



CUADERNO DE PRÁCTICAS

PRÁCTICAS DE LA ASIGNATURA:

BIOESTADÍSTICA

2º Cuatrimestre

Prácticas de Bioestadística mediante ordenador

Guía interactiva de Autoaprendizaje de SPSS:

(<http://www.ugr.es/~bioestad/guiaspss/index.html>)

Práctica 1. Introducción al paquete estadístico SPSS (1 hora).

Ejercicios guiado y propuestos de la práctica 1 de la Guía interactiva de Autoaprendizaje de SPSS

Práctica 2. Estadística descriptiva (1 hora)

Ejercicios guiado y propuestos de la práctica 2 de la Guía Interactiva

Práctica 3. Seminario 1: Análisis estadístico de datos reales: Análisis descriptivo. (1 hora)

Se utiliza el fichero de datos: DATOS-ECOLOGIAAlumnos.xls

Práctica 4. Regresión (1 hora)

Ejercicio guiado de la práctica 3 de la Guía interactiva

Práctica 5. Regresión (1 hora)

Ejercicios propuestos de la práctica 3 de la Guía interactiva

Práctica 6. Distribuciones Discretas y Continuas de probabilidad (1 hora)

Ejercicios guiados y propuestos de la práctica 4 de la Guía interactiva

Práctica 7. Intervalos de confianza (1 hora)

Ejercicio guiado y propuesto de la práctica 5 de la Guía interactiva

Práctica 8. Contrastes de hipótesis (1 hora)

Ejercicios guiados de la práctica 6 de la Guía interactiva

Práctica 9. Contrastes de hipótesis (1 hora)

Ejercicios guiados de la práctica 6 de la Guía interactiva

Práctica 10. Contrastes de hipótesis (1 hora)

Ejercicios propuestos de la práctica 6 de la Guía interactiva

Práctica 11. Introducción al diseño estadístico de experimentos: Diseño completamente aleatorizado (1 hora)

Ejercicios guiado y propuesto de la práctica 7 de la Guía interactiva

Práctica 12. Seminario 2: Análisis estadístico de datos reales: Análisis inferencial. (1 hora)

Aplicación sobre los datos reales de la simetría de la hoja de la encina (DATOS-ECOLOGIAAlumnos.xls)

Bibliografía: Lara Porras, A.M., Román-Montoya, Y, y Pérez Bueno, F. (2012). **Guía interactiva de autoaprendizaje de SPSS versión 4.1** (CD-ROM). Editorial Proyecto Sur. Granada (España). (<http://www.ugr.es/~bioestad/guiaspss/index.html>)

Práctica 1. Introducción al paquete estadístico de SPSS

Ejercicio Guiado de la Práctica 1 (Guía SPSS. Vr. 4.1)

Los siguientes datos muestran información sobre un grupo de opositores:

Sexo	Edad	Estado civil	Puntuación examen
Mujer	27	Soltero	7,3
Hombre	30	Soltero	4,7
Mujer	22	Casado	9,9
Hombre	27	Separado	11,6
Mujer	31	Separado	4,9
Hombre	22	Soltero	8,6
Mujer	21	Casado	7,7
Hombre	24	Casado	5,9
Hombre	29	Casado	10,7
Hombre	32	Viudo	8,8
Mujer	29	Casado	7,4
Hombre	24	Soltero	3,3
Mujer	28	Casado	2,3
Mujer	31	Separado	4,6
Mujer	37	Viudo	5,1
Hombre	22	Soltero	9,2
Mujer	28	Casado	4,7
Hombre	24	Casado	4,5

Se pide:

- Introducir los nombres y características de cada una de las variables en **Vista de variables** (las variables *sexo* y *estado civil* tienen etiquetas de valor).
- Introducir los datos de cada una de las variables en el **Editor de datos SPSS**
- Presentar las variables con sus respectivas **Etiquetas de valor** en el **Editor de datos**
- Introducir una nueva variable con el nombre *cálculo* igual al doble de la variable *edad* más el triple de la variable *puntuación*
- Ordenar de forma creciente los datos según la *edad* de los individuos
- Seleccionar en la variable *puntuación* aquellos casos que tengan un valor mayor que 5.

Práctica 1. Introducción al paquete estadístico de SPSS

Ejercicio Propuesto 1 de la práctica 1 (Guía SPSS. Vr. 4.1)

Los siguientes datos proceden de un grupo de estudiantes

Nombre	Sexo	Edad	Estatura	ColorPelo	FechaNac	LugarNac
Felipe	Hombre	18	1,72	Moreno	11/23/1986	Granada
Pedro	Hombre	20	1,91	Rubio	3/11/1985	Granada
Ana	Mujer	18	1,67	Pelirrojo	4/1/1987	Sevilla
Ismael	Hombre	17	1,82	Castaño	9/9/1987	Jaén
Eloy	Hombre	19	1,77	Rubio	4/22/1986	Granada
Julia	Mujer	19	1,81	Rubio	9/11/1985	Jaén
Eva	Mujer	19	1,73	Moreno	2/28/1986	Málaga
José	Hombre	18	1,84	Castaño	1/1/1987	Granada
Patricia	Mujer	18	1,87	Pelirrojo	9/27/1988	Granada
Miriam	Mujer	21	1,60	Moreno	3/3/1984	Granada

Se pide:

- Introducir los nombres y características de cada una de las variables en **Vista de variables** (las variables *sexo*, *colorPelo* y *LugarNac* tienen etiquetas de valor)
- Introducir los datos de cada una de las variables en el **Editor de datos SPSS**
- Presentar las variables con sus respectivas **Etiquetas de valor** en el **Editor de datos SPSS**
- Introducir una nueva variable con el nombre *cálculo* igual al triple de la variable *Edad* menos el doble de la variable *Estatura*
- Ordenar de forma decreciente los datos según la *Estatura* de los individuos
- Seleccionar en la variable *Estatura* aquellos casos que tengan un valor menor que 1.77.

Práctica 1. Introducción al paquete estadístico de SPSS

Ejercicio Propuesto 2 de la práctica 1 (Guía SPSS. Vr. 4.1)

Los siguientes datos proceden de un grupo de opositores

Sexo	Edad	Estatura	Puntuación 1	Puntuación 2	Calificación
Mujer	25	1,82	6	9	Aprobado
Hombre	30	1,91	5	7	Aprobado
Mujer	28	1,77	5	4	Suspenso
Mujer	27	1,82	7	5	Aprobado
Hombre	37	1,8	3	4	Suspenso
Mujer	29	1,82	5	8	Aprobado
Hombre	31	1,79	7	6	Aprobado
Hombre	28	1,89	5	5	Aprobado
Hombre	38	1,87	6	5	Aprobado
Mujer	28	1,79	4	4	Suspenso
Mujer	32	1,81	6	7	Aprobado
Hombre	35	1,88	7	6	Aprobado
Hombre	28	1,70	4	2	Suspenso
Mujer	21	1,72	3	4	Suspenso
Mujer	21	1,72	5	4	Suspenso

Datos 1

Sexo	Edad	Estatura	Puntuación1	Puntuación2	Calificación
Hombre	55	1,92	5	4	Suspenso
Hombre	32	1,81	5	6	Aprobado
Mujer	29	1,79	8	7	Aprobado
Hombre	37	1,83	7	5	Aprobado
Mujer	33	1,82	7	9	Aprobado
Mujer	29	1,72	9	8	Aprobado
Hombre	32	1,89	4	3	Suspenso
Mujer	27	1,89	6	6	Aprobado
Hombre	38	1,67	5	7	Aprobado
Hombre	29	1,90	3	4	Suspenso
Mujer	30	1,81	4	4	Suspenso
Hombre	35	1,88	6	5	Aprobado
Mujer	38	1,76	7	1	Suspenso
Hombre	25	1,92	8	6	Aprobado
Mujer	21	1,72	3	4	Suspenso

Datos 2

Se pide:

- Introducir los nombres y características de cada una de las variables en **Vista de variables** (las variables *Sexo* y *Calificación* tienen etiquetas de valor)
- Introducir los datos de cada una de las variables en el *Editor de datos SPSS*

- c) Presentar las variables con sus respectivas **Etiquetas de valor** en el *Editor de datos SPSS*
- d) Introducir una nueva variable, en el fichero de Datos1, con el nombre *Media* que recoja la media de las variables Puntuación1 y Puntuación2. Guardar el nuevo fichero con el nombre de Datos3
- e) Ordenar de forma creciente según la *Edad* de los individuos en el fichero de Datos3
- f) Generar un nuevo fichero de datos denominado Datos4 que contenga todos los opositores de los ficheros Datos3 y Datos2
- g) En el fichero Datos4 seleccionar sólo los casos en que los opositores estén aprobados
- h) Seleccionar en el fichero de Datos4 sólo los casos que tengan una Puntuación1 superior a 6 y una Puntuación2 inferior a 5
- i) Seleccionar en el fichero de Datos4 sólo los casos que tengan una Puntuación1 superior a 6 o una Puntuación2 inferior a 5
- j) Seleccionar en el fichero de Datos4 los opositores que sean mujeres y calcular la media de las variables Puntuación1 y Puntuación2
- k) En el fichero Datos4 seleccionar sólo los casos de los aprobados en que los opositores sean hombres

Práctica 2. Estadística descriptiva

Ejercicio Guiado de la Práctica 2 (Guía SPSS. Vr. 4.1)

Para estudiar el efecto de las aguas residuales de las alcantarillas que afluyen a un lago, se toman 40 muestras distintas en las que se mide la concentración de nitrato en el agua. Los datos obtenidos son los siguientes:

x_i	25	30	40	75	80	120	150	200
n_i	3	3	5	6	5	7	6	5

1. Introducir los datos
2. Determinar:
 - Tabla de frecuencias
 - Número de datos
 - Percentil 30
 - Valores máximo y mínimo
 - Media
 - Mediana
 - Moda
 - Desviación típica
 - Varianza
 - Histograma de frecuencias
 - Histograma de frecuencias interactivo con las siguientes características:
 1. Representación horizontal.
 2. **Efecto 3D.**
 3. Considerar 5 clases.
 4. **Título:** Histograma.
Subtítulo: Efecto 3D
Pie: Concentración
Aspecto: Acero

Práctica 2. Estadística descriptiva

Ejercicio propuesto1 de la Práctica 2 (Guía SPSS. Vr. 4.1)

Las autoridades sanitarias de un municipio están interesadas en evaluar la calidad del agua para consumo, en términos de colonias de bacterias tróficas, en un acuífero próximo a la ciudad. Se consideran dos zonas diferentes del acuífero y se obtienen los siguientes resultados (número de colonias por 1000 mm de agua):

zona 1: 194 199 191 202 215 214 197 204 199 202 230 193 194 209

zona 2: 158 161 143 174 220 156 156 156 198 161 188 139 147 116

Se pide:

Un estudio comparativo de la calidad del agua en ambas zonas utilizando medidas estadísticas y gráficos. Para ello realizar los siguientes apartados:

1. Estudiar media, varianza, moda, mediana, máximo y mínimo de ambas zonas.
2. Representar gráficamente los datos con un histograma para cada variable.
3. Representar gráficamente los datos con un diagrama de caja simple (donde los datos del gráfico son resúmenes para distintas variables).

Ejercicio propuesto2 de la práctica 2 (Guía SPSS. Vr. 4.1)

Un agrónomo mide el contenido de humedad en una variedad de trigo tras ser secados especialmente. Para ello, hace las mediciones en 56 muestras de una tonelada de trigo. Los resultados se muestran en la tabla adjunta

Contenido humedad	7.2	7.1	7.3	7.4	6.8	6.7	6.9	7.6	7.5
Número	3	6	13	4	5	2	7	9	8

Se pide:

- a) Resumen estadístico: Tabla de frecuencias, estadísticos de tendencia central, de dispersión, de posición (Percentil 60)
- b) Gráfico de tallo y hojas
- c) Gráfico de cajas
- d) Estratificar la población en tres grupos de Contenido de humedad
- e) Determinar la media, mediana y desviación típica del *Contenido de humedad* en cada grupo
- f) Construir el gráfico de barras y de sectores para la variable de agrupación.

Práctica 3. Seminario 1. Análisis estadístico de datos reales: Análisis descriptivo.

Con el fichero de datos (DATOS-ECOLOGIAAlumnos.xls) relativos al estudio de la simetría de las hojas de encina en distintas zonas de la provincia de Granada realiza las siguientes acciones:

1. Construye un fichero con las siete primeras variables, introduciendo los nombres y las características de cada variable.

Zona	Parte	Año	Hoja	Longitud	Anchura izq.	Anchura der.
-------------	--------------	------------	-------------	-----------------	-------------------------	-------------------------

2. Introduce los datos de estas variables.
3. Introduce las variables.

Dif. I-D	 Dif. I-D 	Anchura total	Asimetría
-----------------	-------------------	----------------------	------------------

calculando su expresión en función de los datos introducidos.

4. Selecciona aquellos casos cuyo año sea superior a 1996.
5. Genera una tabla de frecuencias para la variable “zona”.
6. Genera una tabla de frecuencias para la variable “longitud”.
7. Calcula media, varianza, moda, mediana, máximo, mínimo asimetría y curtosis de la longitud de la hoja.
8. Calcula lo mismo del apartado anterior pero distinguiendo según la zona (realiza también un gráfico de Box-Whisker).

Práctica 4. Regresión

Ejercicio Guiado de la práctica 3 (Guía SPSS. Vr. 4.1)

Se realiza un estudio para investigar la relación entre el nivel de humedad del suelo y la tasa de mortalidad en lombrices de tierra. La tasa de mortalidad, y , es la proporción de lombrices de tierra que mueren tras un periodo de dos semanas. El nivel de humedad, x , viene medido en milímetros de agua por centímetro cuadrado de suelo. Se obtuvieron los siguientes datos:

x	y
0,000	0,5
0,000	0,4
0,000	0,5
0,316	0,2
0,316	0,3
0,316	0,3
0,632	0,0
0,632	0,1
0,632	0,0
0,947	0,1
0,947	0,2
0,947	0,1
1,260	0,6
1,260	0,5
1,260	0,4

Se pide:

- Nube de puntos.
- Rectas de regresión
- Coefficiente de correlación
- Grado de ajuste
- Interpretación

Práctica 5. Regresión

Ejercicio Propuesto1 de la práctica 3 (Guía SPSS. Vr. 4.1)

Se realiza un estudio para establecer una ecuación mediante la cual se pueda utilizar la concentración de estrona en saliva (x), para predecir la concentración de estrona en plasma libre (y). Se obtuvieron los siguientes datos de 14 hombres sanos:

x	7.4	7.5	8.5	9	9	11	13	14	14.5	16	17	18	20	23
y	30	25	31.5	27.5	39.5	38	43	49	55	48.5	51	64.5	63	68

Se pide:

- Diagrama de dispersión
- Recta de regresión de la concentración de estrona en plasma libre en función de la concentración de estrona en saliva. Estudiar la bondad del ajuste
- Regresión parabólica. Estudiar la bondad del ajuste.

Ejercicio Propuesto2 de la práctica 3 (Guía SPSS. Vr. 4.1)

Se realiza un estudio para investigar la relación entre el nivel de humedad del suelo y la tasa de mortalidad en lombrices. La tasa de mortalidad, Y , es la proporción de lombrices de tierra que mueren tras un periodo de dos semanas; el nivel de humedad, X , viene medido en milímetros de agua por centímetro cuadrado de suelo. Los datos se muestran en la siguiente tabla.

x	0.31	0.31	0.56	0.56	0.89	0.89	0.96	0.96	1.15	1.15	1.25
y	0.2	0.1	0.0	0.2	0.3	0.5	0.0	0.6	0.4	0.2	0.5

Se pide:

- ¿Muestran los datos una tendencia lineal?
- Determinar la recta de regresión Y/X , el grado de asociación lineal entre la tasa de mortalidad y el nivel de humedad y la bondad del ajuste realizado en la recta de regresión. ¿Cuánto explica el modelo?
- Mediante la línea de regresión estimada, predecir el nivel de humedad del suelo si la tasa de mortalidad de las lombrices es 0.7
- Determinar el coeficiente de correlación lineal de las rectas de regresión Y/X y X/Y
- Ajustar los datos mediante una regresión curvilínea
- ¿Qué ajuste es mejor ¿Lineal? ¿Curvilíneo?

Práctica 6. Distribuciones de probabilidad

Ejercicio Guiado 1 de la práctica 4 (Guía SPSS. Vr. 4.1)

Se pretende comprobar la efectividad de una determinada vacuna contra la gripe. Para ello se administra dicha vacuna a un grupo de 15 pacientes. La probabilidad de que el paciente vacunado contraiga la gripe es 0.3.

Calcula las siguientes probabilidades:

- Ningún paciente contraiga la gripe.
- Más de dos pacientes contraigan la gripe.
- Contraigan la gripe entre tres y cinco pacientes, ambos inclusive.
- Generar una muestra aleatoria de tamaño 20 de valores de una distribución Binomial de parámetros $n = 10$ y $\text{prob} = 0.2$.

Ejercicio Guiado 2 de la práctica 4 (Guía SPSS. Vr. 4.1)

En un servicio de urgencias de un determinado hospital se sabe que por término medio llegan diez pacientes durante una hora.

Calcula la probabilidad de que:

- Lleguen exactamente cinco pacientes en una hora.
- Lleguen menos de quince pacientes en dos horas.
- Lleguen más de cuatro y menos de ocho pacientes en una hora.
- Generar una muestra de tamaño 15 para una distribución de Poisson de parámetro media igual a 30.

Ejercicio Guiado 3 de la práctica 4 (Guía SPSS. Vr. 4.1)

Se ha estudiado el nivel de glucosa en sangre en ayunas en un grupo de diabéticos. Esta variable se supone que sigue una distribución Normal, con media 106 mg/100 ml y desviación típica 8 mg/100 ml.

Se pide:

- Obtener la probabilidad de que el nivel de glucosa en sangre en un diabético sea inferior a 120 mg/100 ml.
- ¿Qué porcentaje de diabéticos tienen niveles de glucosa en sangre comprendidos entre 90 y 130 mg/100 ml?
- Hallar el valor de la variable caracterizado por la propiedad de que el 25% de todos los diabéticos tiene un nivel de glucosa en ayunas inferior a dicho valor.
- Generar una muestra de tamaño 12 para la una distribución Normal con media igual a 5 y desviación típica igual a 3.

Práctica 6. Distribuciones de probabilidad

Ejercicio Propuesto 1 de la práctica 4 (Guía SPSS. Vr. 4.1)

Una prueba de laboratorio para detectar heroína en sangre tiene un 92% de precisión. Si se analizan 72 muestras en un mes.

Calcular la probabilidad de que:

- 60 o menos estén correctamente evaluadas.
- Menos de 60 estén correctamente evaluadas.
- Exactamente 60 estén correctamente evaluadas.
- Generar una muestra de tamaño 12.

Ejercicio Propuesto 2 de la práctica 4 (Guía SPSS. Vr. 4.1)

En una cierta población se ha observado que el número medio anual de muertes por cáncer de pulmón es 12. Si el número de muertes causadas por la enfermedad sigue una distribución de Poisson, calcular la probabilidad de que:

- a) Haya exactamente 10 muertes por cáncer de pulmón en un año.
- b) 15 o más personas mueran a causa de la enfermedad durante un año.
- c) 10 o menos personas mueran a causa de la enfermedad en 6 meses.

Ejercicio Propuesto 3 de la práctica 4 (Guía SPSS. Vr. 4.1)

En cierta especie de aves, se ha detectado una contaminación apreciable de mercurio (Hg) en sangre. La concentración de mercurio en sangre está distribuida normalmente con media 0.25 ppm (partes de Hg por millón, en plasma) y desviación típica 0.08 ppm.

- a) ¿Cuál es la probabilidad de que un ave presente un nivel de mercurio en sangre superior a 0.40 ppm ?
- b) ¿Cuál es la probabilidad de que un ave tenga un nivel de mercurio en sangre entre 0.20 y 0.50 ppm?
- c) ¿Cuál es el nivel máximo de concentración de mercurio en sangre del 40% de las aves menos contaminadas?
- d) Generar una muestra de tamaño 10.

Práctica 7. Intervalos de confianza.

Ejercicio Guiado 1 de la práctica 5 (Guía SPSS. Vr. 4.1)

Se desea comprobar el efecto de 2 fertilizantes (A y B) sobre la producción de unos árboles frutales, para ello se toman dos grupos de 8 y 10 árboles seleccionados aleatoriamente y se le añade al agua de riego de cada uno de los grupos de árboles el fertilizante A y B, respectivamente. La producción en ese año fue la siguiente (en Kg):

Fertilizante A	30	25	28	29	30	31	24	22	25	27
Fertilizante B	28	27	28	28	26	27	26	29		

Se pide:

1. Obtener un intervalo de confianza al 99% para la producción media de los árboles tratados con el Fertilizante A y para la producción media de los árboles tratados con el Fertilizante B.
2. Obtener un intervalo de confianza al 98% para la diferencia entre la producción media de los árboles tratados con el Fertilizante A y con el Fertilizante B.
3. La producción de los árboles tratados con el Fertilizante A en el año anterior viene reflejada en la siguiente tabla. Obtener un intervalo de confianza al 99% para la diferencia de medias, en la producción antes y después de tratar los árboles con dicho fertilizante.

Antes	25	20	25	28	30	30	26	15	18	22
Después	30	25	28	29	30	31	24	22	25	27

Ejercicio Propuesto 1 de la práctica 5 (Guía SPSS. Vr. 4.1)

Se desea estudiar si la longitud del pico en una especie de loro es distinta entre los machos y las hembras. Para ello se selecciona una muestra de 14 machos y 12 hembras, cuyos resultados, expresados en milímetros, se muestran en la siguiente tabla:

Machos	57	58	60	58	61	62	61	59	57	63	58	55	59	60
Hembras	55	56	58	54	53	55	57	53	54	54	55	55		

Se pide:

- a) Obtener un intervalo de confianza a un nivel del 99%.
 - para la longitud media del pico en los machos.
 - para la longitud media del pico en las hembras.
- b) Obtener un intervalo de confianza a un nivel del 99% para la diferencia entre la longitud media del pico de los machos y de las hembras.

Práctica 7. Intervalos de confianza.

Ejercicio Propuesto 2 de la práctica 5 (Guía SPSS. Vr. 4.1)

Para comprobar si un determinado pienso puede mejorar la producción de lana de las ovejas, se selecciona una muestra aleatoria simple de 10 ovejas para ser alimentadas con dicho pienso. En la tabla siguiente se muestra el peso (en Kgr) de la lana producida antes y después del experimento

Antes	10	8	7	5	9	12	10	9	8	8
Después	10	9	9	7	10	12	11	12	11	10

Obtener un intervalo de confianza al 98% para la diferencia de los pesos medios de la lana producida antes y después del experimento.

Práctica 8. Contraste de hipótesis

Ejercicio Guiado 1 de la práctica 6 (Guía SPSS. Vr. 4.1)

En 5 zonas de la provincia de Granada (Ladikhonda y Fazares, zonas muy secas y Cortijuela, Molinillo y Fardes, zonas húmedas) se hacen una serie de mediciones sobre las hojas de las encinas a lo largo de 3 años consecutivos: 1995, muy seco y 1996 y 1997, muy lluviosos.

El objetivo es medir la simetría fluctuante en dichas hojas como indicador de stress en la planta. Bajo condiciones de stress (sequía, herbivoría, limitación por nutrientes...), la hipótesis es que la asimetría aumente. Contamos con la siguiente información:

- Localización árboles: 5 zonas, dos en zonas muy secas (Hoya Guadix-Baza, Ladikhonda y Fazares) y tres en zonas con mayor precipitación (Cortijuela, Molinillo, Fardes). En esta última, Fardes, son árboles situados en la ladera de un río (presumiblemente poco afectados por años más o menos secos).
- Años de climatología diferente: 1995 año muy seco y años 1996 y 1997, años muy lluviosos.
- Situación de la hoja: Canopy (copa de los árboles) y Sprouts (rebrotos, hojas nuevas que salen desde la parte inferior del tronco).

Disponemos de un total de 2101 casos, cedidos por el Departamento de Ecología de la Universidad de Granada (España), de los que hemos seleccionado aleatoriamente una muestra de tamaño 15 que se presenta en la siguiente tabla:

Zona	Parte	Año	Longitud	Asimetría
Cortijuela	Canopy	1995	26,51	0,028
Cortijuela	Canopy	1996	30,17	0,010
Molinillo	Canopy	1995	34,24	0,080
Molinillo	Canopy	1996	31,04	0,340
Molinillo	Canopy	1996	34,99	0,087
Fardes	Canopy	1995	30,48	0,040
Fardes	Canopy	1996	25,07	0,010
Ladikhonda	Canopy	1995	25,04	0,021
Ladikhonda	Canopy	1996	29,16	0,135
Fazares	Canopy	1995	35,12	0,010
Fazares	Canopy	1996	25,41	0,094
Fazares	Canopy	1996	27,02	0,153
Cortijuela	Sprouts	1995	23,04	0,156
Fazares	Sprouts	1995	27,69	0,172
Fazares	Sprouts	1996	34,71	0,077

Se pide:

- a) ¿Se puede admitir que la longitud de las hojas de encina se distribuye normalmente?
- b) ¿Se puede admitir que la longitud media de las hojas es igual a 30 cm a un nivel de significación del 5%?

- c) Suponiendo que la asimetría de las hojas sigan una distribución Normal; comprobar mediante un contraste de hipótesis si existen diferencias significativas en la asimetría de las hojas teniendo en cuenta la situación de la hoja en el árbol.
- d) A un nivel de significación del 5%, ¿es representativo el ajuste lineal entre la longitud y la asimetría? ¿Cuál sería la expresión del modelo? ¿Cuánto explica el modelo?

Ejercicio Guiado 2 de la práctica 6 (Guía SPSS. Vr. 4.1)

Se realiza un estudio para investigar el efecto del ejercicio físico sobre el nivel de triglicéridos, en el que participaron once individuos. Antes del ejercicio se tomaron muestras de sangre para determinar el nivel de triglicéridos de cada participante. Después, los individuos fueron sometidos a un programa físico que se centraba en carreras y marchas diarias. Al final del periodo de ejercicios, se tomaron nuevamente muestras de sangre y se obtuvo una segunda lectura del nivel de triglicéridos en sangre. Los datos se muestran en la siguiente tabla

	Nivel de triglicéridos										
Nivel previo	198	210	194	220	138	220	219	161	210	313	226
Nivel posterior	65	77	94	73	37	131	77	24	99	321	57

Se pensó que el programa de ejercicios físicos podría reducir del nivel de triglicéridos en sangre. ¿Sostienen estos datos el argumento de los investigadores? (Supóngase normalidad).

Ejercicio Guiado 3 de la práctica (Guía SPSS. Vr. 4.1)6

Se realiza un estudio para investigar el efecto de la presencia de una gran planta industrial sobre la población de invertebrados en un río que atraviesa la planta. Se tomaron muestras de siete especies de invertebrados en dos zonas del río: antes de la planta "Aguas arriba" y después de la planta "Aguas abajo". Los datos se muestran en la siguiente tabla

	Especies						
Zonas	A	B	C	D	E	F	G
Aguas arriba	37	12	10	18	11	16	59
Aguas abajo	19	10	7	20	8	12	24

Se pide:

- a) ¿Se puede admitir que el tipo de especies de vertebrados está relacionado con la situación respecto de la planta de "Aguas arriba del río"?
- b) ¿Se puede admitir relación entre la situación respecto a la planta de la zona del río y el tipo de especies halladas en ella?

Práctica 9. Contraste de hipótesis

Ejercicio Guiado 4 de la práctica 6 (Guía SPSS. Vr. 4.1)

Se quiere estudiar si el 85% de los niños con dolor torácico tienen un ecocardiograma normal. Para ello, se toma una muestra de 139 niños con dolor torácico, de ellos 123 presentan un ecocardiograma normal. ¿Apoyan los datos la hipótesis?

Ejercicio Guiado 5 de la práctica 6 (Guía SPSS. Vr. 4.1)

En un proceso de producción de unas píldoras que se fabrican secuencialmente, la periodicidad de rachas de píldoras defectuosas puede ser significativa de la falta de aleatoriedad en la producción y sugeriría la revisión del proceso. Se desea saber si en el proceso de fabricación de las píldoras, la obtención de éstas en mal estado se produce de manera aleatoria. Para ello se anota el estado de 17 píldoras obtenidas en la cadena de producción a una determinada hora: (B: *Buen estado*; D: *Defectuosa*)

B D B D B B B D D B D B D D B D B

Ejercicio Guiado 6 de la práctica 6 (Guía SPSS. Vr. 4.1)

En un estudio sobre el hábito de fumar y sus efectos sobre las pautas del sueño, una de las variables importantes es el tiempo que se tarda en quedarse dormido. Se extrae una muestra, de tamaño 8, de la población de fumadores y otra independiente, de tamaño 10, de la población de no fumadores. Se obtienen los siguientes datos:

	Tiempos que tardan en dormirse (minutos)									
Fumadores	69.30	56.30	22.10	47.60	53.20	48.10	23.20	23.80		
No Fumadores	18.60	25.10	26.40	14.90	29.80	28.40	12.1	10.20	11.60	12.80

¿Indican estos datos que los fumadores tienden a tardar más tiempo en quedarse dormidos que los no fumadores?

Ejercicio Guiado 7 de la Práctica 6 (Guía SPSS. Vr. 4.1)

En un estudio sobre los efectos del ejercicio físico en pacientes con enfermedad coronaria, se mide el máximo de oxígeno consumido por cada paciente, antes de comenzar el entrenamiento. Después de seis meses de hacer ejercicio con bicicleta tres veces por semana, se midió nuevamente el oxígeno consumido por cada persona y se obtuvieron los siguientes resultados.

	Máximo de oxígeno admitido									
Antes	49.60	23.91	48.35	40.60	43.22	42.12	23.20	30.81	34.70	47.42
Después	40.22	25.15	26.40	58.91	39.80	53.42	51.41	20.23	30.62	31.83

¿Se puede concluir que, al nivel de significación del 5%, el ejercicio tiende a aumentar el máximo de oxígeno admitido por los pacientes?

Práctica 9. Contraste de hipótesis

Ejercicio Guiado 8 de la práctica 6 (Guía SPSS. Vr. 4.1)

Se quiere estudiar si el número de bacterias que aparecen en un determinado cultivo al cabo de una semana es aleatorio o por el contrario habría que suponer que hay algo en el cultivo que propicia el desarrollo de tales bacterias. Para ello, se sometió el cultivo a 10 semanas de observación y se obtuvieron los siguientes resultados: 498, 490, 510, 505, 495, 496, 497, 501, 502, 520.

Ejercicio Guiado 9 de la práctica 6 (Guía SPSS. Vr. 4.1)

En un determinado hospital se están realizando diversos estudios comparativos, con el objetivo de estudiar el número pacientes que llegan, durante una semana al hospital, para ser diagnosticado y el número de enfermos con un tipo de carcinoma que reciben una determinada terapia. Para ello se dispone de la siguiente información:

Sexo	H	H	H	H	H	M	H	M	H	M	H	H	M	M	M	M	H	H	M	M
Terapia	NO	SI	NO	SI	NO	SI	NO	SI	SI	SI	NO	SI	NO	SI	SI	NO	SI	NO	SI	NO

Se pide, a un nivel de confianza del 95%:

- a) Estudiar si el porcentaje de hombres que llegan, durante una semana al hospital, para ser diagnosticado es del 52%.
- b) Comparar la proporción de mujeres con carcinoma que reciben o no reciben la terapia.

Práctica 10. Contraste de hipótesis

Ejercicio Propuesto 1 de la práctica 6 (Guía SPSS. Vr. 4.1)

En una unidad de investigación hospitalaria se está realizando un estudio para conocer si la tolerancia a la glucosa en sujetos sanos tiende a decrecer con la edad. Para ello se realizó un test oral de glucosa a dos muestras de pacientes sanos, unos jóvenes y otros adultos. El test consistió en medir el nivel de glucosa en sangre en el momento de la ingestión (nivel basal) de 100 grs. de glucosa y a los 60 minutos de la toma. Los resultados fueron los siguientes

Jóvenes

Basal	90	82	80	75	74	97	76	89	83	77
60min	136	151	148	138	141	157	154	156	147	141

Adultos

Basal	94	96	93	88	79	90	86	89	81	90
60min	198	191	190	185	184	159	170	197	183	178

Responder a las siguientes cuestiones:

- a. ¿Los niveles de glucosa en sangre en el momento de la ingestión (nivel basal) siguen una distribución normal en las dos poblaciones?
- b. ¿Se puede admitir que el nivel medio de glucosa en sangre en el momento de la ingestión en los jóvenes es menor que 85?
- c. ¿Se detecta una variación significativa del nivel de glucosa en sangre en cada grupo?
- d. Estudiar donde es mayor la concentración de glucosa en sangre:
 - d1.** ¿La concentración de glucosa es mayor en adultos que en jóvenes?
 - d2.** ¿La concentración de glucosa es mayor a los 60 minutos en adultos que en jóvenes?
 - d3.** ¿La concentración de glucosa es mayor en el momento de la ingestión en adultos que en jóvenes?
- e. A un nivel de significación del 5%,
 - e1.** ¿Es representativo el ajuste lineal, en los jóvenes, entre el nivel de glucosa en sangre en el momento de la ingestión (nivel basal) y a los 60 minutos? ¿Cuál sería la expresión del modelo? ¿Cuánto explica el modelo?
 - e2.** ¿Es representativo el ajuste lineal, en los adultos, entre el nivel de glucosa en sangre en el momento de la ingestión (nivel basal) y a los 60 minutos? ¿Cuál sería la expresión del modelo? ¿Cuánto explica el modelo?
 - e3.** ¿Es representativo el ajuste lineal entre los jóvenes y los adultos?

Práctica 10. Contraste de hipótesis

Ejercicio Propuesto 2 de la práctica 6 (Guía SPSS. Vr. 4.1)

Continuando con el estudio de la asimetría en la hoja de la encina, (Ejercicio guiado 1) se ha detectado la presencia de *Agallas en Encina* (pequeñas agallas en el envés de una hoja de encina causadas por el díptero *Dryomyia lichtensteini*) tanto en las zonas secas como en las zonas con mayor precipitación.

- a. En las dos zonas muy secas (Ladikhonda y Fazares) se pretende comprobar si determinado tratamiento, aplicado durante un mes, ayuda a reducir la presencia de dichas agallas. Para ello, se realiza un estudio a 10 encinas, en las que se selecciona aleatoriamente 10 hojas y se registra el promedio de agallas presentes antes del tratamiento y después del tratamiento (se supone normalidad). Los resultados se muestran a continuación:

Antes	10,5	9,7	13,3	7,5	12,8	15,2	11,2	10,7	5,2	18,9
Después	11,2	7,8	9,2	3,4	8,9	10,8	11,4	8,5	6,2	11,1

- b. Se quiere estudiar la asociación entre el nivel de dióxido de sulfúrico del aire y el número medio de *Agallas en Encina* en las zonas de los árboles de Molinillo. Se elige una muestra de 10 zonas de las que se sabe que tienen una alta concentración de dióxido de sulfúrico, 10 zonas que se sabe que tienen un nivel normal y 10 zonas que tienen una baja concentración. Dentro de cada zona se seleccionan aleatoriamente 20 encinas y se determina para cada encina el promedio de agallas en las hojas. Sobre esta base se clasifica cada encina según tenga un recuento bajo, normal o alto de agallas. Se obtienen los datos que se muestran en la siguiente tabla

	Número medio de agallas		
SO ₂	Bajo	Normal	Alto
Bajo	3	10	7
Normal	5	9	6
Alto	8	6	6

Ejercicio Propuesto 3 de la práctica 6 (Guía SPSS. Vr. 4.1)

Se quiere estudiar si el número de bacterias que aparecen en un determinado cultivo al cabo de una semana es aleatorio o por el contrario habría que suponer que hay algo en el cultivo que propicia el desarrollo de tales bacterias. Para ello, se sometió el cultivo a 10 semanas de observación y se obtuvieron los siguientes resultados: 498, 490, 510, 505, 495, 496, 497, 501, 502, 520.

Práctica 10. Contraste de hipótesis

Ejercicio Propuesto 4 de la práctica 6 (Guía SPSS. Vr. 4.1)

En un determinado hospital se están realizando diversos estudios comparativos, con el objetivo de estudiar el número pacientes que llegan, durante una semana al hospital, para ser diagnosticado y el número de enfermos con un tipo de carcinoma que reciben una determinada terapia. Para ello se dispone de la siguiente información:

Sexo	H	H	H	H	H	M	H	M	H	M	H	H	M	M	M	M	H	H	M	M
Terapia	NO	SI	NO	SI	NO	SI	NO	SI	SI	SI	NO	SI	NO	SI	SI	NO	SI	NO	SI	NO

Se pide, a un nivel de confianza del 5%:

- a. Estudiar si el porcentaje de hombres que llegan, durante una semana al hospital, para ser diagnosticado es del 52%.
- b. Comparar la proporción de mujeres con carcinoma que reciben o no reciben la terapia.

Práctica 11. Introducción al Diseño Estadístico de Experimentos.

Ejercicio Guiado 1 de la práctica 7 (Guía SPSS. Vr. 4.1)

Se realiza un estudio del contenido de azufre en cinco yacimientos de carbón. Se toman muestras aleatoriamente de cada uno de los yacimientos y se analizan. Los datos del porcentaje de azufre por muestra se indican en la tabla adjunta.

Yacimientos	Porcentaje de azufre								
	151	192	108	204	214	176	117		
1	169	64	90	141	101	128	159	156	
2	169	64	90	141	101	128	159	156	
3	122	132	139	133	154	104	225	149	130
4	75	126	69	62	90	120	32	73	
5	80	90	124	82	72	57	118	54	130

Para un nivel de significación del 5%.

1. ¿Se puede confirmar que el porcentaje de azufre es el mismo en los cinco yacimientos?
2. Si se rechaza la hipótesis nula que las medias de porcentaje de azufre en los cinco yacimientos es la misma, determinar que medias difieren entre sí utilizando el método de comparaciones múltiples de Tukey.
3. Estudiar las hipótesis de modelo: Homocedasticidad (Homogeneidad de las varianzas por grupo), Independencia y Normalidad.

Ejercicio Propuesto 1 de la práctica 7 (Guía SPSS. Vr. 4.1)

La convección es una forma de transferencia de calor por los fluidos debido a sus variaciones de densidad por la temperatura; las partes calientes ascienden y las frías descienden formando las corrientes de convección que hacen uniforme la temperatura del fluido. Se ha realizado un experimento para determinar las modificaciones de la densidad de fluido al elevar la temperatura en una determinada zona. Los resultados obtenidos han sido los siguientes:

Temperatura	Densidad				
100	21.8	21.9	21.7	21.6	21.7
125	21.7	21.4	21.5	21.4	
150	21.9	21.8	21.8	21.6	21.5
175	21.9	22.1	21.85	21.9	

Responder a las siguientes cuestiones:

1. ¿Afecta la temperatura a la densidad del fluido?
2. Determinar qué temperaturas producen modificaciones significativas en la densidad media del fluido.
3. Estudiar las hipótesis del modelo: Homocedasticidad, independencia y normalidad.
4. Se puede afirmar que las temperaturas de 100 y 125 producen menos densidades de fluido en promedio que las temperaturas de 150 y 175.

Práctica 12. Seminario 2. Análisis estadístico de datos reales: **Análisis inferencial.**

Considerando los datos reales del fichero de las encinas, suponiendo un comportamiento normal, responde razonadamente a las siguientes cuestiones:

1. Estima mediante un intervalo de confianza la anchura media de las hojas de las encinas al 90% de confianza.
2. Estima mediante un intervalo de confianza la diferencia de anchura media entre la parte izquierda y derecha de las hojas de las encinas al 90% de confianza.
3. ¿Pueden considerarse las anchuras de la parte izquierda y derecha iguales al nivel de significación del 5%?
4. ¿Pueden considerarse las anchuras de las hojas de la parte Canopy y Sprouts iguales al nivel de significación del 5%?
1. ¿Existen diferencias significativas de longitud según la zona donde se mida al 1%?
2. ¿Existen diferencias significativas de longitud según el año en el que se mida al 5%?
3. ¿Existen diferencias significativas de longitud según la hoja al 10%?
4. Responde a las cuestiones anteriores sobre la anchura.