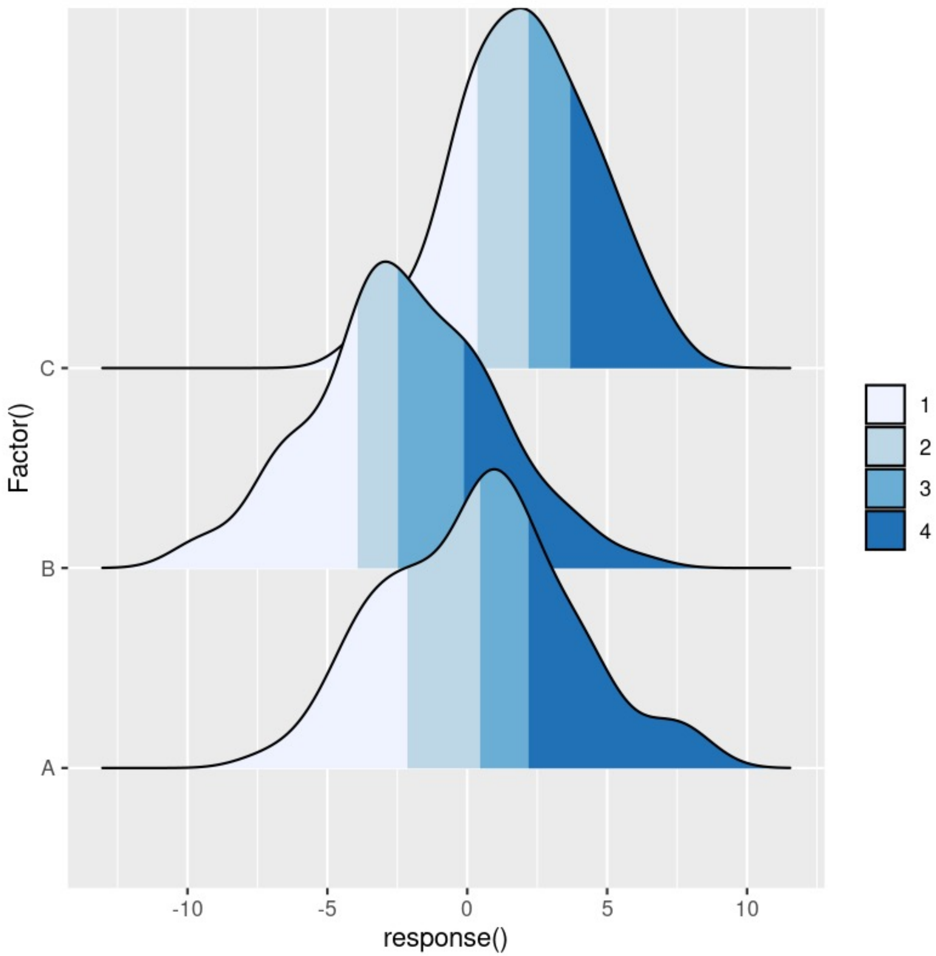
ANOVA

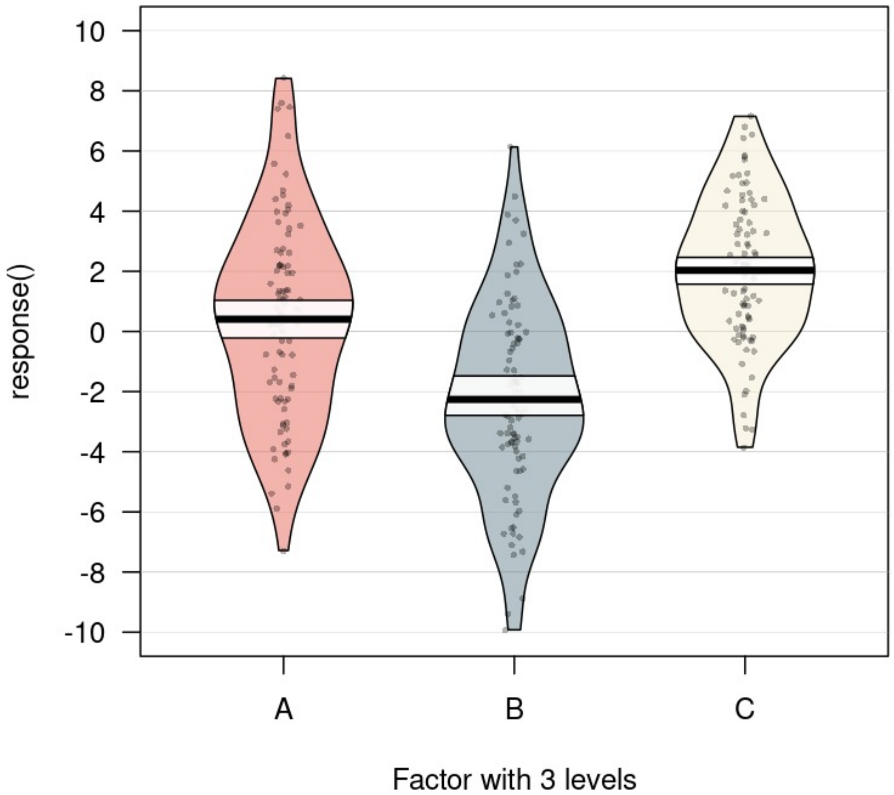
Generalizando …

	1	2	Niveles del factor	k
1	$X_{1,1}$	$X_{2,1}$	…	$X_{k,1}$
2	$X_{1,2}$	$X_{2,2}$	$X_{i,j}$	$X_{k,2}$
j	$X_{1,j}$	$X_{2,j}$	…	$X_{k,j}$
n	$X_{1,n1}$	$X_{2,n2}$	…	$X_{k,nk}$

$i = 1,2,3,\ldots,k$ $j = 1,2,3,\ldots, n_k$ (no balanceado)

Análisis de la varianza de un factor:





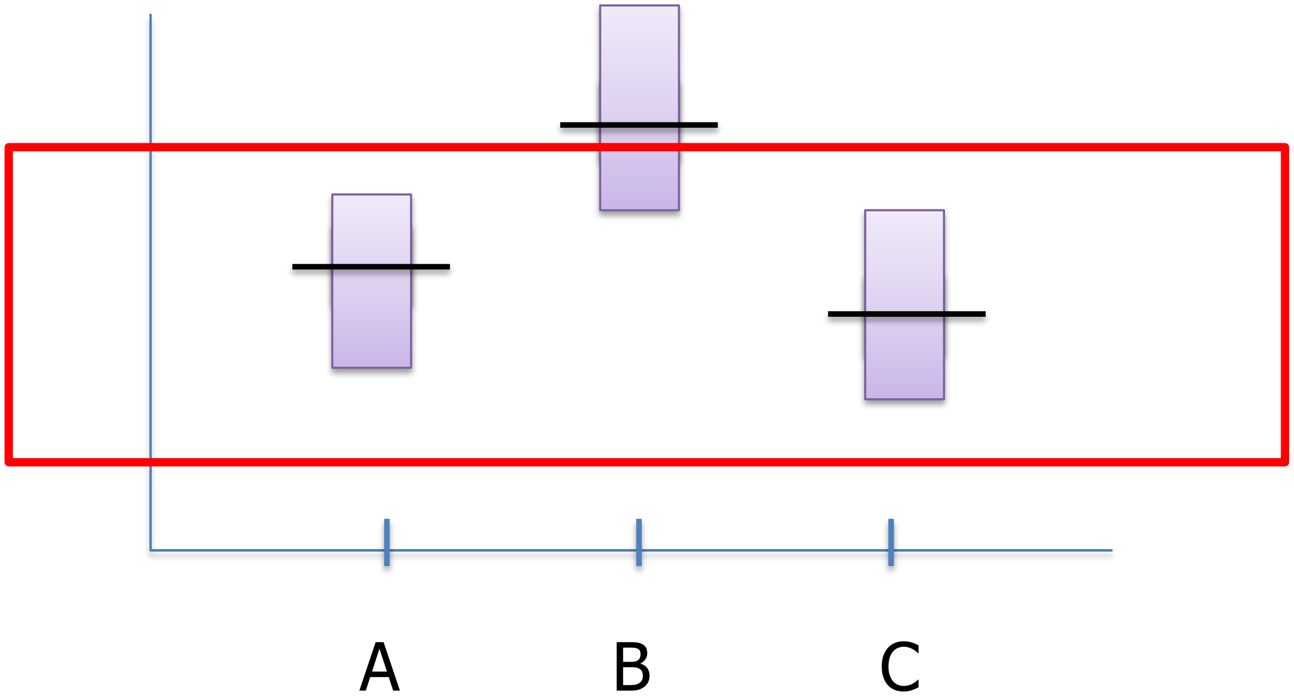
https://josezubcoff.shinyapps.io/Understanding_ANOVA

Para entender el comportamiento y sensibilidad de ANOVA

ANOVA

Fuentes de variación	Sumas de cuadrados	G.L.	Cuadrados Medios	F	Signif.
FACTOR	$Q_E = \sum_{i=1}^{k} n_i \left(X_{i.} - X_{..}\right)^2$	k–1	$Q_E / k-1 = S_E^2$	$F = \dfrac{Q_E / k-1}{Q_D / n-k}$	***p-valor***
Error aleatorio ó Residual	$Q_D = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - X_{i.}\right)^2 = \sum_{i=1}^{k} (n-1) S_i^2$	n–k	$Q_D / n-k = S_D^2$		
TOTAL	$Q = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(X_{ij} - X_{..}\right)$	n–1	$Q / n-1 = S^2$		

¿Que ocurre cuando el resultado de ANOVA es significativo?



$$\left.\begin{array}{l} H_0: \mu_A = \mu_B = \mu_C = \ldots \\ H_1: \textit{Al menos una igualdad no es cierta} \end{array}\right\}$$

ANOVA

Contrastes a posteriori

Análisis *a posteriori*

Dos tipos de definición de H_A

- *a priori* (antes de realizar el experimento)
- *a posteriori* (no propongo alternativas hasta haber realizado el experimento)

Ejemplo de comparación *a priori*:

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

H_A: $\mu_3 > \mu_1$; $\mu_3 > \mu_2$; $\mu_3 > \mu_4$

Los tests *a priori* son más potentes, pero están sometidos a mayor riesgo de error

Ej: Dunn Sidak

Tests *a posteriori*

- Comparan todos los posibles pares de medias entre sí

- Uno de los tests más utilizados es el de Student-Newman-Keuls (SNK)

- Otros tests:

  Scheffe
  Tukey HSD
  LSD
  Bonferroni
  SNK
  … etc.

Tests *a posteriori*

Test SNK:

Ejemplo: 5 hábitats (a = 5)
7 réplicas por hábitat (n = 7)

Hábitat	**a**	**b**	**c**	**d**	**e**
Media	6,4	7,1	3,8	2,6	4,1
S^2	2,18	0,67	1,77	1,35	2,06

Tabla de ANOVA

Fuente de var.	**SC**	**gl**	**MC**	F_{expt}	**p-valor**
Entre	99,26	4	24,82	15,38	***
Dentro	48,18	30	1,60		
Total	147,44	34			

Tests *a posteriori*

$ET = \sqrt{(MC_{DENTRO} / n)} = \sqrt{(1{,}60 / 7)} = 0{,}48$

Test SNK:

$Q_{ij} = X_i - X_j / ET$ está tabulado para H_0 verdadera

Rango	**1**	**2**	**3**	**4**	**5**		**test SNK**		
Medias	2,6	3,8	4,1	6,4	7,1	g	Q	D=QxET	
Comparaciones	**5-1** 4,5						5	4,10	1,97
	4-1 3,8	**5-2** 3,3					4	3,84	1,84
	3-1 1,5	**4-2** 2,6	**5-3** 3,0				3	3,49	1,68
	2-1 1,2	**3-2** 0,3	**4-3** 2,3	**5-4** 0,7			2	2,89	1,39

$\alpha = 0{,}05$
gl = 30

Tests *a posteriori*

$ET = \sqrt{(MC_{DENTRO} / n)} = \sqrt{(1,60 / 7)} = 0,48$

Test SNK:

$Q_{ij} = X_i - X_j / ET$ está tabulado para H_0 verdadera

Rango	**1**	**2**	**3**	**4**	**5**		**test SNK**		
Medias	2,6	3,8	4,1	6,4	7,1	g	Q	D=QxET	
Comparaciones	5-1 4,5						5	4,10	1,97
	4-1 3,8	5-2 3,3					4	3,84	1,84
	3-1 1,5	4-2 2,6	5-3 3,0				3	3,49	1,68
	2-1 1,2	3-2 0,3	4-3 2,3	5-4 0,7			2	2,89	1,39

$\alpha = 0,05$
gl = 30

Tests *a posteriori*

$ET = \sqrt{}(MC_{DENTRO} / n) = \sqrt{}(1{,}60 / 7) = 0{,}48$

Test SNK:

$Q_{ij} = X_i - X_j / ET$ está tabulado para H_0 verdadera

Rango	**1**	**2**	**3**	**4**	**5**		**test SNK**		
Medias	2,6	3,8	4,1	6,4	7,1	g	Q	D=QxET	
Comparaciones	5-1 **4,5***						5	4,10	1,97
	4-1 **3,8***	5-2 **3,3***					4	3,84	1,84
	3-1 1,5	4-2 **2,6***	5-3 **3,0***				3	3,49	1,68
	2-1 1,2	3-2 0,3	4-3 **2,3***	5-4 0,7			2	2,89	1,39

Si $(X_i - X_j) > D$, la diferencia es significativa (*)

$\alpha = 0{,}05$
gl = 30

Ordenamos para obtener subconjuntos homogéneos

5 > 1 5 > 2 5 > 3 5 = 4
4 > 1 4 > 2 4 > 3
3 = 1 3 = 2
2 = 1

$\underline{5 = 4} > \underline{3 = 2 = 1}$

Hay 2 Subconjuntos Homogéneos

Tests *a posteriori*

Presentación de resultados

Test SNK:

Un resultado tal que:

5 4 3 2 1

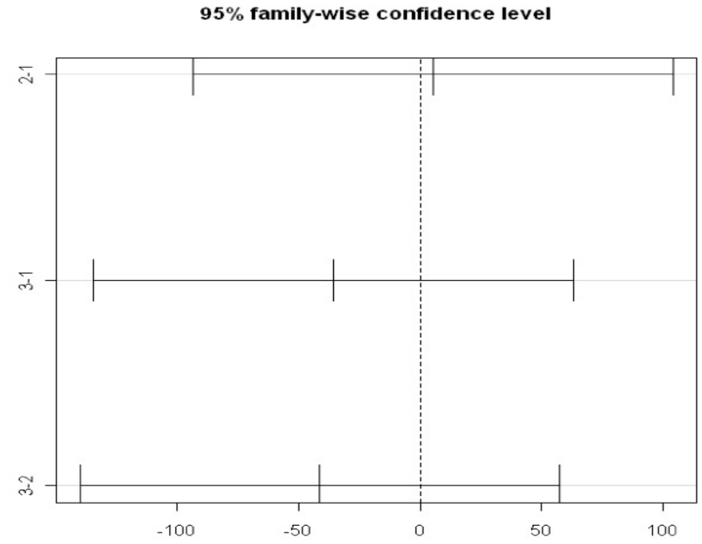
¿Hay Subconjuntos Homogéneos?

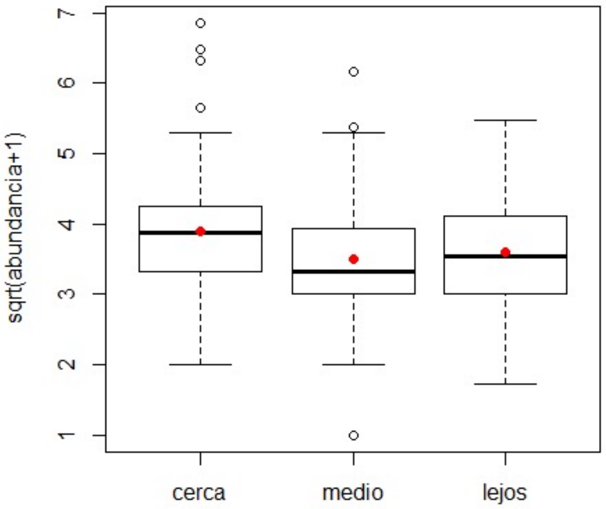
… no resulta lógico (no hay una H_A identificable) !!

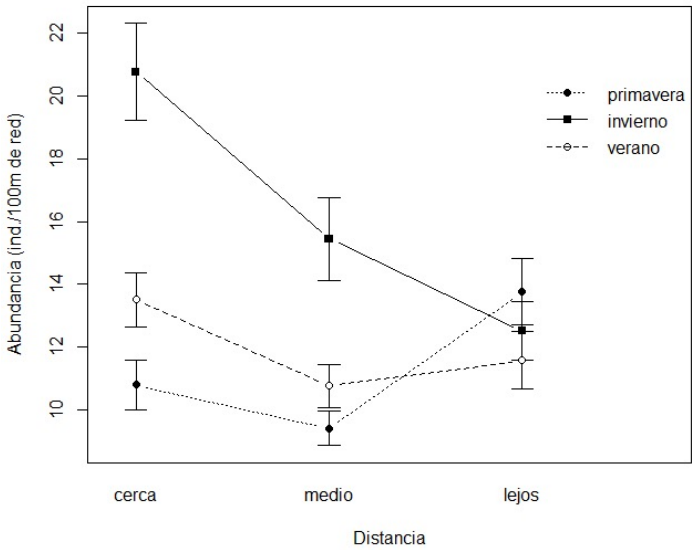
Tests *a posteriori*

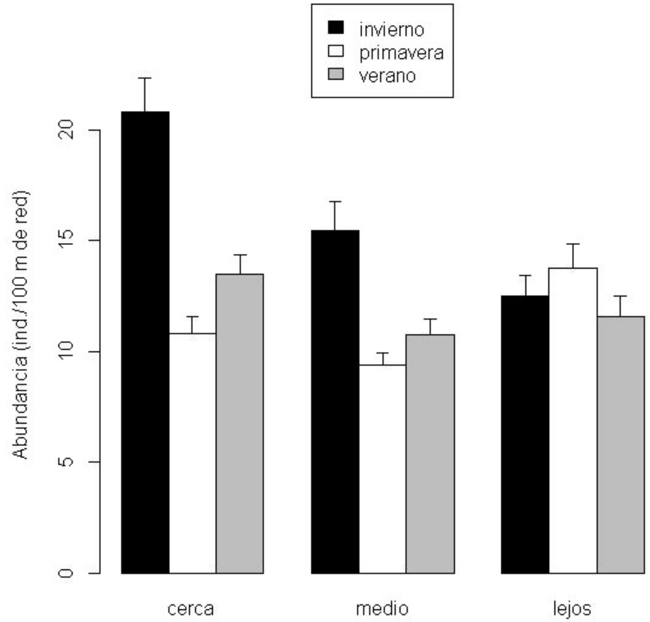
Presentación de resultados

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = datos$ABUND ~ factor(datos$ZONA))

$`factor(datos$ZONA)`
          diff        lwr       upr     p adj
2-1   5.462963  -93.32526 104.25118 0.9906082
3-1 -35.685185 -134.47340  63.10303 0.6696206
3-2 -41.148148 -139.93637  57.64007 0.5871203
```









ANOVA

Requisitos

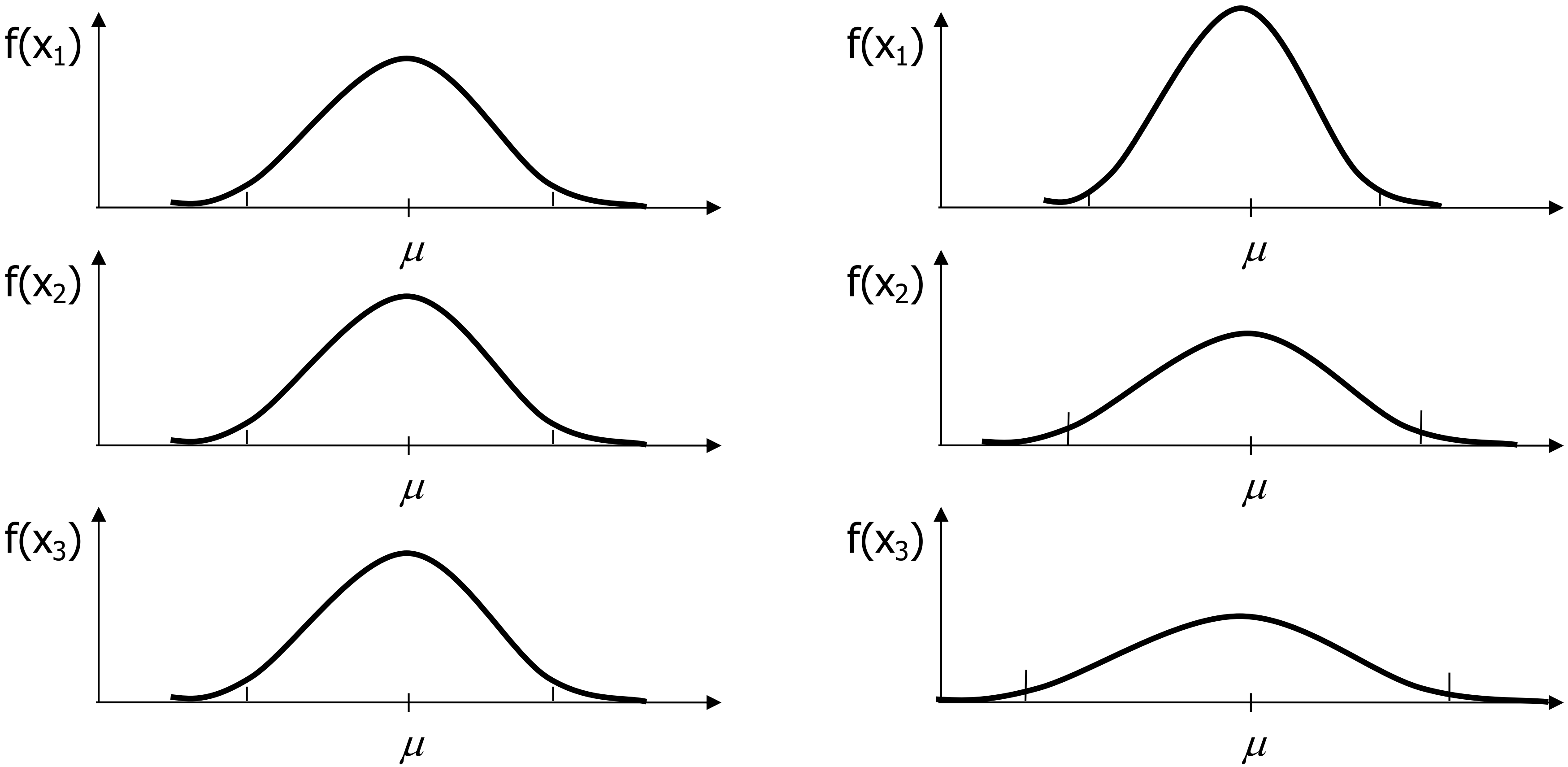
ANOVA

Hipótesis necesarias para realizar un ANOVA

a) Homogeneidad de las varianzas
b) Normalidad de la respuesta en cada nivel
c) Independencia de los valores obtenidos

Asunciones del análisis de la varianza:

Homogeneidad de varianzas:



Cuando varianzas ≠, se incrementa el error Tipo I

Asunciones del análisis de la varianza:

Homogeneidad de varianzas:

Test de heterogeneidad de varianzas:

Bartlett -> sensible a FALTA de normalidad
Levene -> es una ANOVA para VAR. Asume var iguales!!
Scheffe -> insensible a FALTA de normalidad, pero no lo recomnieda …
Hartley -> problema cuando 1 var es pequeña….
… etc.

Test de Cochran

$$C = \frac{\text{mayor } s_i^2}{\sum s_i^2}$$

Asunciones del análisis de la varianza:

<u>Homogeneidad de varianzas:</u>

Test de homogeneidad → Varianzas homogéneas → ANOVA

Test de homogeneidad → Varianzas no homogéneas → Transformación de datos → Test de homogeneidad

Conteos (o datos que siguen una dist. de Poisson) ➔ $\sqrt{(X + 1)}$

Ratios, tasas, concentraciones, etc. ➔ log (X) o log (X + 1)

Porcentajes y proporciones ➔ $\text{sen}^{-1} \sqrt{X}$ (= arcsen X)

Asunciones del análisis de la varianza:

Si NO hay homocedasticidad (aún transformando datos)

Si el experimento está bien replicado (n>30 y balanceado) =>

ANOVA es suficientemente robusto

(*) References

Glass, G.V., P.D. Peckham, and J.R. Sanders. 1972. Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance. Rev. Educ. Res. 42: 237-288.

Harwell, M.R., E.N. Rubinstein, W.S. Hayes, and C.C. Olds. 1992. Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases. J. Educ. Stat. 17: 315-339.

Lix, L.M., J.C. Keselman, and H.J. Keselman. 1996. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. Rev. Educ. Res. 66: 579-619.

Schmider, Emanuel; Ziegler, Matthias; Danay, Erik; Beyer, Luzi; Bühner, Markus. 2010. *Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption*.

Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, Vol 6(4), 2010, 147-151. doi: 10.1027/1614-2241/a000016

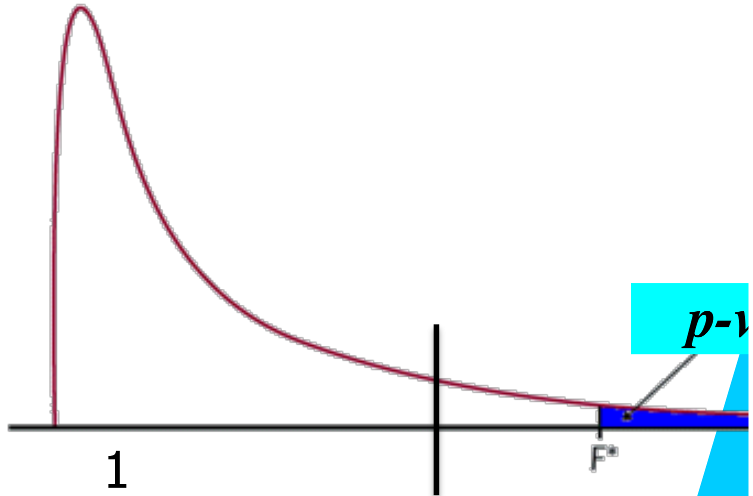
Asunciones del análisis de la varianza:

Si la transformación de datos no corrige la Heterocedasticidad

Se usa α = 0.01 (se minimiza el error tipo I)

Se interpreta el resultado (siempre es mas conservador (dado *α = 0.01)*

Utilizar un test **no paramétrico** no soluciona el problema de heterogeneidad de varianzas



Por tanto:

- Debemos intentar **explicar** dicha heterogeneidad
- Podemos usar ANOVA (o un test no paramétrico)

(*) References

Glass, G.V., P.D. Peckham, and J.R. Sanders. 1972. Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance. Rev. Educ. Res. 42: 237-288.

Harwell, M.R., E.N. Rubinstein, W.S. Hayes, and C.C. Olds. 1992. Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases. J. Educ. Stat. 17: 315-339.

Lix, L.M., J.C. Keselman, and H.J. Keselman. 1996. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. Rev. Educ. Res. 66: 579-619.

Schmider, Emanuel; Ziegler, Matthias; Danay, Erik; Beyer, Luzi; Bühner, Markus. 2010. *Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption*.

Asunciones del análisis de la varianza:

Homogeneidad de varianzas:

Test de homogeneidad

Varianzas homogéneas

ANOVA

Varianzas no homogéneas

Transformación de datos

Si la transformación de datos no corrige la Heterocedasticidad

Podemos usar ANOVA pero:

Si se acepta H_0 => explicamos la heterogeneidad de varianzas!!

Si se rechaza H_0, ($\alpha = 0.01$) => explicamos la heterogeneidad de varianzas!!

Utilizar un test no paramétrico (p.ej. Kruskal-Wallis) no soluciona el problema!
Se debería explicar además la heterogeneidad de varianzas

Asunciones del análisis de la varianza:

Transformación de datos

- Útil para eliminar heterogeneidad de la varianza.
- Sólo es efectivo si la media tiene una relación constante con la varianza.
- La transformación debe ser monotónica.
- Deben mantenerse las medias en el mismo orden.
- La transformación debe utilizarse únicamente para evitar el problema de heterogeneidad de la varianza.
- Transformaciones sistemáticas son perjudiciales.

Asunciones del análisis de la varianza:

Transformación de datos

Raíz cuadrada

- Poblaciones que siguen una **distribución de Poisson**: medias y varianzas son iguales
- Frecuencias o recuentos por unidad de tiempo o superficie.
- Principalmente con abundancias muy pequeñas.

$$\sqrt{X + 1}$$

Asunciones del análisis de la varianza:

Transformación de datos

Logarítmo

- Muestreos con valores muy altos: medias mayores y varianza mucho mayores.
- Datos distribuidos log-normal
- Medidas de tasas, concentraciones, relaciones,...
- Ej: Relación entre el número de presas comida por depredador, cantidad de clorofila por peso algal,...
- Independiente del tipo de logaritmo usado.
- Sumar una constante (1) para aplicar logaritmos por los valores que son 0.
- Problema en valores muy pequeños: solo cuando son mayores de 10.

$$\log (X+1)$$

Asunciones del análisis de la varianza:

Transformación de datos

Arcoseno

- **Porcentajes y proporciones**
- Distribución Binomial.
- Ej. Porcentaje de cobertura de *Posidonia*

$$\mathrm{sen}^{-1} \sqrt{X}$$

Asunciones del análisis de la varianza:

Transformación de datos

logit transformation

- Transforma los valores de **porcentajes** desde - ∞ hasta + ∞

- La transformación de arcoseno limita los valores desde 0 hasta π / 2 radianes (0 hasta 90 º) .

$$(1/2) \log ((p / 1-p))$$

Asunciones del análisis de la varianza:

Normalidad de los datos

ANOVA es suficientemente robusto a las desviaciones de la normalidad (Glass et al. 1972, Harwell et al. 1992, Lix et al. 1996)*, sobre todo cuando:

- hay un gran número de réplicas (n>30)
- los datos están equilibrados (balanceados)

Las transformaciones a menudo corrigen el apuntamiento, pero cuidado, solo usar para corregir heterogeneidades de varianzas.

(*) References

Glass, G.V., P.D. Peckham, and J.R. Sanders. 1972. Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance. Rev. Educ. Res. 42: 237-288.

Harwell, M.R., E.N. Rubinstein, W.S. Hayes, and C.C. Olds. 1992. Summarizing Monte Carlo results in methodological research: the one- and two-factor fixed effects ANOVA cases. J. Educ. Stat. 17: 315-339.

Lix, L.M., J.C. Keselman, and H.J. Keselman. 1996. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. Rev. Educ. Res. 66: 579-619.

Schmider, Emanuel; Ziegler, Matthias; Danay, Erik; Beyer, Luzi; Bühner, Markus. 2010. *Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption*.

Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, Vol 6(4), 2010, 147-151. doi: 10.1027/1614-2241/a000016

Requisitos de ANOVA

- INDEPENDENCIA de los valores observados:

Problemas de la FALTA de independencia de datos

Correlación positiva

Cuando las observaciones se PARECEN más de lo debido a su propia variabilidad

Correlación negativa

Cuando las observaciones se DISPERSAN más de lo debido a su propia variabilidad

Resumen del **problema** de la FALTA de independencia de datos

Ejemplo: evaluar el Factor Método de enseñanza (3 niveles)

¿Qué pasa si no se vigila el examen? ¿Las notas reflejan realmente el conocimiento del alumnado?

¿Se parecen más las notas dentro un aula? (amenaza de **correlación positiva dentro)**

¿Y si dejamos usar un dispositivo móvil conectado? Se parecen más (entre sí) TODAS las notas? (amenaza de **correlación positiva ENTRE los niveles y dentro de cada nivel)**

Decisiones que podríamos adoptar para lidiar con estas amenazas a la independencia…

Resumen del **problema** de la FALTA de independencia de datos

Ejemplo: evaluar el Factor Método de enseñanza (3 niveles)

¿Qué pasa si además del método de enseñanza, la persona que imparte la clase es distinta?

¿Puede influir la persona que imparte la clase? (amenaza de **correlación negativa ENTRE**)

Decisiones que podríamos adoptar para lidiar con esta amenaza a la independencia…

Resumen del **problema** de la FALTA de independencia de datos

Ejemplo: evaluar el Factor Método de enseñanza (3 niveles)

¿Qué pasa si en cada clase usamos 3 TIPOS de examenes para evitar copiarse?

¿Puede influir el TIPO de examen? (amenaza de **correlación negativa DENTRO de cada clase**)

Decisiones que podríamos adoptar para lidiar con esta amenaza a la independencia…

Resumen del problema de la FALTA de independencia de datos

No independencia DENTRO de los tratamientos	No independencia ENTRE los tratamientos

Correlación positiva

σ^2_e dentro de las muestras es subestimado	σ^2_e entre de las muestras es subestimado
F ratio excesivo	F ratio muy pequeno
Incremento del error Tipo I	Incremento del error Tipo II
Diferencias sin importancia son detectadas	Diferencias reales no son detectadas

Correlación negativa

σ^2_e dentro de las muestras es sobreestimado	σ^2_e entre de las muestras es sobreestimado
F ratio demasiado pequeno	F ratio excesivo
Incremento del error Tipo II	Incremento del error Tipo I
Diferencias reales no son detectadas	Diferencias sin importancia son detectadas

Esquema para apuntar amenazas a la independencia de datos

No hay independencia DENTRO de los tratamientos

No hay independencia ENTRE los tratamientos

Correlación positiva

Correlación negativa

Análisis de las amenazas a la Independencia de las observaciones

Trabajo grupal

Proponer un experimento real y explicar la metodología que garantice la independencia de los valores observados

- Definir el **objetivo** del estudio
- Definir la **variable** a medir
- Definir el **Factor** (un factor principal) y sus **niveles**
- Definir **metodología de muestreo** (debes explicarlo como si tuvieras un equipo de varias personas que realizarán el muestreo)
- Estudiar y **enumerar las posibles amenazas a la independencia** de valores observados
- Definir una **estrategia para lidiar con las amenazas** antes descritas
    - Definir un nuevo factor con sus niveles que permitirá conocer la variabilidad achacable a dicho factor
    - Fijar solo un nivel de ése factor (cuando decidas que el resto de niveles del factor no son de interés para este estudio). Las conclusiones de tu factor principal se ajustarán a esta realidad.
    - ~~No hacer nada~~

Análisis de las amenazas a la Independencia de las observaciones

Trabajo grupal

Proponer un experimento real y explicar la metodología que garantice la independencia de los valores observados

- **Objetivo:**
- **Variable:**
- **Factor** y **niveles:**
- **Metodología de muestreo:**

- **Enumerar las posibles amenazas a la independencia** de valores observados

- **Estrategia para lidiar con las amenazas (una para cada amenaza):**

ANOVA

Introducción, conceptos fundamentales y "como funciona por dentro"

Universitat d'Alacant
Universidad de Alicante
Departamento de Ciencias del Mar y Biología Aplicada
Jose Jacobo Zubcoff Vallejo