



## **CUADERNO DE PRÁCTICAS**

**PRÁCTICAS DE LA ASIGNATURA:**

**BIOESTADÍSTICA**

**Curso 2021/22**

**2º Cuatrimestre**

## **Prácticas de Bioestadística mediante ordenador**

*Práctica 1. Introducción al paquete estadístico SPSS. (1h)*

*Práctica 2. Estadística descriptiva I: Representaciones numérica y gráfica. Resúmenes numéricos. (2h)*

*Práctica 3. Análisis estadístico descriptivo aplicado a unos datos reales: DATOS-ECOLOGIA.xls (1h)*

*Práctica 4. Estadística descriptiva II: Regresión y Correlación. (2h)*

*Práctica 5. Distribuciones de probabilidad. (2h)*

*Práctica 6. Intervalos de confianza. (2h)*

*Práctica 7. Contrastes de hipótesis I. (2h)*

*Práctica 8. Contrastes de hipótesis II. Aplicación a datos reales. (3h)*

## Práctica 1. Introducción al paquete estadístico SPSS.

### Ejercicio 1

Los siguientes datos proceden de un grupo de estudiantes

<b>Nombre</b>	<b>Sexo</b>	<b>Edad</b>	<b>Estatura</b>	<b>Peso</b>	<b>ColorPelo</b>	<b>FechaNac</b>	<b>LugarNac</b>
Felipe	Hombre	18	1,72	75.4	Moreno	11/23/1986	Granada
Pedro	Hombre	20	1,91	82.3	Rubio	3/11/1985	Granada
Ana	Mujer	18	1,67	55.5	Pelirrojo	4/1/1987	Sevilla
Ismael	Hombre	17	1,82	73.0	Castaño	9/9/1987	Jaén
Eloy	Hombre	19	1,77	75.6	Rubio	4/22/1986	Granada
Julia	Mujer	19	1,81	63.1	Rubio	9/11/1985	Jaén
Eva	Mujer	19	1,73	65.2	Moreno	2/28/1986	Málaga
José	Hombre	18	1,84	79.0	Castaño	1/1/1987	Granada
Patricia	Mujer	18	1,87	68.4	Pelirrojo	9/27/1988	Granada
Miriam	Mujer	21	1,60	54.5	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*.
- Crear la variable **IMC** (Índice de Masa Corporal) definida, a partir de las variables **Peso** (en kg) y **Estatura** (en metros), como el cociente  $\text{Peso} / (\text{Estatura}^2)$ .
- Ordenar de forma decreciente los datos según la **Estatura** de los individuos.

f) Seleccionar en la variable ***Estatura*** aquellos casos que tengan un valor menor de 1,77.

## Ejercicio 2

Las siguientes dos tablas estadísticas corresponden a la información disponible relativa a dos grupos de opositores

<b><i>Sexo</i></b>	<b><i>Edad</i></b>	<b><i>Estatura</i></b>	<b><i>Puntuación1</i></b>	<b><i>Puntuación2</i></b>	<b><i>Calificación</i></b>
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

**Datos1**

<b>Sexo</b>	<b>Edad</b>	<b>Estatura</b>	<b>Puntuación1</b>	<b>Puntuación2</b>	<b>Calificación</b>
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

**Datos2**

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*.
- Presentar las variables con sus respectivas *Etiquetas de valor* en el *Editor de datos SPSS*.

- d) En el fichero de Datos1, introducir una nueva variable con el nombre **Media** cuyos valores sean la media de las **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:
- g.1) estén aprobados.
  - g.2) tengan una Puntuación1 superior a 6 y una Puntuación2 inferior a 5.
  - g.3) tengan una Puntuación1 superior a 6 o una Puntuación2 inferior a 5.
  - g.4) estando aprobados sean hombres.

## Práctica 2. Estadística descriptiva I.

### Ejercicio 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 datos relativos al *número de colonias por 100 ml 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

A fin de realizar un estudio comparativo de la calidad del agua en ambas zonas a través del uso de distintas medidas estadísticas y gráficos, se pide:

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 (sugerencia: elegir la opción del gráfico “resúmenes para distintas variables”).

### Ejercicio 2

Un agrónomo mide el *contenido de humedad* en una variedad de trigo tras ser secados especialmente. Para ello, hace las mediciones en 57 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) Tabla estadística de frecuencias, medidas resumen de tendencia central, percentiles 60 y 90, medidas de dispersión.
- b) Gráfico de tallo y hojas.
- c) Gráfico de cajas.
- d) Estratificar la población en tres grupos según *contenido de humedad* atendiendo a la siguiente estratificación: Grupo 1 (valores  $\leq 6.9$ ), Grupo 2 (valores en  $(6.9, 7.4]$ ), Grupo 3 (valores  $> 7.4$ ).

- e) Determinar la media, mediana y desviación típica del *contenido de humedad* en cada grupo.
- f) Construir los gráficos de barras y sectores para la variable de agrupación.

### Ejercicio 3

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

Determinar:

- Tabla de frecuencias.
- Percentiles 35 y 65. ¿Qué porcentaje de muestras se encuentran entre ambos percentiles?
- Valores máximo y mínimo.
- Media, mediana y moda.
- Varianza y desviación típica.
- Histograma de frecuencias.



### Práctica 3. Análisis estadístico descriptivo aplicado a datos reales

El Departamento de Ecología de la Universidad de Granada nos ha cedido la información relativa a un estudio sobre la asimetría de hojas de encina realizado sobre un total de 2101 casos. En dicha investigación, a lo largo de tres años consecutivos el primero muy seco (1996) y los otros dos (1997,1998) muy lluviosos, se tomaron una serie de mediciones sobre hojas de encina procedentes de cinco zonas de la provincia de Granada (dos de ellas, Hoya Guadix-Baza y Ladihonda-Fazares muy secas y las otras tres Cortijuela, Molinillo y Fardes con mayor humedad) registrándose también cuál era la situación de la hoja en el árbol: Canopy (copa de los árboles), Sprouts (rebrotos, hojas nuevas que salen desde la parte inferior del tronco).

El objetivo de tal investigación era 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.

El archivo DATOS-ECOLOGÍA.xls (disponible en la plataforma PRADO) contiene una muestra de 30 casos de los 2101 que fueron objeto de estudio. En base a esta muestra, 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. Crea las variables

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

determinando sus valores 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, moda, mediana, varianza, máximo, mínimo, asimetría y curtosis de la longitud de la hoja.
8. Obtén las mismas medidas del apartado anterior pero diferenciando según la zona. Construye también un gráfico de cajas.

## Práctica 4. Estadística descriptiva II: Regresión y Correlación.

### Ejercicio 1

Se realiza un estudio para establecer una ecuación mediante la cual pueda utilizarse la *concentración de estrona en saliva (x)* para predecir la *concentración de estrona en plasma libre (y)*. Se obtuvieron los siguientes datos correspondientes a 14 individuos varones sanos:

<i>x</i>	<i>y</i>
7,4	30,0
7,5	25,0
8,5	31,5
9,0	27,5
9,0	39,5
11,0	38,0
13,0	43,0
14,0	49,0
14,5	55,0
16,0	48,5
17,0	51,0
18,0	64,5
20,0	63,0
23,0	68,0

Se pide:

- Diagrama de dispersión.
- Ecuación de la recta de regresión que exprese la concentración de estrona en plasma libre en función de la concentración de estrona en saliva. Estudia la bondad del ajuste realizado.
- Manteniendo las mismas variables dependiente e independiente del apartado anterior, obtener la ecuación de la parábola de regresión. Estudia la bondad de este segundo ajuste.

### Ejercicio 2

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:

- ¿Se observa en los datos una tendencia lineal?
- Determina la recta de regresión que predice la tasa de mortalidad en función del nivel de humedad. Cuantifica la bondad del ajuste realizado. ¿Cuánto explica el modelo ajustado?
- Predice el nivel de humedad del suelo si la tasa de mortalidad de las lombrices es 0,7.
- Determina el coeficiente de correlación lineal entre la tasa de mortalidad y el nivel de humedad.
- Ajusta los datos mediante una regresión curvilínea
- ¿Qué ajuste ha resultado ser más adecuado para los datos disponibles, el lineal o el curvilíneo?

## Práctica 5. Distribuciones de probabilidad.

### Ejercicio 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. Se sabe que la probabilidad de que un paciente vacunado contraiga la gripe es de 0.3.

Calcula las siguientes probabilidades:

- a) Ningún paciente contraiga la gripe.
- b) Más de dos pacientes contraigan la gripe.
- c) Contraigan la gripe entre tres y cinco pacientes, ambos inclusive.
- d) Genera una muestra aleatoria de tamaño 20 de valores de una distribución Binomial de parámetros  $n = 10$  y  $p = 0,2$ .

### Ejercicio 2

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

Calcula la probabilidad de que:

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

### Ejercicio 3

Se ha estudiado el nivel de glucosa en sangre en ayunas en un grupo de diabéticos. Supuesto que esta variable sigue un comportamiento normal de media 106 mg/100 ml y desviación típica 8 mg/100 ml.

Se pide:

- a) Obtener la probabilidad de que el nivel de glucosa en sangre en un diabético sea inferior a 120 mg/100 ml.

- b) ¿Qué porcentaje de diabéticos tienen niveles de glucosa en sangre comprendidos entre 90 y 130 mg/100 ml?
- c) 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.
- d) Generar una muestra de tamaño 12 para la una distribución Normal con media igual a 5 y desviación típica igual a 3.

#### Ejercicio 4

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 5

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 que oscile entre 0.20ppm 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.

#### Ejercicio 6

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

## Práctica 6. Intervalos de confianza.

### Ejercicio 1

Se está interesado en comprobar el efecto de dos fertilizantes (A y B) sobre la producción de unos árboles frutales. Para ello, se seleccionan aleatoriamente dos grupos de 8 y 10 árboles cada uno y se le añade al agua de riego de cada uno de ellos el fertilizante A y B, respectivamente. La *producción (en kg)* de ese año fue la siguiente:

<b>Fertilizante A</b>	30	25	28	29	30	31	24	22	25	27
<b>Fertilizante B</b>	28	27	28	28	26	27	26	29		

Se pide:

- Obtener sendos intervalos 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.
- 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.
- 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.

<b>Antes</b>	25	20	25	28	30	30	26	15	18	22
<b>Después</b>	30	25	28	29	30	31	24	22	25	27

### Ejercicio 2

Se quiere estudiar si la *longitud del pico en una especie de loro* es distinta en machos y hembras. A fin de realizar dicho estudio se selecciona una muestra de 14 machos y otra de 12 hembras y las mediciones de sus picos (en mm) fueron las que se recogen en la tabla siguiente:

<b>Macho</b>	57	58	60	58	61	62	61	59	57	63	58	55	59	60
<b>Hembra</b>	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.

### Ejercicio 3

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 kg) de la lana producida* antes y después del experimento

<b>Antes</b>	10	8	7	5	9	12	10	9	8	8
<b>Después</b>	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.

### Ejercicio 4

Se desea analizar el *nivel de glucemia*, según tratamientos, en un grupo de enfermos. Para ello, se han llevado a cabo sendos estudios comparativos de los niveles antes y después de cada tratamiento.

De un lado, fue tomada una muestra de 32 pacientes de este grupo a la que se trató con Tibolona (tratamiento A) obteniéndose los siguientes valores relativos a sus niveles de glucemia:

**Antes del tratamiento:** 88, 77, 87, 89, 93, 78, 95, 118, 82, 92, 110, 90, 97, 84, 85, 79, 84, 92, 83, 91, 85, 105, 89, 85, 88, 89, 101, 92, 89, 91, 86, 92.

**Después del tratamiento:** 104, 83, 86, 89, 99, 73, 96, 129, 88, 80, 95, 79, 77, 100, 87, 98, 131, 82, 79, 94, 100, 102, 95, 90, 91, 92, 105, 88, 95, 91, 82, 84.

A una segunda muestra de otros 28 pacientes del grupo se le aplicó Ciclo-Secuencial (tratamiento B), siendo los resultados:

**Antes:** 88, 103, 90, 93, 93, 90, 112, 85, 109, 87, 112, 106, 96, 103, 87, 81, 89, 95, 84, 84, 89, 91, 91, 96, 86, 78, 89, 107

**Después:** 80, 94, 80, 89, 87, 89, 101, 56, 96, 79, 92, 101, 87, 96, 93, 87, 80, 88, 89, 88, 85, 85, 85, 83, 92, 75, 85, 97

Bajo el supuesto de normalidad de las variables, hallar un intervalo de confianza para la diferencia de medias entre los niveles de glucemia antes y después de cada tratamiento. ¿Son efectivos ambos tratamientos?



## Práctica 7. Contrastes de hipótesis I.

### Ejercicio 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 (en minutos) que un individuo tarda en quedarse dormido. Se extrae una muestra de tamaño 8 de la población de fumadores y otra muestra independiente de tamaño 10 de la población de no fumadores. Se obtienen los siguientes datos:

	<i>Tiempos que tardan en dormirse</i>									
<b>Fumador</b>	69.30	56.30	22.10	47.60	53.20	48.10	23.20	23.80		
<b>No Fumador</b>	18.60	25.10	26.40	14.90	29.80	28.40	12.1	10.20	11.60	12.80

A la vista de los datos, y supuesto normalidad de la variable bajo estudio en ambas poblaciones:

- ¿Puede aceptarse que, en media, el tiempo que tarda en dormirse un no fumador es a lo sumo de 20 minutos? ¿Y en el caso de un fumador?
- ¿Hay evidencia suficiente a favor de que los fumadores tienden a tardar más tiempo en quedarse dormidos que los no fumadores?

### Ejercicio 2

Se realiza una investigación para analizar 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

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

La hipótesis de investigación es que el programa de ejercicios físicos reduce el nivel de triglicéridos en sangre. ¿Existe evidencia significativa que sustente dicha hipótesis? (supóngase normalidad).

### Ejercicio 3

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.

	<i>Máximo de oxígeno admitido</i>										
<b>Antes</b>	49.60	23.91	48.35	40.60	43.22	42.12	23.20	30.81	34.70	47.42	
<b>Después</b>	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? (supóngase normalidad)

## Práctica 8. Contrastes de hipótesis II. Aplicación a datos reales.

### Ejercicio 1

A partir de la información disponible en el fichero DATOS-ECOLOGIA.xls, resuélvanse las siguientes cuestiones:

A un nivel de significación del 5%,

- ¿Se puede asumir que la longitud de las hojas de encina se distribuye normalmente? ¿y la asimetría de las hojas?
- ¿Puede aceptarse que la longitud media de las hojas es igual a 30 cm?
- Examínese si existen diferencias significativas en la asimetría de las hojas teniendo en cuenta la situación de la hoja en el árbol.

### Ejercicio 2

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: *Aguas arriba* (antes de la planta industrial) y *Aguas abajo* (después de la planta industrial). Los datos se muestran en la siguiente tabla

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

Se pide:

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

### Ejercicio 3

Se quiere examinar 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 los cuales

123 presentan un ecocardiograma normal. ¿Puede aceptarse tal suposición? ¿Existe evidencia significativa de que dicho porcentaje sea mayor?

#### Ejercicio 4

En relación al estudio de la asimetría en la hoja de la encina 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 (Ladihonda 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:

<b>Antes</b>	10,5	9,7	13,3	7,5	12,8	15,2	11,2	10,7	5,2	18,9
<b>Después</b>	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		
<b>SO<sub>2</sub></b>	<i>Bajo</i>	<i>Normal</i>	<i>Alto</i>
<i>Bajo</i>	3	10	7
<i>Normal</i>	5	9	6
<i>Alto</i>	8	6	6

## Ejercicio 5

A partir de la información recogida en el archivo DATOS-ECOLOGIA.xls, examínese al 5% de significación si resulta representativo el ajuste lineal entre la longitud y la asimetría. ¿Cuál sería la expresión del modelo? ¿Cuánto explica el modelo?

## Ejercicio 6

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

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

### Adultos

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

Responder a las siguientes cuestiones:

- ¿Los niveles basales de glucosa en sangre siguen una distribución normal en las dos poblaciones? ¿Y los de glucosa a los 60min?
- ¿Puede admitirse que el nivel basal medio de glucosa en sangre en los jóvenes está por debajo de 85?
- Para cada población, ¿se detecta una variación significativa entre los niveles medios de glucosa basal y tras 60min?
- ¿Existe evidencia significativa de que los niveles basales de glucosa son mayores en adultos que en jóvenes?
- Resuélvanse nuevamente los apartados c) y d) bajo el supuesto de no-normalidad de las variables.
- En la población de jóvenes, a un nivel de significación del 5%, ¿es representativo el ajuste lineal entre el nivel basal de glucosa en sangre y a los 60 minutos? ¿Cuál sería la expresión de la recta ajustada? ¿Cuánto explica el modelo?