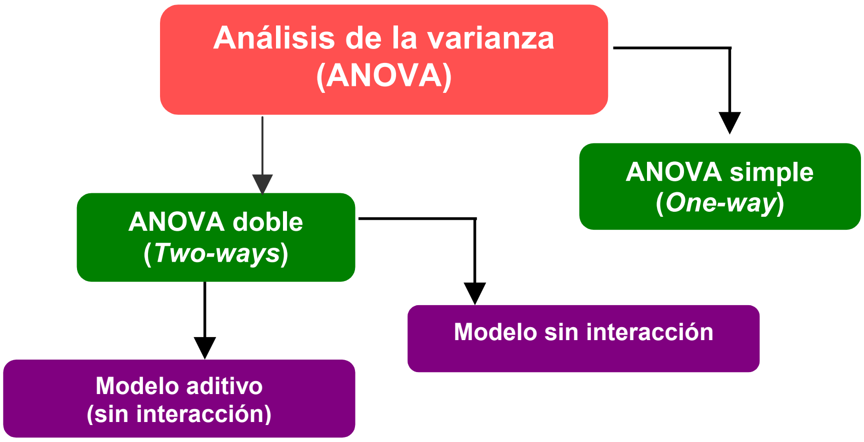
ANÁLISIS DE LA VARIANZA (ANOVA)

Autores: Manuel Terrádez (mterradez@uoc.edu), Ángel A. Juan (ajuanp@uoc.edu)

ESQUEMA DE CONTENIDOS



INTRODUCCIÓN

Usaremos el **análisis de la varianza (ANOVA)** para contrastar la hipótesis nula de que las medias de distintas poblaciones coinciden. Por ejemplo, en el caso de 5 poblaciones, el contraste a realizar sería:

$H_O: \mu_1 = \mu_2 = ... = \mu_5$ vs. H_A: no todas las medias poblacionales son iguales

En *math-block* EST-I16 se estudia cómo se utiliza la distribución t-Student (o la Normal) para contrastar la hipótesis nula de que dos medias poblacionales coinciden. Usando esta técnica, podríamos realizar los siguientes 10 tests para contrastar la hipótesis nula anterior:

$H_{01}: \mu_1 = \mu_2$
$H_{02}: \mu_2 = \mu_3$
$H_{03}: \mu_3 = \mu_4$
$H_{04}: \mu_4 = \mu_5$
$H_{05}: \mu_1 = \mu_3$
$H_{06}: \mu_2 = \mu_4$
$H_{07}: \mu_3 = \mu_5$
$H_{08}: \mu_1 = \mu_4$
$H_{09}: \mu_2 = \mu_5$
$H_{010}: \mu_1 = \mu_5$

En este caso, rechazar cualquiera de las 10 hipótesis nulas implicaría rechazar la hipótesis nula inicial de que las cinco medias coinciden. Por el contrario, si no rechazásemos ninguna de las 10 hipótesis, tampoco rechazaríamos la hipótesis inicial. El problema de este método es doble: (1) por un lado, se requiere de un mayor esfuerzo computacional, y (2) por otro, al hacer un mayor número de contrastes aumenta el error de tipo I (la probabilidad de rechazar la hipótesis nula siendo ésta cierta). El uso de las técnicas ANOVA nos permiten eludir ambos problemas.

Financiado por la Secretaría de Estado de Educación y Universidades (MECD)

OBJETIVOS

- Entender qué es y por qué es importante un contraste de hipótesis.
- Saber distinguir en qué situaciones es útil realizar un análisis de la varianza.
- Conocer pautas para elegir el modelo más adecuado para nuestro problema.
- Saber aplicar el ANOVA, con ayuda de Minitab.
- Interpretar los distintos valores que aparecen en una tabla ANOVA.
- Dar respuesta al problema del contraste de hipótesis planteado.

CONOCIMIENTOS PREVIOS

Aparte de estar iniciado en el uso del paquete estadístico Minitab, resulta muy conveniente haber leído con profundidad los siguientes *math-blocks*:

- Estadística descriptiva.
- Intervalos de confianza y contraste de hipótesis para 1 población.
- Intervalos de confianza y contraste de hipótesis sobre 2 poblaciones.
- Muestreo en poblaciones finitas.

CONCEPTOS FUNDAMENTALES

Análisis simple de la varianza (*One-Way* ANOVA)

El objetivo principal de muchos experimentos consiste en determinar el efecto que sobre alguna variable dependiente Y tienen distintos niveles de algún factor X (variable independiente y discreta). El factor puede ser la temperatura, la empresa que ha producido el bien, el día de la semana, etc.
Esencialmente, el diseño para el análisis simple de la varianza consistirá en obtener muestras aleatorias e independientes del valor de Y asociado a cada uno de los distintos niveles del factor X_1, X_2,..., Xn . Entonces podremos determinar si los diferentes niveles del factor tienen un efecto significativo sobre el valor de la variable dependiente.

El funcionamiento de la técnica ANOVA simple es, a grandes rasgos, el siguiente: a fin de comparar las medias de Y asociadas a los distintos niveles del factor (X_1, X_2,..., Xn), compararemos una medida de la **variación entre diferentes niveles (MS-factor)** con una medida de la **variación dentro de cada nivel (MS-error)**. Si el MS-factor es significativamente mayor que el MS-error, concluiremos que las medias asociadas a diferentes niveles del factor son distintas. Esto significa que el factor influye significativamente sobre la variable dependiente Y. Si, por el contrario, el MS-factor no es significativamente mayor que el MS-error, no rechazaremos la hipótesis nula de que todas las medias, asociadas a diferentes niveles del factor, coinciden.

Supuestos

De forma similar a lo que ocurre con la regresión lineal, aquí también hay un modelo para los datos. El modelo asociado al i-ésimo nivel del factor X será:

$$Y = \mu_i + \varepsilon$$

donde:

- Los errores ε están normalmente distribuidos con media 0
- Los errores ε son independientes
- Los errores ε tienen varianza constante σ^2

Para verificar estos supuestos suele ser útil realizar un gráfico que muestre la distribución de las observaciones por niveles: si en el gráfico se aprecian diferencias entre niveles por lo que a la variación de las observaciones se refiere, es muy probable que tengamos un problema con el supuesto de **varianza constante**; si aparecen "outliers", puede que no se cumpla el supuesto de **normalidad**; por otra parte, si el tiempo fuese un factor importante a la hora de registrar observaciones, podría ocurrir que observaciones consecutivas estuviesen correlacionadas, con lo que no se cumpliría el supuesto de **independencia**.

Análisis doble de la varianza (*Two-Way* ANOVA)

Usaremos el **análisis doble de la varianza** para estudiar los posibles efectos causados por diferentes niveles de dos factores sobre la variable dependiente. Así, por ejemplo, en agricultura estaremos interesados en estudiar qué efectos tendrán, sobre el crecimiento de las patatas, variaciones en los niveles de potasio y nitrógeno de la tierra; en medicina, estaremos interesados en estudiar los efectos, sobre el dolor de cabeza, del medicamento y de la dosis empleados; en educación, buscaremos conocer qué efectos, sobre el tiempo necesario para adquirir unos conocimientos, tendrán los factores nivel de estudios y sexo; en una campaña de marketing, estaremos interesados en conocer los efectos del presupuesto y del medio usado (televisión, revistas, ...) sobre las ventas; etc.

Usaremos ANOVA doble para contrastar, para cada uno de los dos factores, la hipótesis nula de que el resultado de la variable dependiente (crecimiento de patatas, intensidad del dolor de cabeza, tiempo en adquirir conocimientos, ventas, etc.) no depende del factor.

Modelo aditivo (sin interacción)

El **modelo aditivo** supone que la variación total en los datos puede ser expresada como suma de variaciones procedentes de fuentes diversas:

(Variación total en los datos) = (Variación debida al primer factor) + (Variación debida al segundo factor) + (Variación debida al error aleatorio)

En el modelo anterior, si la variación debida al primer factor fuese mucho mayor que la variación debida al error aleatorio, dispondríamos de evidencia estadística contra la hipótesis nula de que los distintos niveles del primer factor tienen el mismo efecto sobre la variable dependiente; de forma similar, si la variación debida al segundo factor fuese mucho mayor que la variación debida al error aleatorio, deberíamos rechazar la hipótesis nula de que la variable dependiente no depende de los diversos niveles del segundo factor.

Modelo con interacción

Un aspecto al que debemos prestar especial atención es el **nivel de interacción entre ambos factores**; es decir, el efecto que cada uno de los factores tiene sobre el otro. Así, por ejemplo, en la campaña de marketing citada anteriormente, el incremento en las ventas debido a un aumento del presupuesto podría ser el mismo independientemente del medio usado, o bien podría variar dependiendo del medio usado.

En este último caso deberíamos estudiar también la interacción entre los factores presupuesto y medio. Para ello usaremos los llamados **modelos con interacción**.

Este modelo supone que la variación total de los datos puede descomponerse de la siguiente forma:

(Variación total en los datos) = (Variación debida al primer factor) + (Variación debida al segundo factor) + (Variación debida a la interacción entre factores) + (Variación debida al error aleatorio)

CASOS PRÁCTICOS CON SOFTWARE

Inflamabilidad de pijamas

La inflamabilidad de los pijamas para niños ha sido un tema de preocupación constante durante las últimas décadas. Hoy en día, hay toda una serie de controles de seguridad que garantizan que las telas con que se fabrican los pijamas no sean fácilmente inflamables.

Tras seleccionar un determinado fabricante de pijamas y 5 laboratorios diferentes, hemos enviado a cada laboratorio 11 prendas de dicho fabricante. La idea es que en cada laboratorio se les aplique un test de inflamabilidad. Los resultados (expresados en un determinado índice de inflamabilidad) se muestran a continuación:

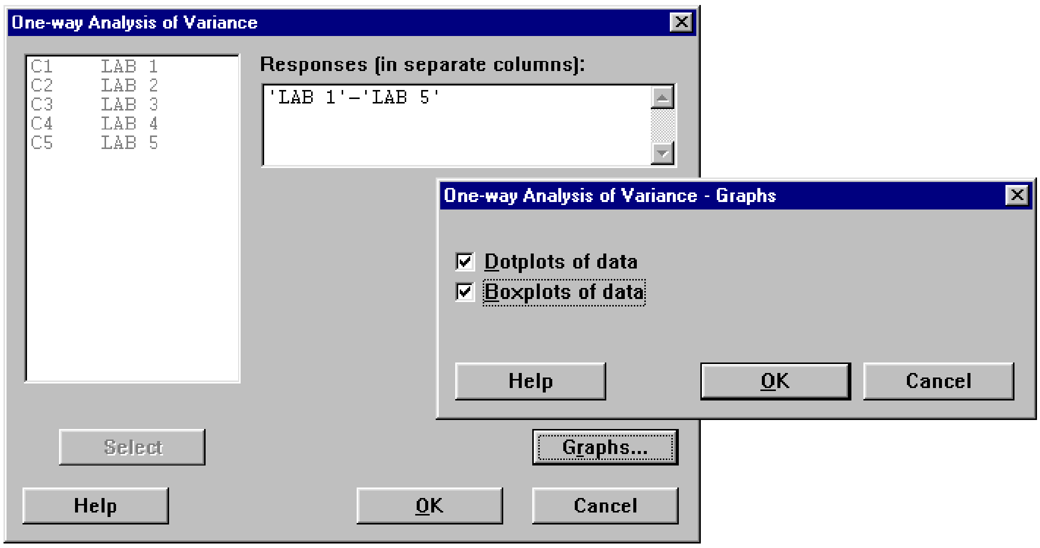
Worksheet 1 ***

	C1	C2	C3	C4	C5	C6
↓	LAB 1	LAB 2	LAB 3	LAB 4	LAB 5	
1	2,9	2,7	3,3	3,3	4,1	
2	3,1	3,4	3,3	3,2	4,1	
3	3,1	3,6	3,5	3,4	3,7	
4	3,7	3,2	3,5	2,7	4,2	
5	3,1	4,0	2,8	2,7	3,1	
6	4,2	4,1	2,8	3,3	3,5	
7	3,7	3,8	3,2	2,9	2,8	
8	3,9	3,8	2,8	3,2	3,5	
9	3,1	4,3	3,8	2,9	3,7	
10	3,0	3,4	3,5	2,6	3,5	
11	2,9	3,3	3,8	2,8	3,9	
12						

Nuestro objetivo será determinar si las medias obtenidas por cada laboratorio son aproximadamente iguales (es decir, pretendemos saber si hay o no diferencias significativas entre los laboratorios a la hora de determinar la capacidad de inflamación de una determinada prenda).

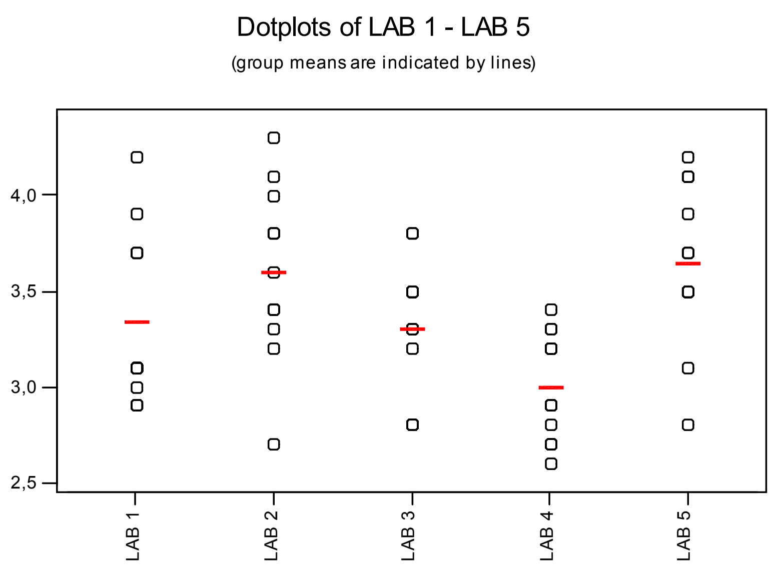
Stat → Anova → One-way (Unstacked)*...*

Le pediremos al programa que nos aplique la técnica ANOVA simple y, además, que nos represente un diagrama de puntos y un *boxplot* de los datos.



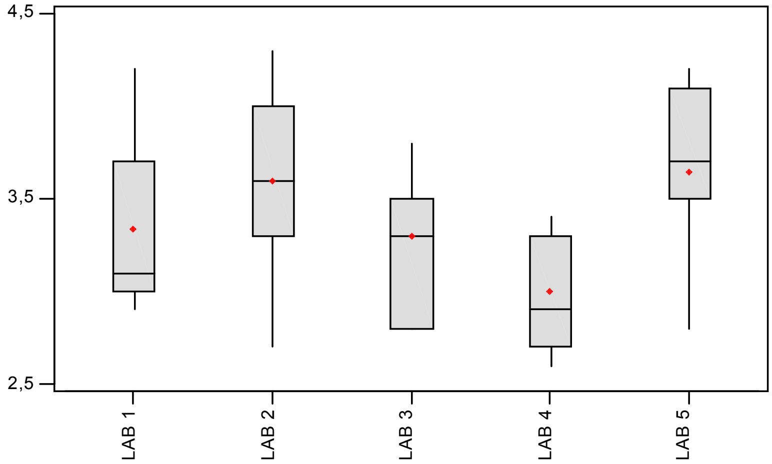
En las dos siguientes gráficas, podemos observar la **variación dentro de cada grupo** (laboratorio) y **la variación entre diferentes grupos** (laboratorios). La pregunta que habría que responder es: ¿resulta la variación entre diferentes grupos <u>significativamente</u> mayor que la variación existente dentro de los grupos?

Notar que el laboratorio 4 parece tener índices mayoritariamente bajos, mientras que los grupos 2 y 5 presentan índices bastante mayores.



Boxplots of LAB 1 - LAB 5

(means are indicated by solid circles)



En el cuadro siguiente se presenta el “output” numérico del programa. La primera parte del mismo es la llamada tabla ANOVA. Entre otros datos importantes (como los MS-factor y MS-error, el valor del estadístico de contraste F = (MS-factor)/(MS-error), los grados de libertad, etc.), ésta nos proporciona el p-valor del contraste anterior. En este caso, el p-valor = 0,003 y, por tanto, rechazaremos la hipótesis nula de que todas las medias son iguales.

A partir del gráfico que proporciona los intervalos de confianza (a un nivel del 95%) para la media, parece obvio que, al menos, las medias procedentes de los laboratorios 4 (3,000) y 5 (3,6455) son significativamente diferentes:

One-way Analysis of Variance

```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      4     2,987     0,747     4,53    0,003
Error      50     8,233     0,165
Total      54    11,219
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
LAB 1      11    3,3364    0,4523            (------*------)
LAB 2      11    3,6000    0,4604                    (------*------)
LAB 3      11    3,3000    0,3715           (------*------)
LAB 4      11    3,0000    0,2864   (------*------)
LAB 5      11    3,6455    0,4321                     (------*------)
                                   --+---------+---------+---------+----
Pooled StDev =   0,4058            2,80      3,15      3,50      3,85
```

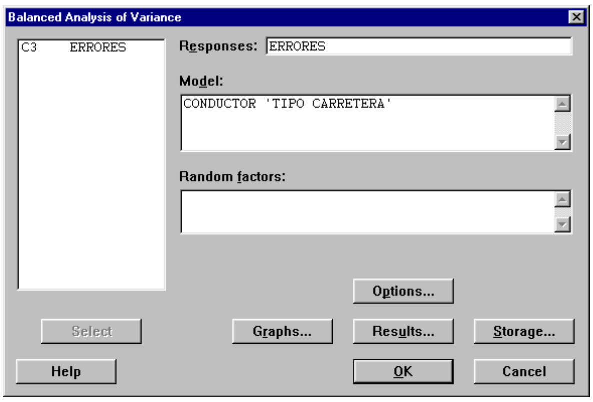
Conducción de vehículos

A continuación se muestran los datos obtenidos en un experimento en el que se comprobaron las habilidades de dos grupos de conductores, los inexpertos y los expertos. Doce conductores de cada grupo tomaron parte en el experimento. Se usaron tres tipos de carreteras: autopista, nacional y comarcal. Mediante un proceso aleatorio, se asignaron a cada tipo de carretera cuatro conductores expertos y cuatro inexpertos. Cada conductor estuvo al volante durante 2 kilómetros, en los cuales se registraron los siguientes "errores de conducción" cometidos:

Row	CONDUCTOR	TIPO CARRETERA	ERRORES
1	Inexperto	Autopista	4
2	Inexperto	Autopista	18
3	Inexperto	Autopista	8
4	Inexperto	Autopista	10
5	Experto	Autopista	6
6	Experto	Autopista	4
7	Experto	Autopista	13
8	Experto	Autopista	7
9	Inexperto	Nacional	23
10	Inexperto	Nacional	15
11	Inexperto	Nacional	21
12	Inexperto	Nacional	13
13	Experto	Nacional	2
14	Experto	Nacional	6
15	Experto	Nacional	8
16	Experto	Nacional	12
17	Inexperto	Comarcal	16
18	Inexperto	Comarcal	27
19	Inexperto	Comarcal	23
20	Inexperto	Comarcal	14
21	Experto	Comarcal	20
22	Experto	Comarcal	15
23	Experto	Comarcal	8
24	Experto	Comarcal	17

Plantearemos un ANOVA doble para contrastar, para cada uno de los dos factores, la hipótesis nula de que el número de errores de conducción cometidos no depende del factor; utilizaremos el modelo aditivo.

Stat → Anova → Balanced Anova...



```
Analysis of Variance (Balanced Designs)

Factor      Type Levels Values
CONDUCTO  fixed      2   Experto Inexperto
TIPO CAR  fixed      3 Autopista  Comarcal  Nacional


Analysis of Variance for ERRORES

Source       DF         SS         MS        F       P
CONDUCTO      1     228,17     228,17     8,56  0,008
TIPO CAR      2     308,33     154,17     5,78  0,010
Error        20     533,33      26,67
Total        23    1069,83
```

En el "output" anterior, cabe destacar los p-valores asociados a cada factor. En este caso, ambos son bastante pequeños (y, por tanto, significativos), por lo que deberemos rechazar las hipótesis nulas asociadas a cada factor; es decir, los datos demuestran que tanto el tipo de carretera como la experiencia del conductor son factores que influyen decisivamente en el número de errores de conducción cometidos.

Calidad en pastelería

Se ha llevado a cabo un experimento para determinar los efectos de dos ingredientes, harina y azúcar (factores), sobre la calidad final de un pastel.

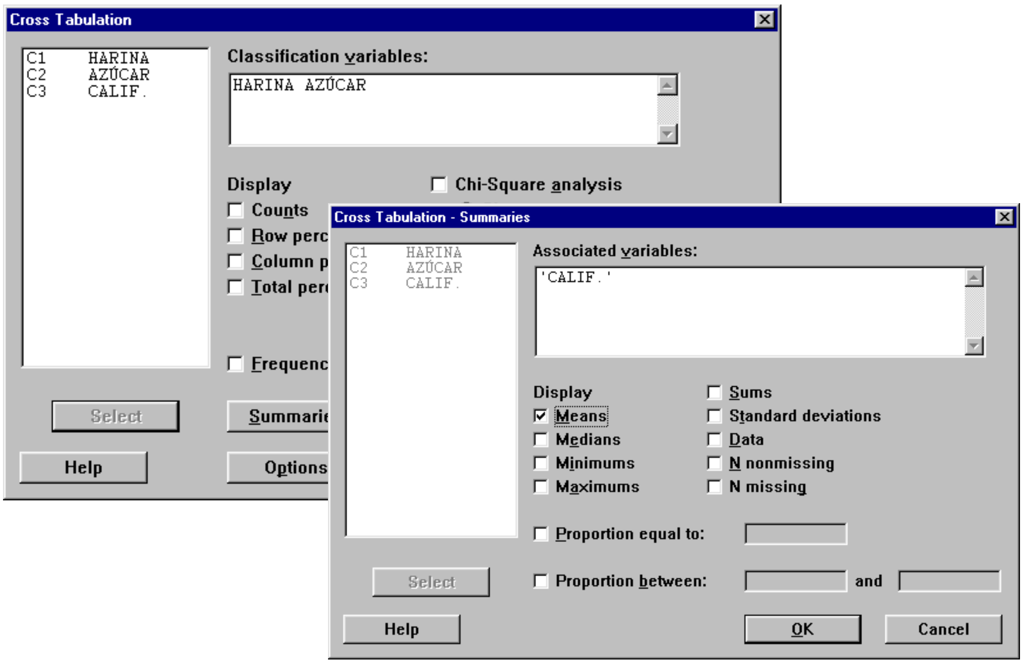
Hay cuatro niveles para la cantidad empleada de harina (0%, 10%, 20% y 30%), y dos niveles para la cantidad empleada de azúcar (1 = con azúcar, y 0 = sin azúcar). Para cada una de las 8 posibles combinaciones, se elaboraron 3 pasteles.

Cada uno de los 24 pasteles fue calificado por el mismo experto con una nota entre 3 (flojo) y 6 (excelente). A continuación se muestran los resultados:

Row	HARINA	AZÚCAR	CALIF.
1	0	0	4,4
2	0	0	4,5
3	0	0	4,3
4	0	1	3,3
5	0	1	3,2
6	0	1	3,1
7	10	0	4,6
8	10	0	4,5
9	10	0	4,8
10	10	1	3,8
11	10	1	3,7
12	10	1	3,6
13	20	0	4,5
14	20	0	4,8
15	20	0	4,8
16	20	1	5,0
17	20	1	5,3
18	20	1	4,8
19	30	0	4,6
20	30	0	4,7
21	30	0	5,1
22	30	1	5,4
23	30	1	5,6
24	30	1	5,3

En primer lugar, mostraremos en forma tabular las medias asociadas a cada una de las 8 combinaciones posibles. Ello nos permitirá saber si hay o no interacción entre ambos factores:

Stat → Tables → Cross Tabulation...




```
Tabulated Statistics

 Rows: HARINA      Columns: AZÚCAR

           0        1       All

 0     4,4000   3,2000   3,8000
 10    4,6333   3,7000   4,1667
 20    4,7000   5,0333   4,8667
 30    4,8000   5,4333   5,1167
 All   4,6333   4,3417   4,4875

  Cell Contents --
           CALIF.:Mean
```

Analicemos el efecto promedio de cada factor:

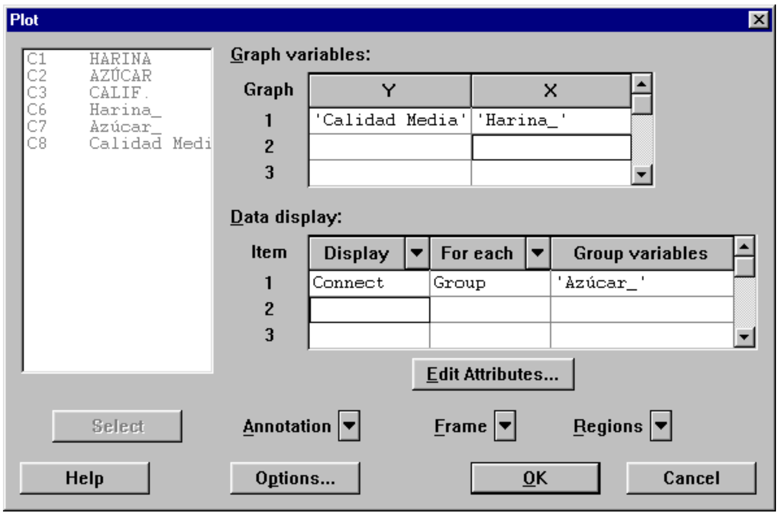
- La calidad de los pasteles parece incrementarse conforme lo hace el nivel de harina empleado: los seis pasteles elaborados para el nivel 0% tienen una calidad media de 3,8; los seis elaborados para el nivel 10% muestran una calidad media de 4,2; los elaborados para el nivel 20% tienen una calidad media de 4,9; y los elaborados para el nivel 30% muestran una calidad media de 5,1.

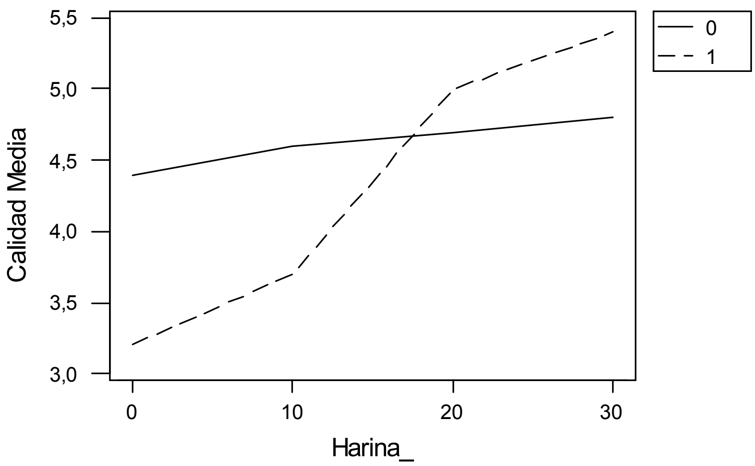
- Por lo que se refiere al azúcar, no parece haber mucha diferencia entre los valores promedio obtenidos: la calidad media para los 12 pasteles elaborados sin azúcar (4,6) es sólo ligeramente superior a la cantidad media de los otros 12 (4,3).

Financiado por la Secretaría de Estado de Educación y Universidades (MECD)

En el siguiente gráfico se muestra la evolución de la calidad media en función de la cantidad de harina empleada (descompuesto por niveles del factor azúcar.

Graph → Plot...

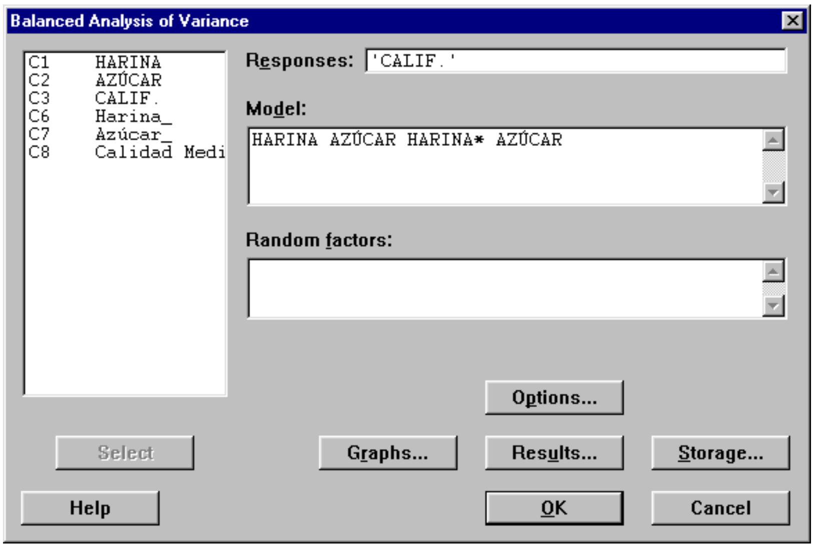




Según se aprecia, el uso del azúcar incrementa la calidad media de los pasteles cuando estamos en los niveles superiores de harina (20% y 30%), mientras que ocurre todo lo contrario para niveles bajos de harina (0% y 10%). Por tanto, que el azúcar mejore o no la calidad del pastel dependerá del nivel de harina que éste contenga. Esto significa que ambos factores interactúan. En general, si los factores no interactuasen las líneas del gráfico anterior serían (aproximadamente) paralelas.

Aplicaremos ahora la técnica ANOVA usando un modelo con interacción.

Stat → ANOVA → Balanced Anova…



Analysis of Variance (Balanced Designs)

Factor	Type	Levels	Values			
HARINA	fixed	4	0	10	20	30
AZÚCAR	fixed	2	0	1		

Analysis of Variance for CALIF.

Source	DF	SS	MS	F	P
HARINA	3	6,6912	2,2304	74,35	0,000
AZÚCAR	1	0,5104	0,5104	17,01	0,001
HARINA*AZÚCAR	3	3,7246	1,2415	41,38	0,000
Error	16	0,4800	0,0300		
Total	23	11,4063			

Analicemos los resultados de los tres contrastes (H_{01}: el factor harina no influye sobre la calidad del pastel, H_{02}: el factor azúcar no influye sobre la calidad del pastel, H_{03}: no hay interacción entre ambos factores):

Lo primero es comprobar si existe interacción, puesto que en tal caso deberemos interpretar con cautela los resultados de los otros tests. En este caso p-valor = 0,000. Por tanto, hay fuertes evidencias de la existencia de interacción entre ambos factores.

Obtenemos también p-valores significativos en los otros dos contrastes; es decir, tanto el nivel de harina como el nivel de azúcar son determinantes para la calidad esperada de un pastel. Notar que la calidad media de un pastel sin azúcar (4,6333) es mayor que la calidad media de un pastel con azúcar (4,3417). Esto es consistente con el resultado de nuestro test. Sin embargo, según hemos observado anteriormente, el uso del azúcar tiende a aumentar la calidad media para niveles altos de harina, y viceversa. En general, siempre que haya interacción entre los factores convendrá hacer un análisis detallado que vaya más allá del p-valor obtenido en el contraste.

Educación universitaria en la UE

En la siguiente tabla se muestran los ratios (%) de educación universitaria por países de la Unión Europea (no se dispone de los datos de Luxemburgo), distribuidos por sexo y grupos de edad. (Fuente: Informe "Retrato de los europeos", Eurostat, Año 2002).

	18-21 años		22-24 años		25-28 años	
	H	M	H	M	H	M
BÉLGICA	36	47	21	20	6	5
DINAMARCA	7	10	23	33	17	21
ALEMANIA	7	15	20	20	15	11
GRECIA	57	63	9	8	3	3
ESPAÑA	28	40	26	30	10	9
FRANCIA	30	40	22	25	5	7
IRLANDA	30	38	12	12	4	5
ITALIA	12	17	19	27	12	14
HOLANDA	24	30	25	22	8	6
AUSTRIA	11	18	20	20	14	12
PORTUGAL	22	30	17	23	8	10
FINLANDIA	19	26	37	42	21	22
SUECIA	13	19	25	30	14	16
REINO UNIDO	26	31	11	12	6	7

En primer lugar, vamos a determinar si existen diferencias significativas de educación universitaria entre los dos sexos (sin tener en cuenta la edad), para lo cual utilizaremos un modelo ANOVA simple.
Para ello, introducimos los datos en Minitab, en distintas columnas (ver archivo `universitarios.mtw`)

Stat → Anova → One-way ...

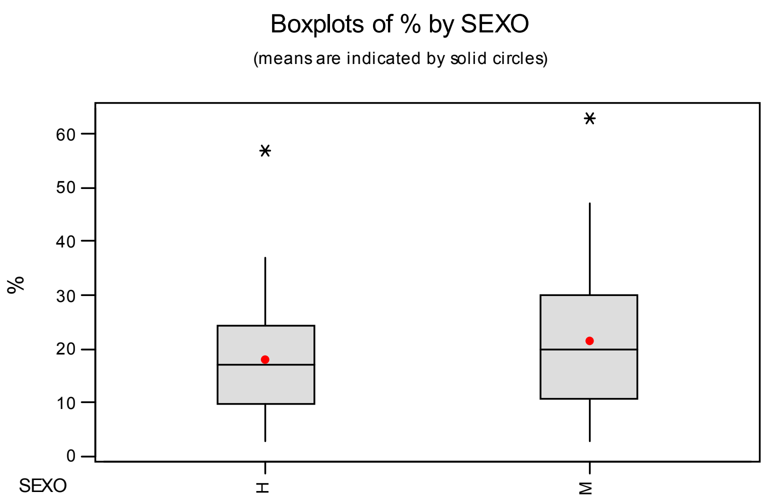
Tomamos como variable respuesta % y como factor SEXO, y solicitamos los gráficos de cajas (también valen los diagramas de puntos).
El output del Minitab es el siguiente:

One-way Analysis of Variance

```
Analysis of Variance for %
Source      DF        SS        MS        F        P
SEXO         1       247       247     1,73    0,192
Error       82     11711       143
Total       83     11958

    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level        N      Mean     StDev  ---+---------+---------+---------+---
H           42     17,90     10,68  (------------*-----------)
M           42     21,33     13,10              (-----------*-----------)
                                    ---+---------+---------+---------+---
Pooled StDev =     11,95            15,0      18,0      21,0      24,0
```

Mientras que el gráfico resultantes es:



En los gráficos se aprecia que las diferencias entre los dos sexos son mínimas, tanto en las medias como en la variabilidad (en ambos casos mayor en las mujeres), y por tanto no parecen significativas.
En la salida numérica se puede comprobar que el p-valor es 0,192, con lo cual no rechazaríamos la hipótesis nula para cualquier nivel de significación razonable (por ejemplo, α=0,05). De hecho, los intervalos de confianza tienen bastantes valores comunes. Por tanto, concluiríamos que no existen diferencias estadísticamente significativas entre los sexos en cuanto a los ratios de educación universitaria en los países europeos.

Supongamos que ahora queremos tener en cuenta también los grupos de edad, para determinar si hay diferencias en los ratios de educación universitaria entre los grupos de edad y/o el sexo. Para ello, utilizaremos el modelo doble con interacción.
En este caso tenemos un modelo con dos factores (sexo y edad) e interacción, siendo la variable respuesta, obviamente, %. En la variable edad, los grupos de edad se han numerado: 1 (18-21 años), 2 (22-24 años), 3 (25-28 años).
Antes de realizar el ANOVA vamos a presentar los datos de forma tabular para ver si obtenemos alguna conclusión previa.

Stat → Tables → Cross Tabulation...

Tabulated Statistics

```
Rows: EDAD     Columns: SEXO

           H         M       All

 1     23,000    30,286    26,643
 2     20,500    23,143    21,821
 3     10,214    10,571    10,393
 All   17,905    21,333    19,619

  Cell Contents --
                %:Mean
```

Se puede observar que, a medida que aumenta la edad, los ratios decrecen sensiblemente en los dos sexos, y además las diferencias entre los sexos son menores (siempre a favor de las mujeres).
Pasamos a realizar el ANOVA doble.

Stat → Anova → Two-way ...

Two-way Analysis of Variance

```
Analysis of Variance for %
Source          DF       SS        MS        F        P
EDAD             2    3900,6    1950,3    19,92    0,000
SEXO             1     246,9     246,9     2,52    0,116
Interaction      2     174,5      87,3     0,89    0,414
Error           78    7635,9      97,9
Total           83   11957,8


                       Individual 95% CI
EDAD            Mean   ---------+---------+---------+---------+--
1               26,6                            (-----*------)
2               21,8                    (-----*------)
3               10,4   (-----*------)
                       ---------+---------+---------+---------+--
                             12,0      18,0      24,0      30,0

                       Individual 95% CI
SEXO            Mean   -+---------+---------+---------+---------+
H               17,9   (------------*-----------)
M               21,3                 (-----------*-----------)
                       -+---------+---------+---------+---------+
                     15,0      17,5      20,0      22,5      25,0
```

Basándonos en la salida anterior, podemos afirmar que no existe interacción entre los dos factores (p-valor=0,414).
Así mismo, obtenemos un p-valor prácticamente nulo (0,000) para el factor edad, lo cual nos indica que existen diferencias significativas entre los tres grupos, tal y como habíamos detectado en la tabla cruzada. Esto también se puede comprobar observando los intervalos de confianza, ya que si bien entre los grupos 1 y 2 comparten muchos valores, ninguno de ellos tiene intersección con el intervalo del grupo 3.
Por otra parte, al igual que en el modelo de un factor, no existen evidencias de que el sexo influya en los ratios de educación, ya que el p-valor es 0,116>0,05.

Rapidez en la impresión

Estamos interesados en comparar la rapidez de dos modelos de impresoras (0 y 1). Para realizar el experimento, medimos el tiempo de impresión (en segundos) de los dos modelos sobre una serie de plantillas estándar.
Los resultados del experimento se muestran en el archivo `impresoras.mtw`. Los datos han sido obtenidos después de medir el tiempo de impresión (recogido en la variable puntuación) de las dos máquinas sobre una muestra de 80 plantillas de iguales características, es decir, plantillas con similar dificultad de impresión.
Queremos determinar si existen diferencias significativas entre los dos tipos de impresoras.

Stat → Anova → One-way ...

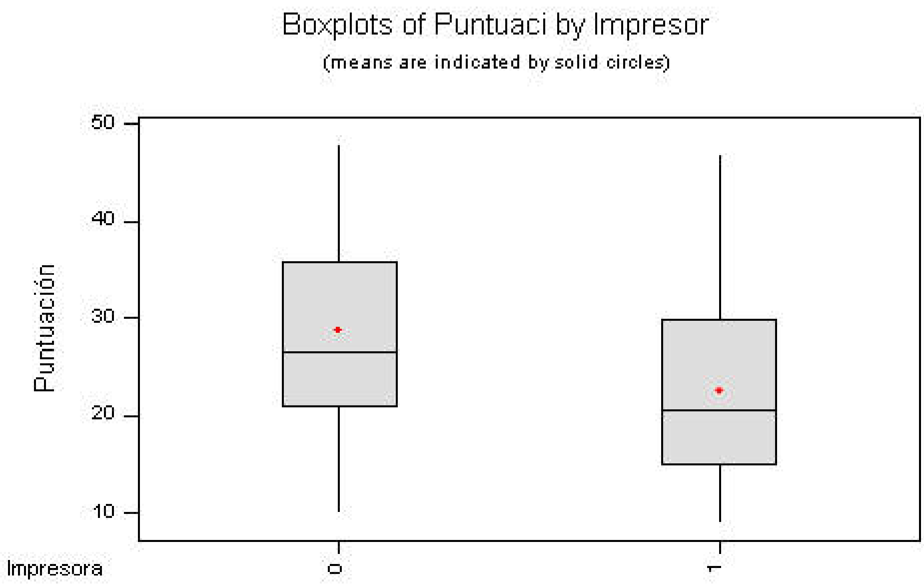
Lógicamente, tomamos como variable respuesta puntuación, y como factor impresora.
Solicitamos los diagramas de puntos y los gráficos de cajas.
El output del Minitab es el siguiente:

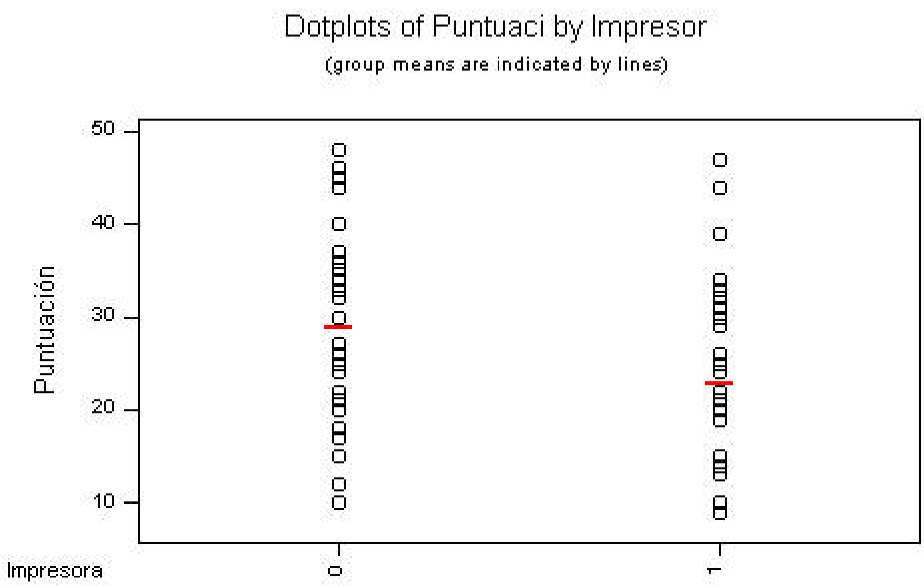
One-way Analysis of Variance

```
Analysis of Variance for Puntuaci
Source     DF        SS        MS        F        P
Impresor    1     732,0     732,0     7,87    0,006
Error      78    7259,9      93,1
Total      79    7992,0
                                  Individual 95% CIs For Mean
                                  Based on Pooled StDev
```

```
Level        N      Mean     StDev  ----+---------+---------+---------+--
0           40    28,750    10,081                 (--------*--------)
1           40    22,700     9,194  (--------*--------)
                                    ----+---------+---------+---------+--
Pooled StDev =     9,648             21,0      24,5      28,0      31,5
```

Mientras que los gráficos resultantes son:





En los gráficos se aprecia que existen diferencias (el modelo 0 ofrece puntuaciones más altas) entre los modelos de impresora, aunque resulta difícil saber si son o no significativas.
En la salida numérica se puede comprobar que el p-valor es 0,006, con lo cual rechazaríamos la hipótesis nula al nivel de significación habitual (α=0,05). De hecho, los intervalos de confianza son prácticamente disjuntos. Por tanto, concluiríamos que existen diferencias significativas entre las impresoras

Supongamos que ahora queremos tener en cuenta también en el experimento la calidad del papel utilizado, ya que consideramos que puede influir en la rapidez de impresión.
Buscamos, pues, determinar si hay diferencias entre los tipos de impresoras y/o entre las calidades en el papel utilizado, y si existe relación entre el tipo de impresora y el tipo de papel utilizado.
Consideraremos un modelo doble (factores: impresora y papel) con interacción, siendo la variable respuesta, obviamente, puntuación.
Como en otras ocasiones, antes de realizar el ANOVA vamos a presentar los datos de forma tabular para ver si obtenemos alguna conclusión previa:

Stat → Tables → Cross Tabulation...

Tabulated Statistics

```
Rows: Impresor     Columns: Papel

          0        1      All

0    30,450   27,050   28,750
1    19,350   26,050   22,700
All  24,900   26,550   25,725

 Cell Contents --
        Puntuaci:Mean
```

Se observa que las impresoras del modelo 0 siempre ofrecen valores mayores que las del modelo 1, pero la diferencia es mucho más acusada con papel de tipo 0 que con papel de tipo 1. Esto nos hace intuir que existe interacción entre ambos factores.

Pasamos a realizar el ANOVA doble:

Stat → ANOVA → Two-way…

Two-way Analysis of Variance

```
Analysis of Variance for Puntuaci
Source        DF        SS        MS        F        P
Impresor       1     732,0     732,0     8,31    0,005
Papel          1      54,4      54,4     0,62    0,434
Interaction    1     510,1     510,1     5,79    0,019
Error         76    6695,4      88,1
Total         79    7991,9

                       Individual 95% CI
Impresor        Mean   -----+---------+---------+---------+------
0               28,8                        (---------*---------)
1               22,7    (---------*---------)
                       -----+---------+---------+---------+------
                         21,0      24,0      27,0      30,0

                       Individual 95% CI
Papel           Mean   -+---------+---------+---------+---------+
0               24,9    (--------------*-------------)
1               26,6            (--------------*--------------)
                       -+---------+---------+---------+---------+
                     22,0      24,0      26,0      28,0      30,0
```

Basándonos en la salida anterior, podemos afirmar que existe interacción entre los dos factores (p-valor=0,019).
Así mismo, obtenemos un p-valor muy pequeño (0,005) para el factor impresora, lo cual nos indica que existen diferencias significativas entre ambos modelos de impresoras (intervalos de confianza disjuntos).
Por otra parte, no existen evidencias de que el tipo de papel influya en la rapidez de la impresión, ya que el p-valor es 0,434>0,05. Además, los intervalos de confianza tienen una intersección de longitud muy amplia.
No obstante, el hecho de que exista interacción entre ambos factores nos impide realizar las afirmaciones anteriores con total rotundidad, ya que sería necesario un análisis más detallado de los datos.

BIBLIOGRAFÍA

[1] Baró, J. y Alemany, R. (2000): “Estadística II”. Ed. Fundació per a la Universitat Oberta de Catalunya. Barcelona.

[2] Peña Sánchez de Rivera, D. (1987): “Estadística. Modelos y Métodos. Volumen 2”. Alianza Editorial. Madrid. ISBN: 84-206-8110-5

[3] Johnson, R. R. (1996): “Elementary statistics”. Belmont, etc. : Duxbury, cop

[4] Martín-Guzmán, P. (1991): “Curso básico de estadística económica”. AC, DL. Madrid. ISBN: 84-7288-142-3

[5] Wonnacott, Thomas H. (1997): "Introducción a la estadística". Limusa, México.

[6] Moore, David S. (1998): "Estadística aplicada básica". Antoni Bosch, Barcelona.

ENLACES

- http://www.uv.es/~lejarza/anova/anova.html
  Lección de ANOVA en HTML (Universitat de València)
- http://e-stadistica.bio.ucm.es/cont_mod_1.html - Anova
  Aula Virtual de Bioestadística (Universidad Complutense de Madrid)
- http://www.ruf.rice.edu/~lane/stat_sim/
  Applets de Java
- http://www.kuleuven.ac.be/ucs/java/version2.0/Content_Anova.htm
  Applets de Java