

ANÁLISIS DE DATOS – 2º de BIOLOGÍA

PRÁCTICA 1: Introducción al SPSS

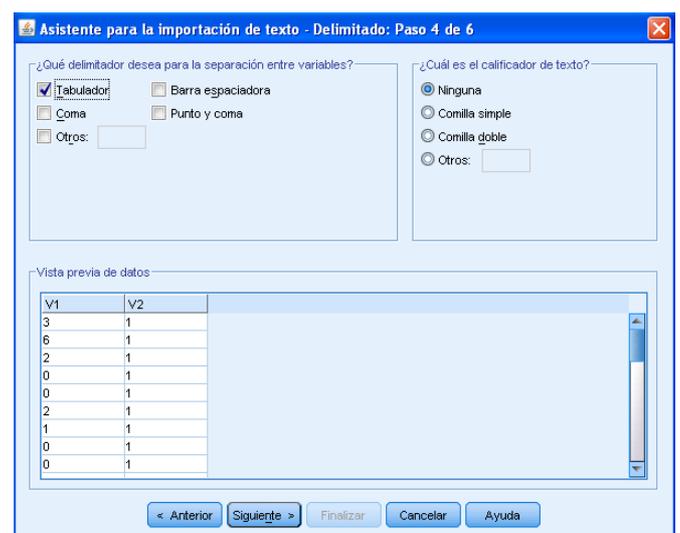
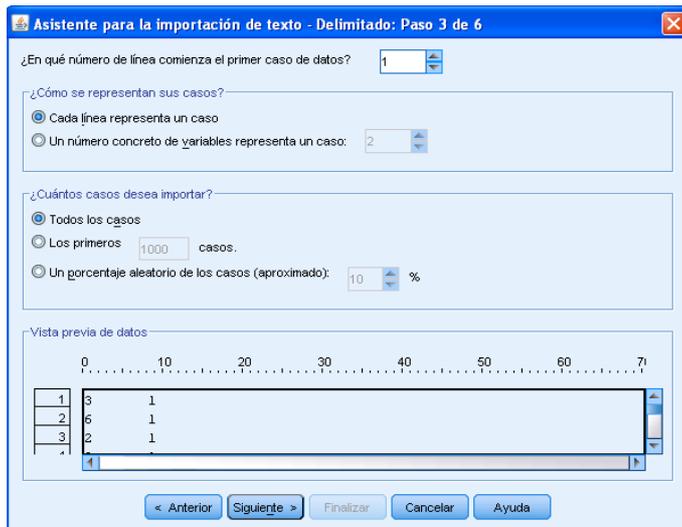
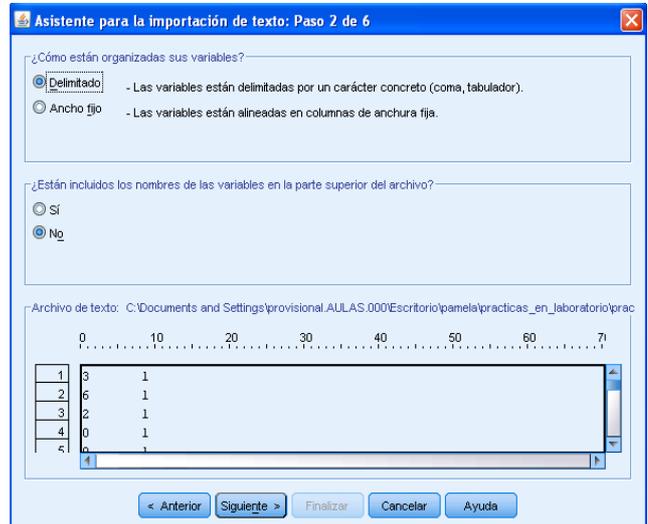
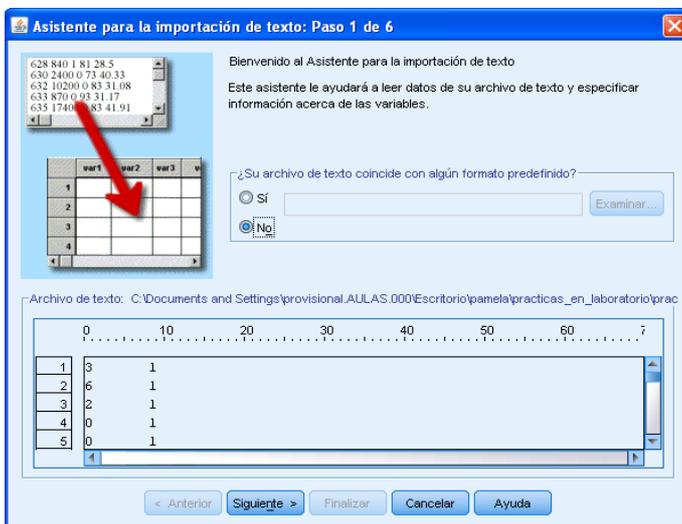
Experimento y datos:

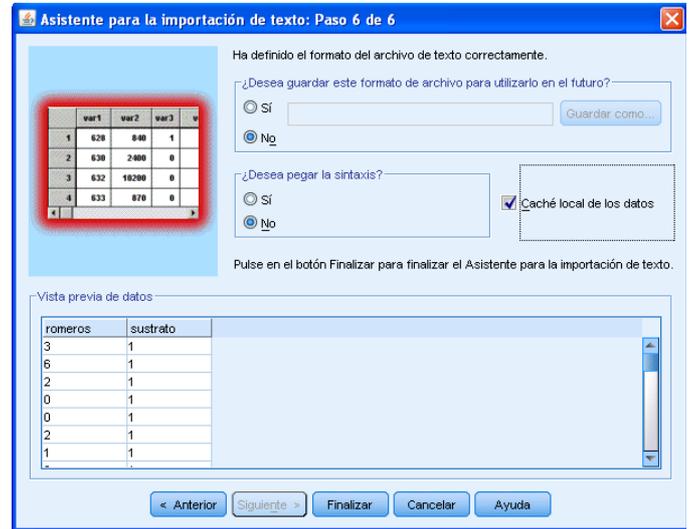
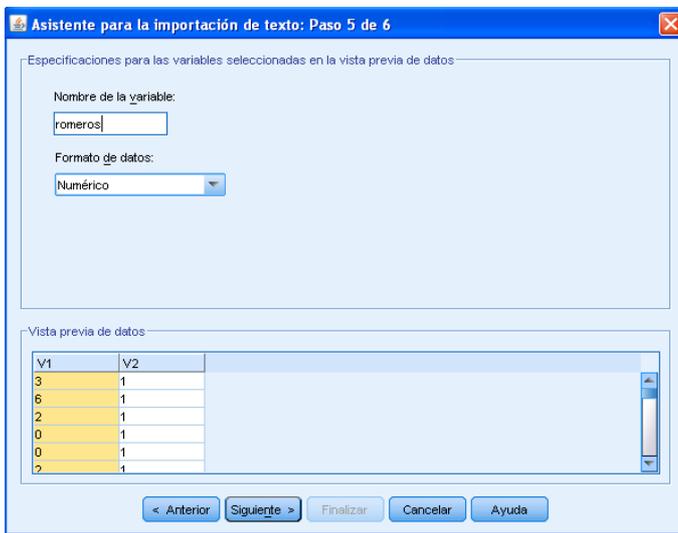
Análisis de la influencia del tipo de suelo sobre la abundancia de romero (*Rosmarinus officinalis*) en el Parque Regional de Calblanque. Se toman 20 unidades de muestreo de 2x2 m² sobre sustrato de pizarras (sust = 1) y 15 unidades sobre sustrato calizo (sust = 0). Los datos obtenidos fueron los siguientes:

1	3	6	2	0	0	2	1	0	0	0	2	2	4	5	1	1	2	2	4	0
0	1	9	6	16	12	7	11	12	2	6	10	9	7	3	4					

Guión para realizar la práctica:

1. Descargar los datos de la red al escritorio o a una carpeta del disco duro. Se trata de un fichero romero.txt de texto en el que aparecen la cantidad de romeros en cada unidad experimental por tipo de suelo.
2. Entrar en Inicio → Todos los Programas → SPSS Inc → PASW Statistics 18. Cerrar el primer cuadro de dialogo e ir a Archivo → Abrir → Datos Archivos de tipo: Todos los archivos (o Texto (*.txt,*.dat)) y abrir el fichero romero.txt. Seguir los siguientes pasos:





En el último paso, seleccionar Finalizar:

Aparecen dos pantallas que se mantendrán durante todo el manejo de SPSS:

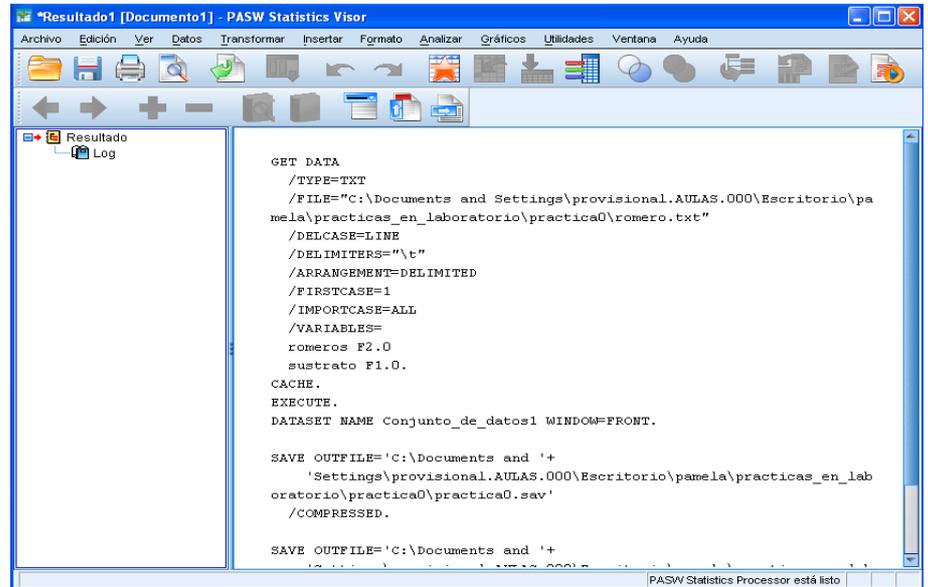
El editor de datos (a la izquierda) y el Visor de resultados (a la derecha):

*Sin título2 [Conjunto_de_datos1] - PASW

Archivo Edición Ver Datos Transformar

1 : romeros 3

	romeros	sustrato
1	3	1
2	6	1
3	2	1
4	0	1
5	0	1
6	2	1
7	1	1
8	0	1
9	0	1
10	0	1
11	2	1
12	2	1
13	4	1
14	5	1
15	1	1



Desde el Visor de Datos es conveniente seleccionar Archivo → Guardar como ... y guardar el fichero como Practica1.sav (que es el formato de fichero de datos en SPSS).

En el Visor de Resultados irán apareciendo los resultados de nuestro análisis estadístico (tablas, gráficos,...). Los resultados se pueden guardar al final en un fichero llamado Practica1.spv. Las tablas se pueden copiar en diversos formatos (por ejemplo, como texto o como tabla para Word) y los gráficos se pueden extraer también en varios formatos (por ejemplo, .eps o .pdf).

En la parte inferior del editor aparecen dos pestañas: Vista de datos y Vista de variables.

La pestaña Vista de datos nos muestra los nombres y valores de las variables (se ve arriba a la izquierda).

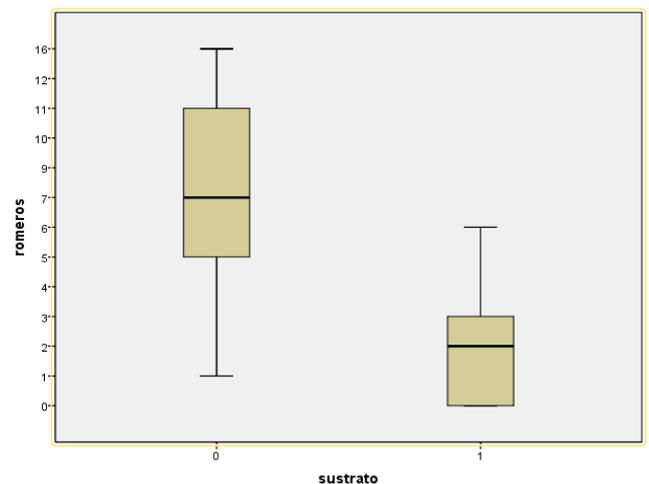
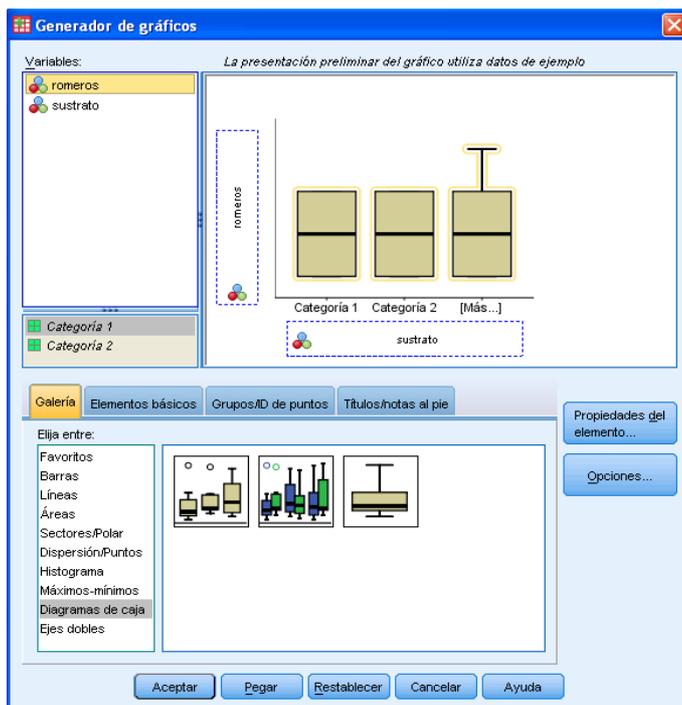
La pestaña Vista de variables muestra las características de cada variable (Nombre, Tipo, Anchura, ...).

	Nombre	Tipo	Anchura	Decimales	Etiqueta
1	romeros	Numérico	2	0	
2	sustrato	Numérico	1	0	
3					
4					

Comenzamos nuestro análisis.

3. En el editor o en el visor seleccionar Gráficos → Generador de gráficos → Diagramas de caja

Arrastrar la variable respuesta (romeros) al eje vertical del gráfico y la el factor (sustratos) al eje horizontal. Pinchar en Aceptar. Haciendo doble clic en el gráfico resultante aparece el Editor de gráficos, que nos permite modificar los límites de los ejes, colores, etc. El gráfico resultante es el que aparece a la derecha



Cuestión 1: ¿Qué observas en el gráfico de caja de los datos?

Responder en hoja adjunta: respuestas.doc

Hay otra manera de obtener el mismo gráfico y los estadísticos descriptivos (media, mediana, ...) de la variable respuesta calculados por grupos. Ir a Analizar → Estadísticos descriptivos → Explorar.

El cuadro de diálogo que aparece es el siguiente, en el que se pueden dejar todas las opciones seleccionadas por defecto o cambiarlas.



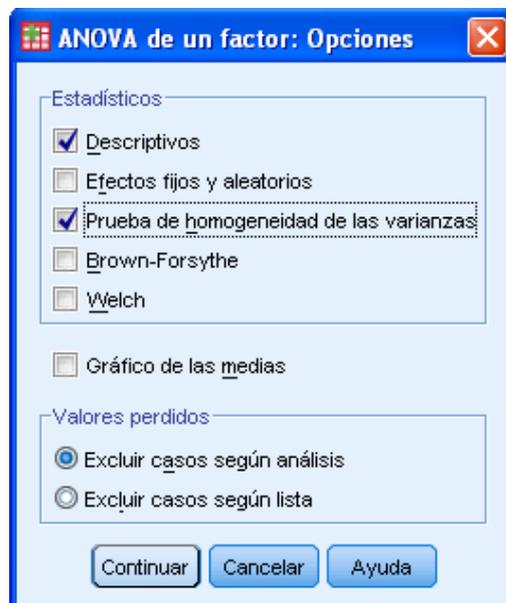
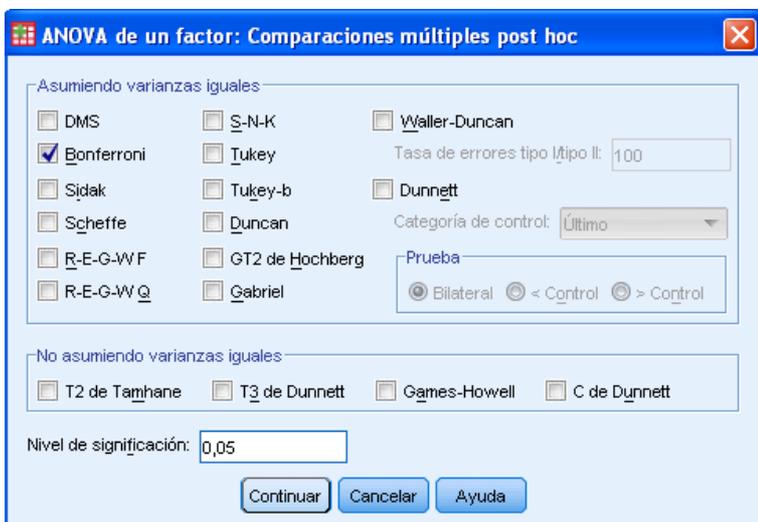
4. Para ver si hay diferencias entre la cantidad de romeros en función del tipo de suelo realizamos un análisis de la varianza con un factor. Para ello, en el Editor de datos o en el Visor de resultados seleccionar Analizar → Comparar medias → ANOVA de un factor.

El cuadro de diálogo que aparece es el siguiente.



El botón Contrastes... no hay que tocarlo.

En Post hoc... marcar la opción Bonferroni y seleccionar el nivel de significación global α_T que deseamos (0,05 por ejemplo). En Opciones marcar Estadísticos descriptivos y Prueba de



Descriptivos

romeros

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95%		Mínimo	Máximo
					Límite inferior	Límite superior		
0	15	7,67	4,186	1,081	5,35	9,98	1	16
1	20	1,85	1,785	,399	1,01	2,69	0	6
Total	35	4,34	4,186	,708	2,90	5,78	0	16

homogeneidad de varianzas. (Nos dará el valor del estadístico de de Levene el cual tiene distribución F y su p-valor). No cambiar nada donde pone Valores perdidos. No marcar la opción Gráfico de las medias.

Al presionar Aceptar obtenemos la tabla ANOVA (donde pone Sig. ha calculado el p-valor del contraste).

Prueba de homogeneidad de varianzas

romeros

Estadístico de Levene	gl1	gl2	Sig.
11,756	1	33	,002

ANOVA

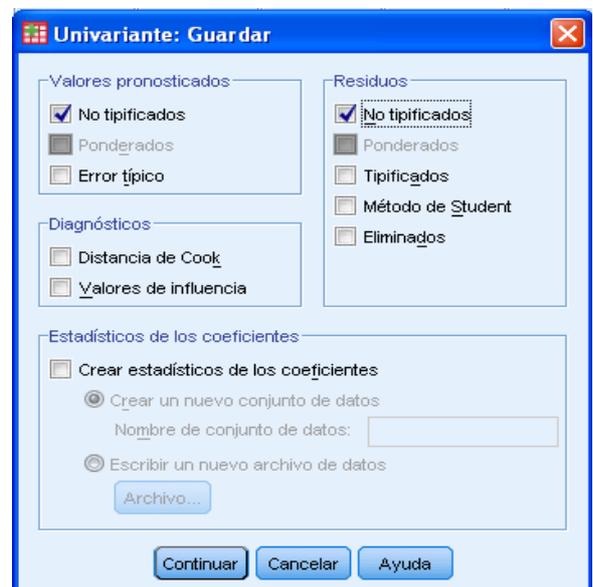
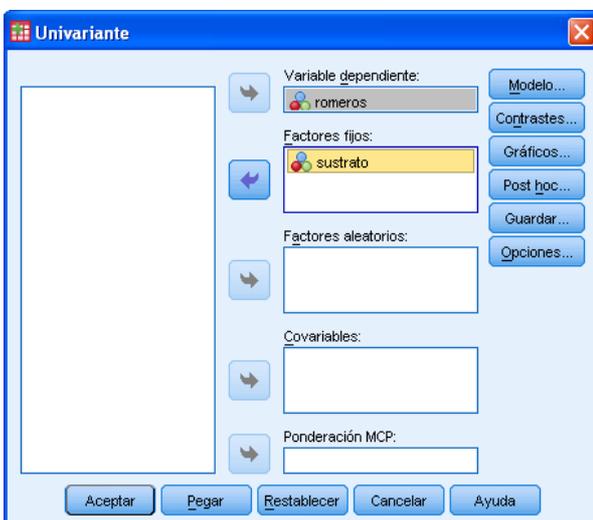
romeros

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	290,002	1	290,002	31,287	,000
Intra-grupos	305,883	33	9,269		
Total	595,886	34			

Cuestión 2: ¿Qué conclusión se obtiene de la tabla ANOVA?

Responder en hoja adjunta: respuestas.doc

El procedimiento ANOVA unifactorial de SPSS no calcula los residuos. Para calcularlos utilizamos el procedimiento del modelo lineal general. Pinchar en Analizar → Modelo lineal general → Univariante.

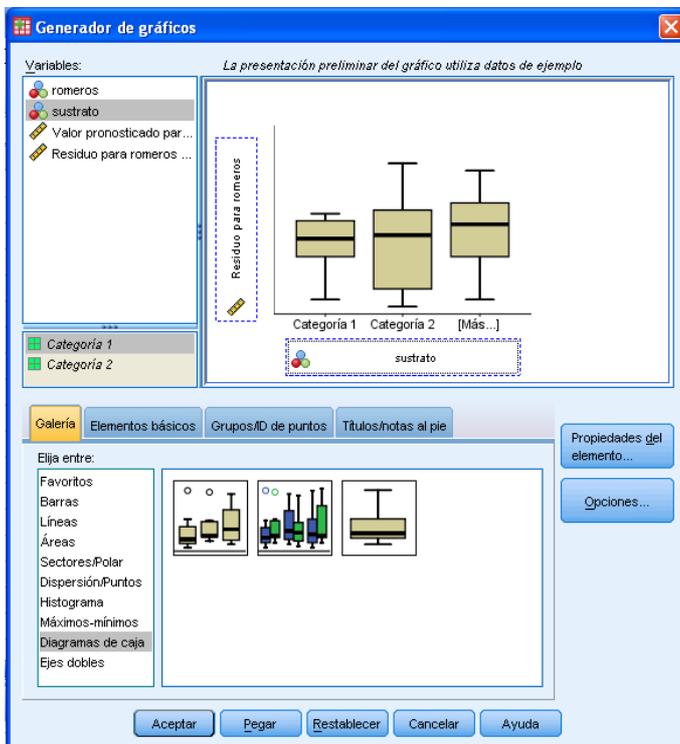


Pasar romeros a Variable dependiente y sustrato a Factor. Pinchar en Guardar y elegir Valores pronosticados no tipificados y Residuos no tipificados. Pinchar en Continuar y luego en Aceptar.

Vemos que en el editor de datos han aparecido dos nuevas columnas con los valores de las predicciones (PRE_1) y de los residuos (RES_1).

	romeros	sustrato	PRE_1	RES_1
1	3	1	1,85	1,15
2	6	1	1,85	4,15
3	2	1	1,85	,15
4	0	1	1,85	-1,85
5	0	1	1,85	-1,85
6	2	1	1,85	,15
7	1	1	1,85	-,85
8	0	1	1,85	-1,85
9	0	1	1,85	-1,85
10	0	1	1,85	-1,85

Para representar los residuos frente a valores previstos pinchar en Gráficos → Generador de gráficos y elegir Diagrama de caja. Arrastrar residuo para romeros al eje vertical del gráfico y sustrato al eje horizontal. Pinchar en Aceptar. Se obtiene el diagrama de cajas de la derecha.

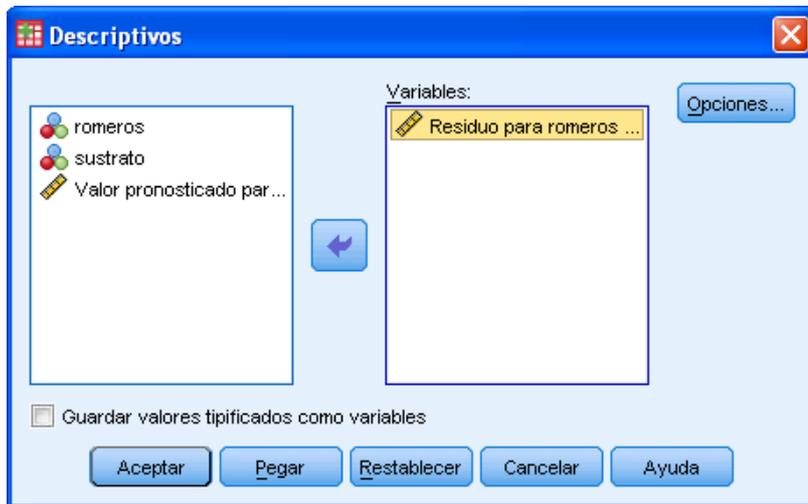


Para dibujar un gráfico de probabilidad de los residuos pinchar en Analizar → Estadísticos descriptivos → Gráficos P-P. Obtendremos el siguiente gráfico

Para dibujar un histograma de los residuos pinchar en **Analizar** → **Estadísticos descriptivos** → **Frecuencias**, pasar residuo para romeros a **Variables**, luego en **Gráficos**, tildar las opciones **Histograma** y **Mostrar curva de distribución normal** y se obtendrá el histograma.



Para calcular el coeficiente de asimetría y la curtosis de los residuos pinchar en **Analizar** → **Estadísticos descriptivos** → **Descriptivos** y pasar el Residuo a **Variables**. Pinchar en el botón **Opciones** y seleccionar los estadísticos descriptivos de interés.



Obtendremos una tabla de este estilo:

Estadísticos descriptivos

	N	Media	Desv. típ.	Asimetría		Curtosis	
	Estadístico	Estadístico	Estadístico	Estadístico	Error típico	Estadístico	Error típico
Residuo para romeros	35	,0000	2,99943	,279	,398	,998	,778
N válido (según lista)	35						

Como regla aproximada, si el valor absoluto de la **asimetría** dividida por su **error típico** es mayor que 1,96 entonces la distribución de los residuos es asimétrica. Asimismo, si el valor absoluto de la **curtosis** dividida por su **error típico** es mayor que 1,96 entonces la distribución de los residuos no sigue una distribución normal.

Cuestión 3: ¿Qué conclusión obtienes de los gráficos de los residuos?

Responder en hoja adjunta: respuestas.doc