

**Título do proxecto:**

**Innovación no ensino práctico da Bioestatística  
aplicando novas tecnoloxías da información**

**Coordenadora: M<sup>a</sup> del Carmen Carollo Limeres**

## Descripción dos materiais producidos

# I.- CONTIDOS TEÓRICOS

## INTRODUCCIÓN Ó PAQUETE ESTATÍSTICO SPSS.

### ○ O programa estatístico SPSS

Os programas estatísticos permítennos a introducción, a representación gráfica e o tratamento estatístico de datos.

Un dos máis completos e de sinxelo manexo é o SPSS (Statistical Package for the Social Sciences) fabricado por SPSS Inc. Dado que permite a rápida manipulación de grandes bases de datos e ofrece a posibilidade de aplicarlles gran cantidade de procedementos gráficos e estatísticos, é empregado como ferramenta para a investigación relacionada co manexo de enquisas, sondaxes de opinión, resultados de test e en xeral, co emprego de calquera tipo de datos procedentes do eido das Ciencias Sociais e Ciencias Biolóxicas e da Saúde.

Das diferentes versións existentes deste programa ocuparémonos da versión 12.0 para WINDOWS.

Sempre que se desexe realizar unha análise con SPSS, deberemos ter en conta os seguintes catro puntos básicos:

- Contar cun arquivo de datos que poidan ser lidos polo SPSS.
- Seleccionar o procedemento que queremos aplicar (obtención de táboas, gráficos, medidas resumen, etc.)
- Selecciona-las variables ás que se desexan aplicar ditos procedementos.
- Obter os resultados.

### ○ Entorno de traballo do SPSS

O programa presenta un gran número de fiestras dende as que, por unha banda, se xestiona a introducción de datos e decídense as análises a realizar e, por outra, accédese a distintos aspectos da manipulación dos resultados xerados. Todas elas presentan as súas propias barras de ferramentas que poden ser, como en case todas as aplicacións do entorno Windows, personalizadas a gusto do usuario. A continuación enuméranse e describíense brevemente as características e contidos das distintas fiestras:

#### • *Vista de datos/variables:*

Pode considerarse a fiestra principal do programa pois ao cerra-la remátase coa execución do mesmo. Esta fiestra conta con dúas vistas diferentes: a vista de datos (na que poden observarse os datos a analizar) e a vista de variables (na que poden observarse as variables do ficheiro de datos). É importante subliñar que só pode haber un ficheiro de datos aberto a un tempo.

#### • *Visor de resultados:*

Os resultados xerados polas distintas análises (gráficos, estatísticos,...) amósanse no visor de resultados. Esta fiestra ábrese de maneira automática no momento no que se executa a primeira análise. Se ben poden coexistir varios visores a un tempo, só hai un activo en cada momento que

delas mediante a opción **designar ventana activa**.

- **Editor de gráficos:**

Pode modificalos gráficos cambiando calquera aspecto do mesmo: cores, tamaños, intercambiar eixes, ademais dunha elegante ferramenta de rotación de nubes de puntos 3-D.

- **Editor de resultados de texto:**

Con este editor pódense modificar os resultados tipo texto (sinxelamente identificables pois non están inmersos nunha táboa). É un simple editor para cambiar aspectos básicos do texto.

Outras fiestras que aparecerán durante a sesión de traballo son os cadros de diálogo. Estes aparecen cada vez que se selecciona unha opción do menú ou barra de ferramentas. Neles introdúcese a información que o programa precisa para completar a operación escollida.

Para o caso das análises estatísticos e creación de gráficos, os cadros de diálogo sempre presentan as seguintes características. Por un lado mostran as variables do arquivo de datos dispoñibles para o procedemento seleccionado. Dependendo do tipo de análise, estarán dispoñibles todas ou algunhas variables. Por outro lado preséntanse un ou varios cadros de texto onde introducir as variables que configurarán a análise (p. ex. variable dependente e independente, nunha análise de regresión lineal simple). Ademais, móstranse varios botóns para, executar a análise (**Aceptar**), abandonar o cadro de diálogo (**Cancelar**), borrar tódalas seleccións realizadas (**Restablecer**), trasladar a sintaxe da acción a realizar o editor de sintaxe (**Pegar**), e seleccionar resultados e opcións adicionais da análise (**Opciones, ...**).

- **Fiestra Editor de datos**

Cando entramos no SPSS aparece a fiestra Editor de datos. Dende ela imos poder acceder a un determinado tipo de arquivos de datos, xa creados, e poderemos modificar os datos deses arquivos. Ademais tamén poderemos introducir directamente os datos para un novo arquivo. Os arquivos de datos noméanse coa extensión *.sav*.

En xeral, nun arquivo de datos de SPSS temos  $n$  filas e  $p$  columnas. Cada fila corresponde a un individuo e contén, para ese individuo, os valores de  $p$  variables. É dicir, cada fila corresponde a un individuo, cada columna a unha variable e cada cuadrícula contén un dato (asociado a unha variable e a un individuo). Una cuadrícula valeira interprétase como un dato perdido.

Na fiestra Editor de datos atopámonos cun menú dende o que se pode acceder a moitas das funcións do SPSS. Ten as seguintes opcións:

- **Archivo:** permite crear, abrir, gardar e imprimir arquivos. Mostra os datos que teñen sido utilizados recentemente.
- **Edición:** permite buscar, copiar, cortar, pegar e borrar datos.
- **Ver:** permite seleccionar ou non a visualización de certo tipo de información, como a cuadrícula, barras de ferramentas e estado. Tamén proporciona información sobre as fontes:

tipo, estilo e tamaño da letra.

- *Datos*: permite facer cambios globais nos datos (engadir novas variables, arranxar ou seleccionar casos,...)
- *Transformar*: permite facer cambios en determinadas variables, crear novas variables a partir das xa existentes ou substituír valores perdidos.
- *Analizar*: permite efectuar diferentes procedementos estatísticos: informes, táboas, medidas resumen, gráficos sinxelos, etc.
- *Gráficos*: permite efectuar diferentes tipos de representacións gráficas.
- *Utilidades*: permite obter información sobre as variables, definir conxuntos de variables ou executar determinados procesos de SPSS.
- *Ventana*: da información sobre as fiestras abertas.
- *?*: contén información sobre o programa SPSS.

Debaixo do menú atopamos unha barra de ferramentas cuns botóns dende os que se pode acceder rapidamente a algunhas das funcións máis habituais. O situar o punteiro sobre cada botón pódese obter unha breve descrición do mesmo.

Podemos distinguir tres tipos de ferramentas: as xerais (que aparecen en todas as fiestras de WINDOWS), as específicas de cada fiestra e as ferramentas de gráficos.

As ferramentas xerais permítennos abrir un arquivo, gardalo, imprimilo, mostrar un listado das operacións máis recentes efectuadas con SPSS, mostrar os datos dun individuo concreto e información sobre as variables.

En canto ás ferramentas específicas da fiestra Editor de datos, permítennos entre outras cousas, buscar un valor concreto dunha variable e todos os individuos que presentan dito valor, engadir unha fila para os datos dun novo individuo ou unha columna para unha nova variable, seleccionar os individuos que verifican algunha condición ou elixir entre a visualización dos valores das variables ou das súas etiquetas.

Na parte inferior da fiestra atópase a denominada barra de estado que proporciona información sobre o estado actual do proceso, o número de casos procesados, o filtrado dos datos (cando se utiliza só unha parte do arquivo de datos), a variable que pondera (cando se usa algún criterio de ponderación) ou a segmentación de datos (cando o arquivo de datos se divide en subgrupos a partir dunha ou varias variables de agrupación).

#### ○ **Apertura dun arquivo dende a fiestra Editor de datos**

Farémolo dende a opción *Archivo/ Abrir/ Datos*. Aparecerá entón un cadro de diálogo no que haberá que seleccionar as seguintes cuestións:

- *Nome do arquivo*: aparece unha lista dos arquivos de datos (con extensión *.sav*) da carpeta de SPSS. Pódese seleccionar algún deles ou ben algún outro gardado nesa mesma carpeta ou

noutra diferente.

- *Arquivos de tipo*: permite seleccionar entre os diferentes tipos de arquivos de datos que recoñece o SPSS.
- *Buscar en*: permítenos seleccionar a carpeta na que queremos buscar o arquivo de datos. Podemos recorrer a “árbore”, cara abaixo (abrindo novas carpetas) ou cara arriba (cerrando carpetas e buscando dentro das que xa están abertas).

Hai varios tipos de arquivos de datos que recoñece o SPSS:

- SPSS(\*.sav): son os arquivos de datos que se crean e/ou gardan co SPSS baixo WINDOWS ou UNIX.
- SPSS/PC+(\*.sys): os arquivos que se crean co SPSS/PC+.
- Transportables de SPSS(\*.por): arquivos creados baixo outros sistemas como Macintosh.
- Excel(\*.xls): arquivos creados pola folla de cálculo Excel.
- Lotus(\*.w\*): arquivos creados pola folla de cálculo Lotus.
- SYLK(\*.slk): arquivos creados por Excel ou Multiplan gardados en formato Syk.
- dBase(\*.dbf): arquivos da base de datos dBase.
- Texto(\*.txt): arquivos de texto.
- Datos(\*.dat): arquivos de texto que están delimitados por tabuladores.
- Todos os arquivos(\*.\*): calquera tipo de arquivo.

*Exemplo 1*: Na dirección *C:\Archivos de programa\SPSS* podemos abrir o arquivo *Supervivencia cancer de pecho.sav*.

En *Vista de datos* podemos observar que temos información sobre 11 variables en 1207 individuos.

En *Vista de variables* podemos ver con detalle as características de cada variable, en particular os posibles valores de cada unha delas. En concreto, se para a variable *tumorcat* imos a cela correspondente os *valores*, ó pulsar sobre o cadrado que aparece (⋮), podemos ver os diferentes valores así coma as etiquetas que identifican a cada un deles.

### o Creación dun novo arquivo de datos

Poderémolo facer dende a opción *Archivo/Nuevo/Datos*. Hai que dar os seguintes dous pasos:

#### ➤ Introducción das variables de interese e as súas características

Despois de entrar nun novo arquivo de datos e dende *Vista de variables* poderemos introducir para as variables, das que imos a ter datos, as seguintes características:

- *Nombre*: Por defecto, SPSS asigna a unha variable o prefixo “var” seguido de cinco díxitos. Cando introducimos nos o nome, debemos ter en conta os seguintes requisitos:
  - Terá un máximo de oito caracteres.
  - Empezará por unha letra e non acabará cun punto.
  - Non pode conter espazos en branco nin caracteres especiais (¡,?,\*,...)
  - Non pode haber dous nomes de variables repetidos.
  - O programa non distingue entre maiúsculas e minúsculas.
  - Non poden usarse as palabras ALL, LT, AND, EN, BY, NOT, EQ, OR, GE, TO, GT, UIT, LE.
- *Tipo*: As variables poden ser:

- Numéricas: estas admiten valores numéricos, signos de “+”, “-“ e decimais. O ancho máximo é de 40 caracteres e o número de decimais é 16.
  - Coma: igual co tipo numérico pero incluíndo unha coma de delimitación para os valores cada tres posicións. Os decimais neste caso sepáranse por un punto.
  - Punto: igual ca no caso anterior, pero cambiando comas por puntos.
  - Cadena: Variables tales que os seus valores non son numéricos.
  - Outros: Notación científica, fecha e moeda.
- *Valores*: Podemos asignar unha etiqueta a cada valor da variable cando os valores son numéricos pero en realidade representan modalidades dunha variable cualitativa. É importante que cada vez que na correspondente fiestra de diálogo se introduce un *valor* e a súa *etiqueta*, debemos pulsar o botón de añadir.
  - *Perdidos*: Aparte de calquera cuadrícula en branco dentro dos datos (que se considerará un dato perdido ou “missing”) podemos indicar qué valores se poden considerar como perdidos: tres valores como máximo nas variables discretas, un intervalo nas variables continuas ou un intervalo e un valor fora do mesmo para ambos tipos de variables.
  - *Otros*: Ademais pódese especificar a *anchura* de caracteres que ocupará a variable, o número de *decimais* que usaremos, por una *etiqueta* á variable, indicar o número de *columnas* que se reservan, la *aliñación* da información dentro da cuadrícula (dereita, centro ou esquerda) e o tipo de *medida* que corresponde á variable (escala, ordinal ou nominal).

#### ➤ **Introducción dos datos**

Unha vez que creou a “plantilla” con todas as variables, se imos a *Vista de datos*, por cada variable creada vemos que aparece unha columna co nome de dita variable. Agora, en cada fila, débense de introducir os datos (correspondentes a todas as variables) dun individuo.

#### ➤ **Transformación e edición dos datos**

Entre as operacións máis usuais de *edición de datos* atópanse:

- Modificar datos: basta situarse na cuadrícula correspondente e introducir o cambio, que pode consistir nunha modificación parcial do dato ou na súa substitución por un novo.
- Copiar, cortar e pegar datos: pódese seleccionar unha cuadrícula para a continuación cortala ou copiala e despois pegala (utilizando a opción *Edición* do menú) na cuadrícula o cuadrículas que previamente seleccionemos. O mesmo pódese facer cunha fila, cunha columna ou cun área completa do arquivo.
- Engadir unha nova variable: a opción *Datos/ Insertar variable* inserta unha nova variable á esquerda de onde se atope situado o punteiro.
- Engadir un novo individuo: a opción *Datos/ Insertar caso* inserta un novo individuo na fila superior a onde se atope situado o punteiro.
- Eliminar variables ou individuos: si se desexa suprimir unha variable basta marcala e pulsar o botón *Suprimir* do teclado do ordenador (ou a opción *Edición/ Eliminar*); da mesma forma procederase cando se desexe suprimir un individuo.
- Encontrar un individuo: utilízase a opción *Datos/ Ir a caso*.
- Encontrar un valor dunha variable: con *Edición/ Buscar*.
- Ordenar os datos: con *Datos/ Ordenar casos*.

Os seguintes exemplos amósannos algunhas das transformacións que se poden realizar cos datos. As operacións máis habituais de *transformación de datos* son as seguintes:

- Crear novas variables a partir das que xa están no arquivo de datos.

- Recodificar variables.
- Categorizar variables.
- Reemplazar valores perdidos.

Exemplo 2: Cos datos do exercicio 2 crear una nova variable que sexa o peso en Kg.

Faremos *Transformar/ Calcular/ Variable de destino* pesokg /*Tipo y etiqueta* numérico e peso en Kg / *Expresión numérica* peso/1000.

Exemplo 3: Cos datos do exercicio 2 crear unha nova variable que teña valor un se o elemento da mostra é un alga e 2 para o resto. Faremos uso da opción Transformar/ Calcular:

*Transformar/ Calcular/ Variable de destino* ntipo / *Expresión numérica:* 1 *Si/ Incluir si el caso satisface la condición:* (tipo=1) e despois

*Transformar/ Calcular/ Variable de destino* ntipo / *Expresión numérica:* 2 *Si/ Incluir si el caso satisface la condición:* (tipo=2) ~=(tipo=3).

Exemplo 4: O anterior exemplo tamén se pode resolver recodificando a variable tipo:

*Transformar/ Recodificar/ En distintas variables/ Variable numérica* tipo / *variable de resultado* ntipo2 / *cambiar/ Valores antiguos y nuevos/ Valor antiguo* 1 / *Valor nuevo* 1 / *Añadir Todos los demás valores* Valor nuevo 2 / *Añadir.*

Si se quere facer unha análise por separado en grupos de individuos, necesitamos dividir o arquivo de datos en tantas partes como grupos.

Exemplo 5: Cos datos do exercicio 2 separar o arquivo de datos en tres: un coas algas, outro cos moluscos, e outro cos peixes.

*Datos/ Segmentar archivo/ Organizar los resultados por grupos/ Grupos basados en tipo.* Para desactivar esta agrupación basta seleccionar *Analizar todos los casos, no crear los grupos.*

O facer certas análises vains interesar en ocasións seleccionar algúns individuos concretos no estudio.

Exemplo 6: Cos datos do exemplo 1: A) seleccionar de entre toda a mostra aproximadamente o 75% de los individuos aleatoriamente. B) Coller de novo tódolos datos. C) Seleccionar agora os que estean vivos.

- A) *Datos/ Seleccionar casos/ Muestra aleatoria de casos/ Muestra/ Aproximadamente/ 75% de todos los casos/ Los casos no seleccionados son filtrados.*
- B) *Datos/ Seleccionar casos/ Todos los casos.*
- C) *Datos/ Seleccionar casos/ Si satisface la condición/ Si estado=0.*

#### o **Remate e novo inicio dunha sesión**

Se queremos gardar o arquivo basta utilizar a opción *Archivo/ Guardar como* e indicar a carpeta e o nome para o arquivo. Cando o arquivo xa ten un nome e se queren gardar novos datos, basta utilizar *Archivo/ Guardar*. Para saír do programa utilizaremos *Archivo/ Salir*. Para volver a traballar cun documento xa creado, despois de entrar no programa utilizaremos *Archivo/ Abrir/*

*Datos* e indicaremos a carpeta e o nome do arquivo.

## TÁBOAS DE CONTINXENCIA E REGRESIÓN LINEAL SIMPLE.

## o Táboas de continxencia

Por **táboas de continxencia** enténdense aquelas táboas de dobre entrada onde se realiza unha clasificación da mostra dacordo a un dobre criterio de clasificación.

Por exemplo, a clasificación de uns individuos dacordo ó seu sexo e ó seu grupo sanguíneo crearía unha táboa onde cada cuadrícula da táboa representaría a **frecuencia bivariante** das características correspondentes á súa fila e á súa columna (p. ex. mulleres de grupo sanguíneo A).

Con estas táboas pódese obter unha descrición cuantitativa das distintas calidades bivariantes da mostra, en forma de frecuencias e porcentaxes. Estes últimos poden ser relativos ó total da mostra, ó total dunha fila ou ó total dunha columna. Ademais do anterior, o SPSS realiza diversos contrastes acerca da distribución das frecuencias observadas en dita táboa, de acordo a distintas hipóteses. O máis clásico destes contrastes é o **contraste de homoxeneidade, de independencia ou de igualdade de proporcións** que propón, como hipótese a rexeitar, que tódalas probabilidades son iguais. É dicir, a pertenza dun individuo a unha clase dunha das variables de clasificación non afecta á probabilidade de pertenza ás distintas clases do outro criterio. En termos dun ensaio clínico, por exemplo, esta homoxeneidade suporía que o feito de que un individuo fose destinado aleatoriamente a un dos brazos do ensaio, non afecta ó resultado clínico que terá, medido este en termos cualitativos (p. ex. curado ou non curado).

Para a realización deste tipo de proba estatística, débese seleccionar do menú **Analizar**, a opción **Estatísticos Descritivos** e dentro dela a última das opcións (**Táboas de continxencia**). No cadro de diálogo correspondente debemos definir qué variables **categorías** definirán as filas e as columnas da táboa. Obsérvase que no mesmo cadro pódense definir, mediante outras variables cualitativas, novas dimensións a esta táboa, é dicir, estratifica-la mostra de acordo a criterios adicionais.

Como configuración adicional da táboa de continxencia podemos seleccionar os estatísticos que desexamos obter, así como a información que desexamos engadir en cada unha das cuadrículas de dita táboa. Entre os estatísticos, ademais do **Chi-cuadrado** para o contraste de independencia, atópase a opción de calcular unha das medidas de asociación (**risco**) máis frecuente como é a **Odds Ratio (OR)** que aparece nos resultados baixo a denominación de razón de vantaxes. Estas medidas de risco só se calcularán no caso particular de táboas 2x2.

## o Regresión lineal simple.

Mediante as técnicas de regresión preténdese modelar a relación existente entre unha **variable dependente** e unha ou varias **variables independentes**. No caso da **regresión lineal simple**, o modelo xerado asume a existencia dunha relación lineal entre unha variable dependente e unha única variable independente. Este será o caso do que nos ocuparemos.

Baixo a opción de **Regresión**, do menú **Analizar**, englobanse todas as técnicas de modelado de entre as que seleccionaremos a opción Lineal. Esta opción agrupa tanto a regresión Simple como a Múltiple. A diferenza entre ambas atópase no número de variables independentes que se introducen no cadro de diálogo.

Unha vez seleccionadas a variable dependente e a independente, podemos definir un filtrado

dos casos. Isto provocará que a análise se realizará sobre o grupo seleccionado. Igualmente, podemos decidir se desexamos que nos gráficos apareza unha identificación dos puntos, o que se

realizará de acordo á variable introducida como **etiquetas de casos**.

Na fiestra de estatísticos podemos solicitar ó programa que mostre distinta información nos resultados numéricos da análise. De eles cabe destacar dous: a opción de **estimacións** e a relativa ós **intervalos de confianza**. Estas opcións refírense ós coeficientes da recta de regresión, é dicir, ó modelo de regresión que é o que estamos buscando.

Na fiestra de resultados móstrase a táboa resultante con estes coeficientes e co resultado do contraste da hipóteses de que estes coeficientes sexan cero.

Para a aplicación desta análise débese asumir a **normalidade dos residuos**. Estes calcúlanse como a diferenza, para cada valor da variable independente, entre o valor da variable dependente e o valor predito polo modelo lineal.

## II.- PRÁCTICAS NO LABORATORIO DE INFORMÁTICA

### PRÁCTICA 1:

Vaise realizar un estudo sobre un residuo nunha ría. Tómase unha mostra de algas, moluscos e peixes nunha determinada zona. De cada elemento da mostra anotarase un identificador (*id*), o tipo (Alga, Pez o Molusco) (*tipo*), tamaño (Pequeno, Medio, Grande) (*tam*), peso en gramos (*peso*) e cantidade de residuo en gramos (*residuo*).

**Exercicio 1:** Crear un arquivo de datos con estas variables e as súas características. Gardalo arquivo co nome *Practical.sav*.

Despois de efectuarse a mostraxe obtéñense os seguintes datos:

id	tipo	tam	peso	residuo
1	1	1	0,30	0,0100
2	1	1	0,09	0,0030
3	3	3	2300,00	1,0000
4	3	2	500,00	0,0010
5	3	3	1700,00	0,5000
6	2	3	30,00	2,0000
7	1	2	10,00	0,0100
8	3	1	200,00	0,0010
9	2	2	10,00	2,0000
10	2	3	25,00	2,0000
11	3	1	150,00	0,0008
12	1	2	1,00	0,0100

**Exercicio 2:** Introdúzanse os datos anteriores no arquivo de datos. Gardalo arquivo.

**Exercicio 3:** Cos datos dos exercicios 1 e 2 pídese:

- Transformar a variable residuo nunha nova variable que sexa “a cantidade de residuo en miligramos”.
- Crear unha nova variable que teña valor 1 para os elementos da mostra que teñan tamaño pequeno e medio e 2 para o resto, a través da opción “transformar/calcular” da variable tamaño e tamén a través dunha “recodificación” da mesma.

- Separar el arquivo de datos en tres: segundo o tipo (alga, molusco ou peixe).
- A) Seleccionar de entre toda a mostra aproximadamente un 90% dos individuos aleatoriamente.  
B) Volver a coller tódolos datos. C) Seleccionar os peixes da mostra.
- Ordenar os individuos de maior a menor peso.

## PRÁCTICA 2:

### Exemplo de táboas de continxencia:

Na dirección *C:\Archivos de programa\SPSS*, abrir o arquivo de datos *Datos de Empleados.sav*. Estudiar se existe asociación entre as variables *SEXO* e *CATEGORÍA LABORAL*.

- ❖ Imos a *Analizar / Estadísticos Descriptivos / Tablas de Contingencia*.
- ❖ Seleccionamos a variable *SEXO* para as filas e a variable *CATLAB (CATEGORIA LABORAL)* para as columnas.
- ❖ No botón *Estadísticos* seleccionamos *chi-cuadrado* e facemos clic en *Continuar*.
- ❖ No botón *Casillas* seleccionamos *Frecuencias Observadas, Porcentajes Filas e Columnas*. Facemos clic en *Continuar*.

Facemos clic en *Aceptar* e obtemos os resultados.

**Exercicio 1:** Cos datos de exemplo anterior estudar se existe asociación entre as variables *MINORÍA* e *CATEGORÍA LABORAL*.

### Exemplo de regresión lineal:

Un investigador está interesado no estudio dun tipo de mel que se produce na provincia de Lugo. En 1990 recolléronse 66 mostras de mel de dita provincia clasificándoa dacordo ás seguintes variables: *Zona*, *Humedad*, *Cenizas*, *Solins (sólidos insolubles)*, *Azured (azucres reductores)*, *Sacar (sacarosa)*, *Conduc (conductividade)*, *AcTot (acidez total)*, *AcLib (acidez libre)*, *AcLab (acidez lactónica)*, *HMF (Hidroximetilfusfural, importante para clasificar as meles)*, *pH*, *Lac\_Lib (AcLac/AcLib)* e *Azu\_Hum (Azured/Humedad)*. Os datos atópanse gardados no arquivo de datos *miel.sav*.

O investigador trata de establecer un modelo que relacione as variables *AcTot* e *AcLib* para poder predicir os valores de *AcTot* en caso de coñecer *AcLib*. Os pasos para estudar este modelo de regresión son os seguintes:

1. APROXIMACIÓN O MODELO: Imos a *Gráficos / Dispersión / Simple*. Seleccionamos a variable **independente** *AcLib* no eixo X, e a variable **dependente** *AcTot* no eixo Y. Facemos clic no botón *Aceptar*, e observámoslos resultados.
2. MODELO DE REGRESIÓN LINEAL SIMPLE: Imos a *Analizar / Regresión / Lineal*. Introducimos como variable dependente *AcTot* e como variable independente *AcLib*. No botón *Estadísticos* marcamos: *estimaciones, intervalos de confianza, matriz de covarianza e ajuste del modelo*. Facemos clic en *Continuar*. No botón *Gráficos* seleccionamos *Gráficos de los residuos: Histograma e gráfico de probabilidad normal*. Facemos clic en *Continuar*. No botón *Guardar* marcamos *valores pronosticados no tipificados, residuos no tipificados, residuos tipificados e intervalos de pronóstico para la media y para cada individuo*.

3. VALIDACIÓN DO MODELO: Un modelo de regresión é válido cando cumpre as seguintes condicións:

- ❖ Linealidade na relación entre as variables (nube de puntos).
- ❖ Independencia das observacións.
- ❖ Os erros distribúense normalmente; equivalentemente os erros tipificados distribúense como unha  $N(0,1)$  (gráficos dos residuos tipificados).

Para unha análise máis exhaustiva do último punto faise un estudio descritivo dos residuos. Imos a *Analizar / Estadísticos Descriptivos / Explorar*, tomamos como variable dependente *ZRE\_1 (Residuo Tipificado)*. No botón de *Estadísticos* seleccionamos *Descriptivos e Valores Atípicos*. No botón *Gráficos* unicamente marcar *Gráficos con prueba de normalidad*. Facemos clic en *Continuar* e en *Aceptar*.

**Exercicio2:** No estudo do exemplo anterior, o investigador ten interese en saber se existe relación entre a *humidade da zoa* e o índice *azuhum*, e en caso afirmativo, quere poder predecir o índice de *azuhum* en función da *humidade da zoa*. Facer un estudo completo de regresión para este modelo.



## Descrición das tarefas realizadas por cada membro

- Nome do membro 1: **M<sup>a</sup> del Carmen Carollo Limeres**
  - Tarefa realizada 1: Coordinación.
  - Tarefa realizada 2 : Elaboración dos materiais producidos.
  - Tarefa realizada 3: Impartición dos contidos, tanto na aula usual de docencia como nos laboratorios de informática.
  
- Nome do membro 2: **Balbina Virginia Casas Méndez**
  - Tarefa realizada 1: Elaboración dos materiais producidos.
  - Tarefa realizada 2: Impartición dos contidos, tanto na aula usual de docencia como nos laboratorios de informática.
  
- Nome do membro 3: **Pedro Faraldo Roca**
  - Tarefa realizada 1: Elaboración dos materiais producidos.
  - Tarefa realizada 2: Impartición dos contidos, tanto na aula usual de docencia como nos laboratorios de informática.
  
- Nome do membro 4: **M<sup>a</sup> Angeles Fernández Fernández**
  - Tarefa realizada 1: Elaboración dos materiais producidos.
  - Tarefa realizada 2: Impartición dos contidos, tanto na aula usual de docencia como nos laboratorios de informática.

## Grao de cumprimento dos obxectivos do proxecto

- Introducción ó paquete informático SPSS: 80%
- Motivación cara á materia: 75%
- Incremento da participación do alumno: 90%

## Grao de cumprimento do plan de traballo

O material foi elaborado durante o primeiro cuadrimestre de acordo coa planificación inicial.

Tratábase de introducir ao alumno no manexo dun paquete informático que lle permitira tratar, de modo sinxelo, problemas relacionados co mundo da Bioloxía que, levando consigo un gran número de datos, fixeran imprescindible o uso dalgunha ferramenta dese tipo.

Para acadar tales obxectivos impartíronse, en cada un dos catro grupos e en horario de clases prácticas, tres horas de docencia na aula usual (coa axuda dun ordenador portátil e un canón de vídeo), e dúas horas no laboratorio de informática (estructuradas en dúas prácticas de 1 hora cada unha e nelas puideron aplicar, a distintos datos, os coñecementos que lle foran expostos nas clases teóricas). Estas últimas impartíronse en grupos reducidos, de xeito que un ordenador non fose compartido por máis de dous estudantes

Acadouse traballar os tres primeiros puntos do plan, é dicir: introducción ao manexo do SPSS, utilización dos módulos básicos de análise exploratorio de datos e a utilización do módulo de regresión univariante. Debido a que as horas dedicadas ás prácticas resultaron ser insuficientes, non foi posible desenrolar o punto catro do plan de traballo (utilización do módulo de inferencia estatística).

## Valoración da experiencia

Non se fixo ningunha proba avaliativa dos coñecementos adquiridos polos estudantes debido a que esta actividade non aparecía, oficialmente, entre os métodos de avaliación do alumno.

Agora ben, tendo en conta o seu carácter non obrigatorio, as datas de realización da experiencia (coincidencia en día, aínda que non en horario, con exames doutras materias), e o número de alumnos que asistían con regularidade ás clases tradicionais, a asistencia ás clases específicas desta experiencia foi aceptable. Durante o proceso de desenrolo os alumnos parecían estar interesados nos contidos que lle amosaba o profesor e realizaron con éxito as tarefas que lles foron encomendadas durante as dúas clases prácticas.

En liñas xerais, os alumnos asimilaron os obxectivos marcados en cada práctica, participaron na realización dos exercicios prantexados durante a hora da clase e incluso realizaron, fora do horario da clase, algunha tarefa de carácter NON obrigatorio que lles foi proposta.

Podemos concluír que os alumnos acolleron ben a experiencia pois a súa asistencia e participación foi considerable. Aínda que a súa duración non foi o suficientemente ampla como para abordar tódolos puntos desexados e algúns estudantes comentaron que as clases tiñan sido un pouco rápidas, creemos que, adecuando un pouco o horario, o desenrolo destas clases con ordenador, tanto as teóricas como as prácticas, serían de gran apoio ó alumno para poder assimilar mellor os contidos da materia.

Consideramos que esta experiencia será tamén de gran utilidade para vindeiro curso 2005-2006 xa que o primeiro curso da titulación de Bioloxía entra a formar parte da experiencia piloto para a converxencia europea.