10. DISEÑOS EXPERIMENTALES

Dr. Edgar Acuña

http://math.uprm.edu/~edgar

UNIVERSIDAD DE PUERTO RICO

RECINTO UNIVERSITARIO DE MAYAGUEZ

Diseños Experimentales de Clasificación Simple

- En un diseño experimental de clasificación simple, se trata de comparar varios grupos generalmente llamados **Métodos** o **Tratamientos,**
- Para hacer la comparación se usa una variable de respuesta cuantitativa Y que es medida en cada uno de los grupos. Los grupos también pueden ser los niveles de una variable cualitativa que es llamada **Factor,** como por ejemplo niveles de conocimiento: básico, intermedio, avanzado.

Datos

Grupo 1	Grupo 2	Grupo 3	…	Grupo k
Y_{11}	Y_{21}	Y_{31}	…	Y_{k1}
Y_{12}	Y_{22}	Y_{32}	…	Y_{k2}
Y_{13}	Y_{23}	Y_{33}	…	Y_{k3}
:	:	:		:
Y_{1n1}	Y_{2n2}	Y_{3n3}	…	Ykn_k

El Grupo 1 tiene n_1 observaciones, el Grupo 2 tiene n_2 observaciones, y así sucesivamente.

Modelo lineal de un diseño experimental

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

Donde:

y_{ij} : Es la *j-ésima* observación del grupo .

μ : Es la media total.

α_i : Es el efecto del grupo .

ε_{ij} : Error aleatorio de la *j-ésima* observación del grupo .

Prueba de hipotesis para comparar los grupos

La comparacion de los grupos se reduce a determinar si hay igualdad de medias poblacionales de la variable de respuesta en todos los grupos.

H_o: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$ versus

H_a: Al menos un grupo tiene distinta media poblacional.

La prueba estadística que se usa para tomar una decisión es la prueba de F, la prueba F es obtenida al completar la tabla del análisis de varianza.

MINITAB da el p-value de la prueba de F.

Análisis de Varianza

Fuentes de Variación	**Grados de Libertad**	**Suma de Cuadrados**	**Cuadrados Medios**	**F**
Entre Grupos	k-1	BSS	BMS = BSS/k-1	BMS/MSE
Dentro de Grupos	n-k	SSE	MSE = SSE/n-k	
Total	n-1	SST		

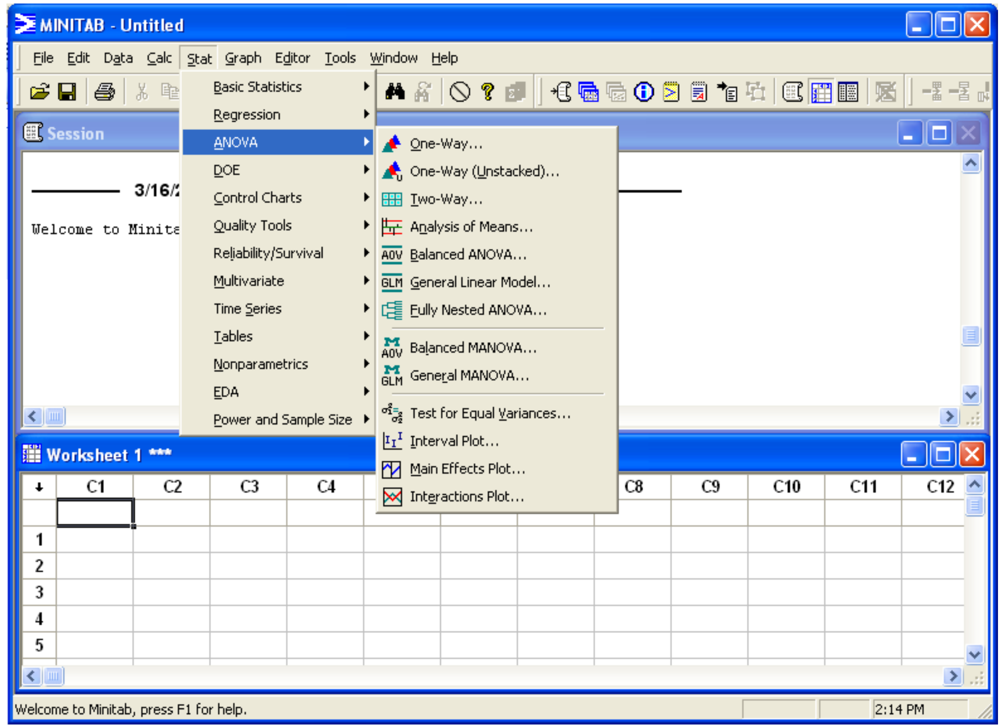
$$n = \sum_{i=1}^{k} n_i \qquad SST = \sum_{i=1}^{k}\sum_{i=1}^{n_k} y_{ij}^2 - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n_k} y_{ij})^2}{n} \qquad BSS = \frac{T_1^2}{n_1} + \frac{T_2^2}{n_2} + ... + \frac{T_k^2}{n_k} - \frac{(\sum_{i=1}^{k}\sum_{j=1}^{n_k} y_{ij})^2}{n}$$

T_i representa el total del *i-ésimo* Grupo

Si la F calculada es mayor que una F con k-1 y n-k grados de libertad, al nivel de significación α entonces, se rechaza la hipótesis nula.

ANOVA en el menú **Stat**

La opción ***One-Way*** del menú **ANOVA** se usa para hacer análisis de varianza de clasificación simple cuando los datos de la variable de respuesta van en una sola columna y los niveles del factor (o Grupos) van en otra columna.



La opción ***One-Way (Unstacked)***, se usa también para hacer diseños de clasificación simple, pero cuando los datos de los grupos a comparar son entrados columna por columna.

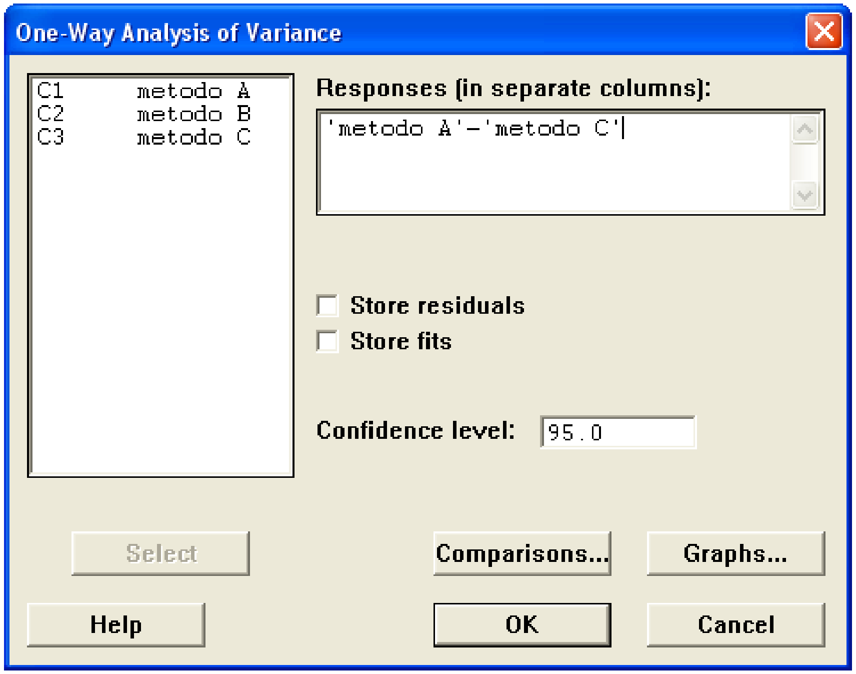
Ejemplo 10.1

Se desea comparar 3 métodos de enseñanza A, B y C, se eligen al azar una muestra de estudiantes de cada método y se le aplica una prueba final común. Los resultados son como sigue:

método A	método B	método C
89	78	64
45	85	69
59	93	82
46	81	74
64	79	79
71	98	
	94	

¿Habrá suficiente evidencia para concluir que hay diferencia entre métodos?

Usando la opción ***One-way [Unstacked]***

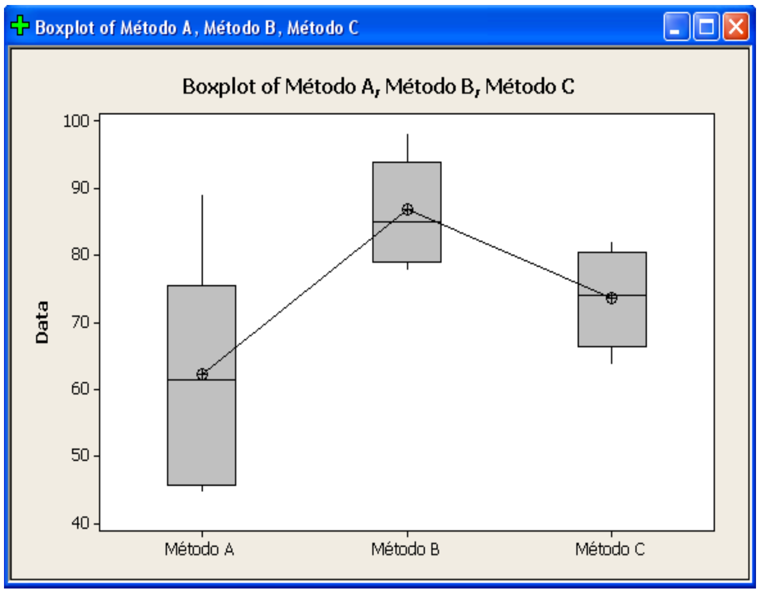


Resultados

```
One-way ANOVA: Método A, Método B, Método C
Source  DF    SS   MS     F      P
Factor   2  1957  978  7.44  0.006
Error   15  1971  131
Total   17  3928
S = 11.46   R-Sq = 49.81%   R-Sq(adj) = 43.12%
                         Individual 95% CIs For Mean Based on
                         Pooled StDev
Level     N   Mean  StDev  ------+---------+---------+---------+---
Método A  6  62.33  16.54  (-------*-------)
Método B  7  86.86   8.07                      (------*-------)
Método C  5  73.60   7.30            (--------*--------)
                           ------+---------+---------+---------+---
                                60        72        84        96
Pooled StDev = 11.46
```

Interpretación: “P-value” = .006 se rechaza la afirmación, “todos los métodos sean iguales”.

Boxplots para comparar los tres métodos del ejemplo



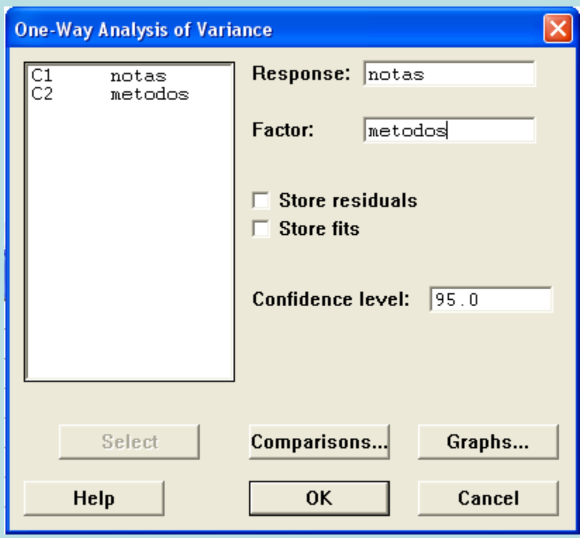
Interpretación: *La posición de la mediana y las medias sugiere que aún cuando los métodos B y C no están muy distantes, si existe una diferencia marcada entre los métodos B y A, lo cual llevará a rechazar la hipótesis de igualdad de medias.*

Hay que notar que la variabilidad del método A es mucho mayor que los otros dos métodos.

Resultados al usar la opción ***One-Way***

Para usar la opción ***One-Way*** los datos deben ser entrados en dos columnas: Una de ellas conteniendo los valores de la variable de respuesta y la otra los valores que indican a que grupo pertenecen dichos datos.

notas	método
89	1
45	1
59	1
46	1
64	1
71	1
78	2
85	2
93	2
81	2
79	2
98	2
94	2
64	3
69	3
82	3
74	3
79	3



One-way ANOVA: notas versus método

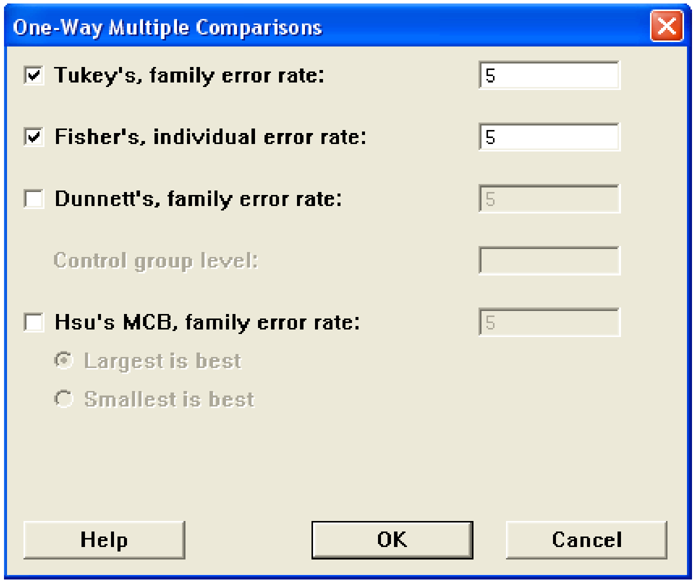
```
Source  DF    SS   MS     F      P
método   2  1957  978  7.44  0.006
Error   15  1971  131
Total   17  3928
S = 11.46   R-Sq = 49.81%   R-Sq(adj) = 43.12%
                    Individual 95% CIs For Mean Based on
                    Pooled StDev
Level  N   Mean  StDev  ------+---------+---------+---------+---
1      6  62.33  16.54  (-------*-------)
2      7  86.86   8.07                     (------*-------)
3      5  73.60   7.30           (--------*--------)
                        ------+---------+---------+---------+---
                             60        72        84        96
Pooled StDev = 11.46
```

Comparaciones Múltiples

Una vez que se ha rechazado que todos los grupos son iguales hay que determinar cuales de ellos son comparables entre si. Tukey y Fisher son metodos más utilizados, aplican el siguiente criterio

Los Grupos *i* y *j* son comparables entre ellos, si se cumple:
| *media del Grupo i -Media del Grupo j* | < *valor crítico*

En **MINITAB** use la opcion **Comparisons** de **Oneway.**



Métodos para Comparaciones Múltiples

Método de Tukey

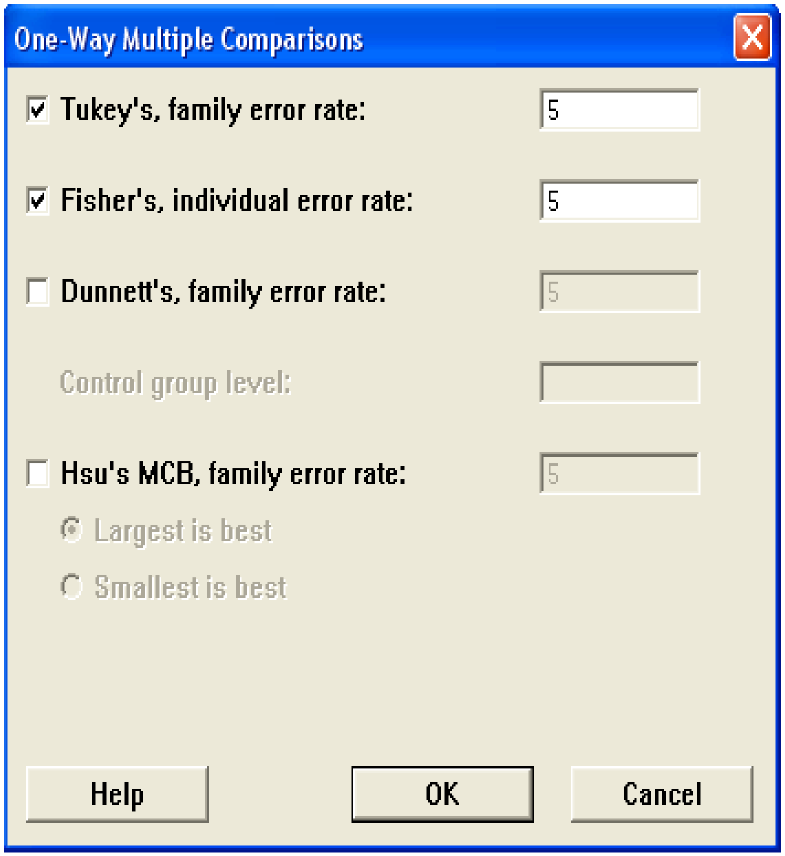
El valor crítico está dado por $\dfrac{Q}{\sqrt{2}}\, s\sqrt{\dfrac{1}{n_i}+\dfrac{1}{n_j}}$

donde: n_i es el tamaño del *i-ésimo* grupo,
n_j es el tamaño del *j-ésimo* grupo,
s es igual a la desviación estándar combinada de los grupos tambien es igual a la raíz cuadrada del cuadrado medio del error (MSE) y Q es el percentil de 100α% de la ***distribución del rango estudentizado*** con parámetros k-1 y n-k.

Método de Fisher

el valor crítico está dado por $t_{(\alpha/2,n-k)}\, s\sqrt{\dfrac{1}{n_i}+\dfrac{1}{n_j}}$

donde, $t(\alpha/2,n\text{-}k)$ representa el valor de la distribución t tal que, el área a la derecha es $\alpha/2$.



Resultados para el ejemplo anterior

```
Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons among Levels of método
Individual confidence level = 97.97%
método = 1 subtracted from:
método  Lower  Center  Upper  -----+---------+---------+---------+----
 2       7.97   24.52  41.08                       (-------*--------)
 3      -6.75   11.27  29.28               (--------*--------)
                              -----+---------+---------+---------+----
                                 -20         0        20        40
método = 2 subtracted from:
método   Lower  Center  Upper  -----+---------+---------+---------+----
   3    -30.68  -13.26   4.16  (-------*--------)
                               -----+---------+---------+---------+----
                                  -20         0        20        40
Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons among Levels of método
Simultaneous confidence level = 88.31%
método = 1 subtracted from:
método  Lower  Center  Upper   ----+---------+---------+---------+-----
    2   10.93   24.52  38.12                       (------*------)
    3   -3.53   11.27  26.06              (-------*------)
                               ----+---------+---------+---------+-----
                                 -20         0        20        40
método = 2 subtracted from:
método   Lower  Center  Upper  ----+---------+---------+---------+-----
   3    -27.56  -13.26   1.05  (------*-------)
                               ----+---------+---------+---------+-----
                                 -20         0        20        40
```

Interpretación:
Por cada combinación de grupos aparecen los límites inferiores y superiores de los intervalos de confianza para la diferencia poblacional de las dos medias.

Los métodos Tukey y Fisher nos llevan a la conclusión de que los métodos de enseñanza A y C son comparables al igual que B y C porque los intervalos de confianza contienen a cero. Pero A y B no lo son porque el intervalo de confianza no contiene a cero. Hay un nivel superior formado por los métodos B y C y un nivel inferior formado por C y A. Notar que C aparece en ambos niveles.

Ejemplo de ANOVA

Los siguientes datos representan los tiempos de sobrevivencia a varios tipos de cáncer, después que se lo ha diagnosticado.

Estómago	Pulmón	Colon	Ovario	Seno
248	124	1234	81	1235
377	42	89	461	24
189	25	201	20	1581
1843	45	356	450	1166
180	412	2970	246	40
537	51	456	166	727
519	1112	63	3808	
455	46	64	791	
406	103	155	1804	
365	876	859	3460	
942	146	151	719	
776	340	166		
372	396	37		
163		223		
101		138		
20		72		
283		245		

Hacer un análisis de varianza para probar si hay igual tiempo de sobrevivencia para los diversos tipos de cáncer. Aplicar los métodos de comparaciones múltiples de Fisher y Tukey para identificar los tipos de cáncer con tiempos de sobrevivencia similares.

Solución:

La hipótesis nula esta dada por
Ho: Los tiempos promedios de sobrevivencia de los pacientes diagnosticados con cáncer de estómago, pulmón, colon, ovario y seno son iguales.
La hipótesis alterna esta dado por *Ha*: Al menos uno de los tipos de cáncer tiene tiempo de sobrevivencia promedio distinto a los otros.

Primero se entran los datos en dos columnas: *Sobrevivencia*, que contiene los tiempos de sobrevivencia y *Organo*, que contiene los órganos donde el cáncer es detectado, se sigue la secuencia **Stat** ▸ **ANOVA** ▸ **One-Way** y oprimiendo el botón **comparisons** se obtienen los siguientes resultados en la ventana **session:**

Resultados

One-way ANOVA: tiempo versus cancer

```
Source  DF        SS       MS     F      P
cancer   4  11535761  2883940  6.43  0.000
Error   59  26448144   448274
Total   63  37983905
S = 669.5   R-Sq = 30.37%   R-Sq(adj) = 25.65%
                          Individual 95% CIs For Mean Based on
                          Pooled StDev
 Level      N    Mean   StDev   --+---------+---------+---------+-------
 colon      6   884.3  1098.6           (----------*----------)
estomago   17   457.4   427.2        (-----*------)
ovario     17   211.6   209.9   (-----*------)
pulmon     13   286.0   346.3   (-------*------)
seno       11  1395.9  1239.0                        (-------*-------)
                                --+---------+---------+---------+-------
                                  0       500      1000      1500
Pooled StDev = 669.5
Tukey 95% Simultaneous Confidence Intervals
All Pairwise Comparisons among Levels of cancer
Individual confidence level = 99.34%
cancer = colon subtracted from:
 cancer      Lower  Center   Upper  ---------+---------+---------+---------+
estomago  -1321.7  -426.9   467.8         (--------*--------)
ovario    -1567.5  -672.7   222.0       (--------*--------)
pulmon    -1528.3  -598.3   331.6        (--------*--------)
seno       -444.7   511.6  1467.9                 (--------*---------)
                                    ---------+---------+---------+---------+
                                          -1000        0      1000      2000
cancer = estomago subtracted from:
 cancer   Lower  Center   Upper    ---------+---------+---------+---------+
  ovario  -892.1  -245.8   400.5            (------*-----)
 pulmon   -865.6  -171.4   522.8            (------*------)
 seno      209.4   938.5  1667.6                      (------*-------)
                                   ---------+---------+---------+---------+
                                         -1000        0      1000      2000
```

```
cancer = ovario subtracted from:
cancer   Lower  Center   Upper   ---------+---------+---------+---------+
pulmon  -619.8    74.4   768.6             (------*------)
seno     455.2  1184.3  1913.4                     (------*------)
                                 ---------+---------+---------+---------+
                                       -1000        0      1000      2000
cancer = pulmon subtracted from:
cancer  Lower  Center   Upper    ---------+---------+---------+---------+
seno    338.0  1109.9  1881.8                      (-------*-------)
                                 ---------+---------+---------+---------+
                                       -1000        0      1000      2000
Fisher 95% Individual Confidence Intervals
All Pairwise Comparisons among Levels of cancer
Simultaneous confidence level = 72.17%
cancer = colon subtracted from:
cancer       Lower  Center   Upper  -------+---------+---------+---------+--
estomago  -1063.1  -426.9   209.3       (------*-----)
ovario    -1308.9  -672.7   -36.6     (-----*------)
pulmon    -1259.6  -598.3    62.9     (------*------)
seno       -168.4   511.6  1191.5             (------*------)
                                    -------+---------+---------+---------+--
                                        -1000        0      1000      2000
cancer = estomago subtracted from:
cancer   Lower  Center   Upper   -------+---------+---------+---------+--
ovario  -705.3  -245.8   213.7           (----*---)
pulmon  -665.0  -171.4   322.2           (----*----)
seno     420.1   938.5  1456.9                   (----*-----)
                                 -------+---------+---------+---------+--
                                     -1000        0      1000      2000
cancer = ovario subtracted from:
cancer   Lower  Center   Upper   -------+---------+---------+---------+--
pulmon  -419.2    74.4   568.0             (----*----)
seno     665.9  1184.3  1702.7                     (----*----)
                                 -------+---------+---------+---------+--
                                     -1000        0      1000      2000
cancer = pulmon subtracted from:
cancer  Lower  Center   Upper    -------+---------+---------+---------+--
seno    561.1  1109.9  1658.8                      (----*-----)
                                 -------+---------+---------+---------+--
                                     -1000        0      1000      2000
```

Interpretación

El "p-value" de la prueba de F es .0000, lo cual sugiere que la hipótesis nula se rechaza y se concluye que hay suficiente evidencia estadística para afirmar que al menos uno de los tipos de cáncer tiene tiempo de sobrevivencia promedio distinto a los otros.

De acuerdo al método de **Tukey**
Los cáncer al pulmón, estómago, colon y ovarios tienen tiempos de sobrevivencia similares, formado una categoría *inferior*.
Los cánceres de ovarios y senos tienen tiempos promedios de sobrevivencias similares, formando una categoría *superior*.

De acuerdo al método de **Fisher**
Los cánceres al pulmón, estómago y colon tienen tiempos de sobrevivencia similares y forman una categoría *inferior*.
Los cánceres al estómago, colon y ovarios tienen tiempos de sobrevivencia similares y forman una categoría *intermedia*.
Los cánceres de ovarios y senos tienen tiempos promedios de sobrevivencias similares y forman la categoría *superior*.

Diseños Experimentales de clasificación Doble

En este caso se trata de comparar grupos (métodos o tratamientos) pero, tomando en cuenta un segundo factor el cual podría afectar la comparación de los mismos.

	Grupo 1	Grupo 2	…	Grupo k
Bloque 1	Y_{111} Y_{112}	Y_{211} Y_{212}	…	Y_{k11} Y_{k12}
Bloque 2	Y_{121} Y_{122}	Y_{221} Y_{222}	…	Y_{k21} Y_{k22}
:	:	:		:
Bloque B	Y_{1B1} Y_{1B2}	Y_{2B1} Y_{2B2}	…	Y_{kB1} Y_{kB2}

Pruebas de Hipótesis

Ho: $\mu1 = \mu2 = \mu3 = \ldots = \mu k$ (Los *k* grupos tienen medias poblacionales iguales), versus

Ha: Al menos un grupo tiene distinta media poblacional que los otros, y

Ho: $\mu1= \mu2=\mu3=\ldots=\mu B$ (Los B bloques tienen medias poblacionales iguales), versus

Ha: Al menos un bloque tiene media poblacional distinta al de los otros.

La **prueba estadística** correspondiente es la prueba de F, la cual es obtenida al completar la tabla del análisis de varianza.

Análisis de Varianza

La siguiente tabla del análisis de varianza es para un diseño con *k* grupos, *b* bloques y *c* observaciones en cada celda.

Fuentes de variación	**Grados de libertad**	**Suma de Cuadrados**	**Cuadrados Medios**	**F**
Grupos	k-1	SSG	MSG=SSG/k-1	MSG/MSE MSB/MSE
Bloques	b-1	SSB	MSB=SSB/b-1	
Error	kbc-k-b+1	SSE	MSE=SSE/kbc-k-b+1	
Total	kbc-1	SST		

MSG es el cuadrado medio de Grupos, **MSB** es el cuadrado medio de Bloques y **MSE** es el cuadrado medio del Error.

Si la *F* calculada es mayor que una *F* con *k*-1 y *kbc*-*k*-*b*+1 al nivel de significación α entonces, se rechaza la hipótesis nula de igualdad de medias de grupos y

Si la *F* calculada es mayor que una *F* con *b*-1 y *kbc*-*k*-*b*+1 al nivel de significación α entonces se rechaza la hipótesis nula de igualdad de medias de bloques.

Ejemplo 10.3

Se trata de comparar 3 métodos de enseñanza (*a*, *b* y *c*) pero tomando en cuenta además el factor turno (*m*, *t* y *n*), es decir el tiempo del día al cual se da clase. Los datos son como siguen:

	a	b	c
m	80.000	65.000	66.000
	78.000	79.000	49.000
t	69.000	50.000	34.000
	72.000	58.000	58.000
n	73.000	62.000	46.000
	74.000	65.000	59.000

Primero se entran los datos en tres columnas:

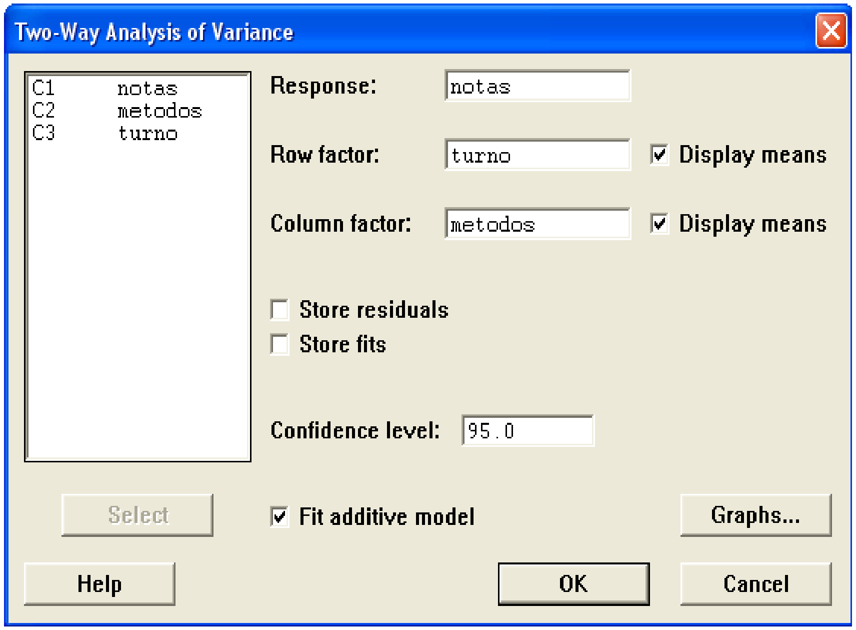
nota	**método**	**turno**
80	a	m
78	a	m
69	a	t
72	a	t
73	a	n
74	a	n
65	b	m
79	b	m
50	b	t
58	b	t
62	b	n
65	b	n
66	c	m
49	c	m
34	c	t
58	c	t
46	c	n
59	c	n

Ejemplo 10.3 cont.

Las hipótesis que se deben probar son:
Ho: No hay diferencia entre los tres métodos de enseñanza.
Ha: Al menos uno de los métodos de enseñanza tiene un rendimiento distinto a los otros y
Ho: Hay igual rendimiento de los estudiantes en los tres turnos.
Ha: En al menos uno de los turnos los estudiantes tienen distinto rendimiento con respecto a los otros dos turnos.

Eligiendo la secuencia **Stat** ▸ **Anova** ▸ **Two-Way** se obtiene la ventana de diálogo de la figura.

Notar que la opción ***Fit Additive model*** *debe ser seleccionada, de lo contario se ajustará un modelo con Interacción.*



Resultados para el ejemplo

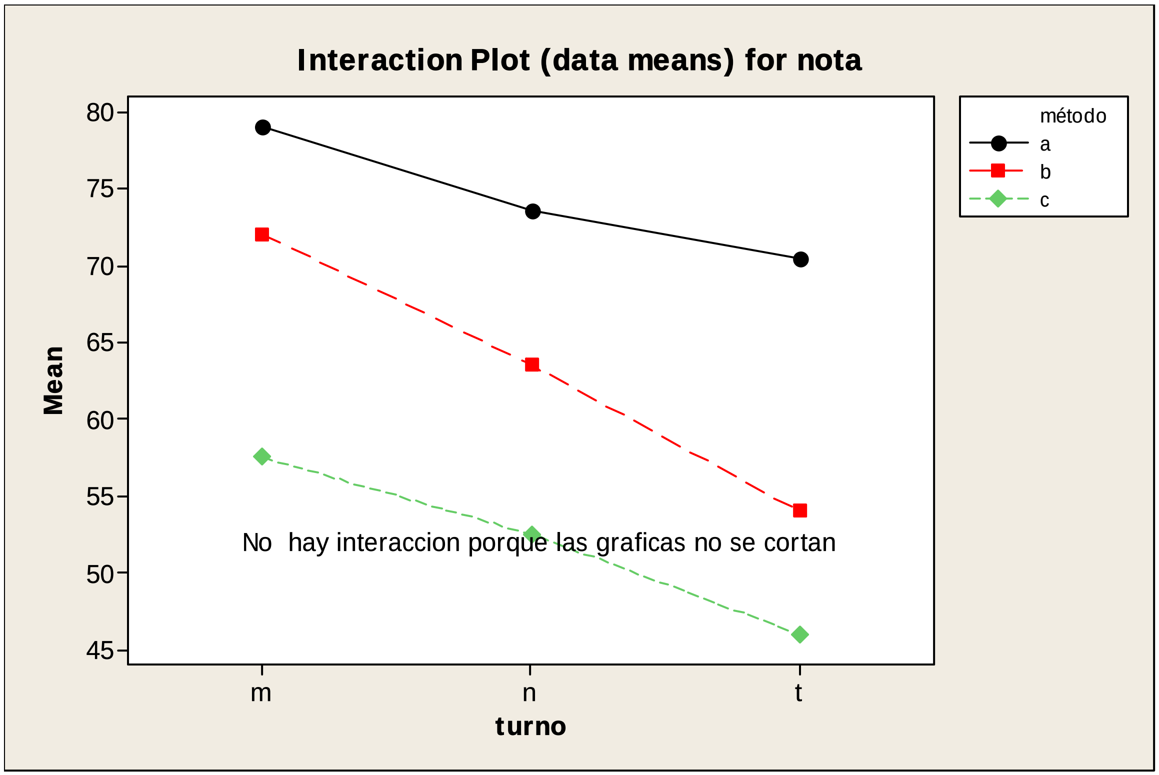
Two-way ANOVA: nota versus turno, método

Source	DF	SS	MS	F	P
turno	2	481.33	240.667	4.41	0.034
método	2	1496.33	748.167	13.72	0.001
Error	13	708.83	54.526		
Total	17	2686.50			

S = 7.384   R-Sq = 73.61%   R-Sq(adj) = 65.50%

```
                Individual 95% CIs For Mean Based on
                Pooled StDev
turno     Mean      --------+---------+---------+---------+-
m       69.5000                      (--------*---------)
n       63.1667             (--------*---------)
t       56.8333     (--------*--------)
                    --------+---------+---------+---------+-
                          56.0      63.0      70.0      77.0
```

```
                Individual 95% CIs For Mean Based on
                Pooled StDev
método    Mean      -----+---------+---------+---------+----
   a    74.3333                           (-----*------)
   b    63.1667                (-----*------)
   c    52.0000     (------*------)
                    -----+---------+---------+---------+----
                        50        60        70        80
```


Interaction Plot (data means) for nota
Mean
80
75
70
65
60
55
50
45
m
n
t
turno
método
a
b
c
No hay interaccion porque las graficas no se cortan

Modelo de clasificacion doble con interaccion

En este caso además de las hipótesis acerca de igualdad de medias de grupos y de igualdad de medias de bloques hay una tercera hipótesis referente a Interacción:

Ho: No hay interacción entre grupos y bloques

Ha: Sí hay interacción.

En **MINITAB** la tabla de Análisis de varianza es obtenida usando **two-way** con la opción **Fit Additve Model** sin ser elegida.

Two-way ANOVA: nota versus turno, método

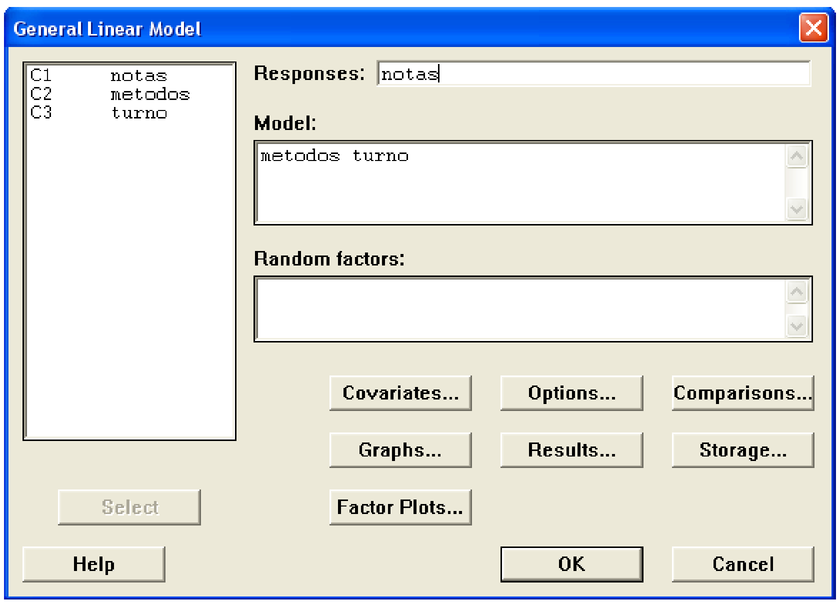
Source	DF	SS	MS	F	P
turno	2	481.33	240.667	3.29	0.085
método	2	1496.33	748.167	10.23	0.005
Interaction	4	50.33	12.583	0.17	0.947
Error	9	658.50	73.167		
Total	17	2686.50			

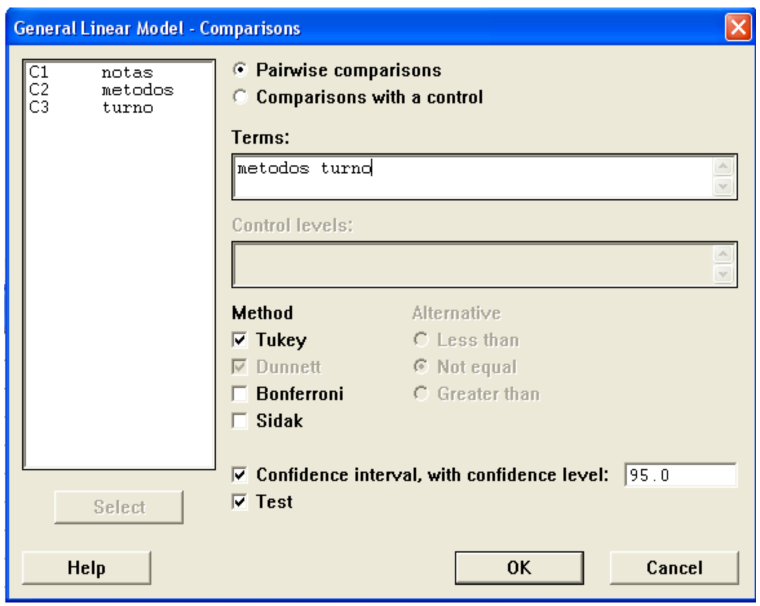
S = 8.554 R-Sq = 75.49% R-Sq(adj) = 53.70%

Interpretación: *El valor del "P-value" para Interacción es .947 que lleva a concluir que se debe aceptar la hipótesis nula de que no existe interacción entre los factores, lo cual ya se había concluido gráficamente.*

Usando General Linear Model

La opción ***General Linear Model*** del menú **ANOVA** permite analizar diseños de clasificación doble aún cuando no haya igual número de observaciones por celda y además tiene una opción que permite hacer comparaciones múltiples,





General Linear Model

General Linear Model: nota versus método, turno

Factor	Type	Levels	Values
método	fixed	3	a, b, c
turno	fixed	3	m, n, t

Analysis of Variance for nota, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
método	2	1496.33	1496.33	748.17	13.72	0.001
turno	2	481.33	481.33	240.67	4.41	0.034
Error	13	708.83	708.83	54.53		
Total	17	2686.50				

S = 7.38415   R-Sq = 73.61%   R-Sq(adj) = 65.50

Interpretación: *Viendo los "P-values" correspondientes a ambos factores se llega a la conclusión de que en al menos uno de los métodos de enseñanza el rendimiento es distinto y que en al menos uno de los turnos los estudiantes rinden distinto a los de los otros dos turnos.*

Resultados

```
Tukey 95.0% Simultaneous Confidence Intervals
Response Variable Nota
All Pairwise Comparisons among Levels of Método
Método = a  subtracted from:
Método   Lower  Center   Upper  ----+---------+---------+----
     -----+--
b      -22.41  -11.17    0.08            (----------*----------)
c      -33.58  -22.33  -11.09  (-----------*----------)
                          ----+---------+---------+---------+--
                            -30       -20       -10         0
Método = b  subtracted from:
Método   Lower  Center    Upper  ----+---------+---------+---
     ------+--
c      -22.41  -11.17  0.07766            (----------*----------)
                          ----+---------+---------+---------+--
                            -30       -20       -10         0
Tukey Simultaneous Tests
Response Variable Nota
All Pairwise Comparisons among Levels of Método
```

Método = a subtracted from:

Método	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
b	-11.17	4.263	-2.619	0.0520
c	-22.33	4.263	-5.239	0.0004

Método = b subtracted from:

Método	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
c	-11.17	4.263	-2.619	0.0520