



# SeLA

Aplicación para la enseñanza y el aprendizaje  
de Inferencia Estadística

<http://www.ugr.es/local/bioest/sela>

---

Pedro Femia Marzo

Sep/2012

Bioestadística – Dep. de Estadística e I.O.

Universidad de Granada

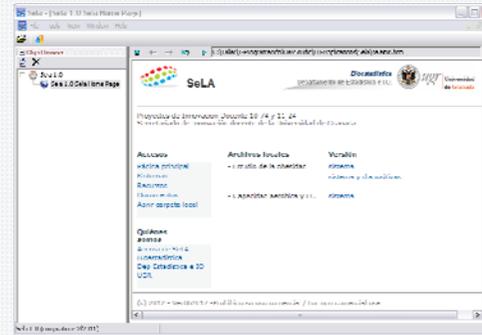
Proyecto de Innovación docente (10-74 / 11-24)

Secretariado de Innovación docente de la Universidad de Granada



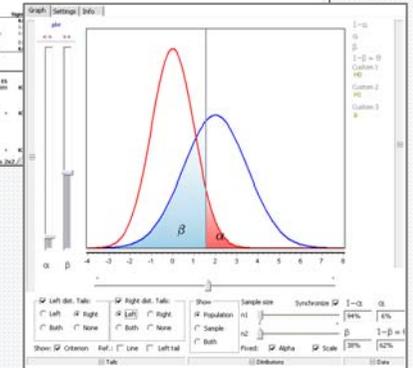
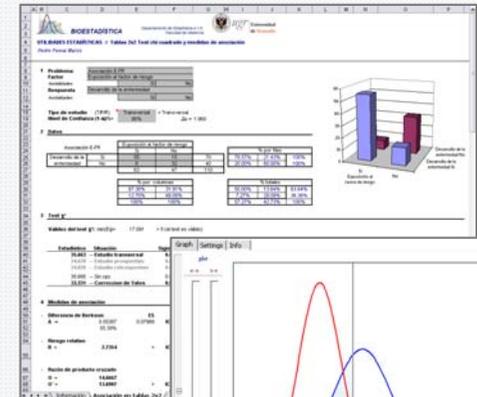
Portal Web de la aplicación

<http://www.ugr.es/~bioest/sela/>



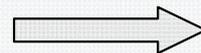
Núcleo de la aplicación (simulador + organizador de recursos)

SeLA.exe



Recursos

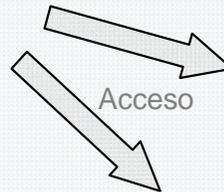
Acceso



Acceso



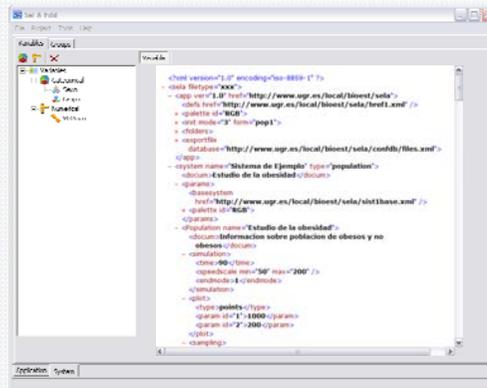
Acceso



Lectura



codificación



Editor de sistemas

SeLAEdit.exe



Sistemas (XML)



Proyecto de Innovación Docente

**SeLA**

Statistical e-Learning Application



ugr

Universidad  
de Granada

**Página de inicio**

**Noticias**

**Documentos**

**Descargas**

**Biblioteca de sistemas**

**Recursos adicionales**

**Acerca de SeLA**

**Lo mas reciente**

- Versión actual (exe)
- Versión actual (zip)

**Innovación docente UGR**

- > Portal SID
- > Acceso identificado

**Applets de Estadística**

- > Virtual Lab in Statistics
- > Cybergnostics project
- > L Green's pages
- > Holmes'-Swanson

**Enlaces Institucionales**

Grupo de Bioestadística  
Dep. Estadística e I.O.  
Universidad de Granada

## Presentación

*En la era de la tecnología  
es necesaria una nueva forma de enseñar y de aprender.  
Marc Prensky.-*

## SeLA

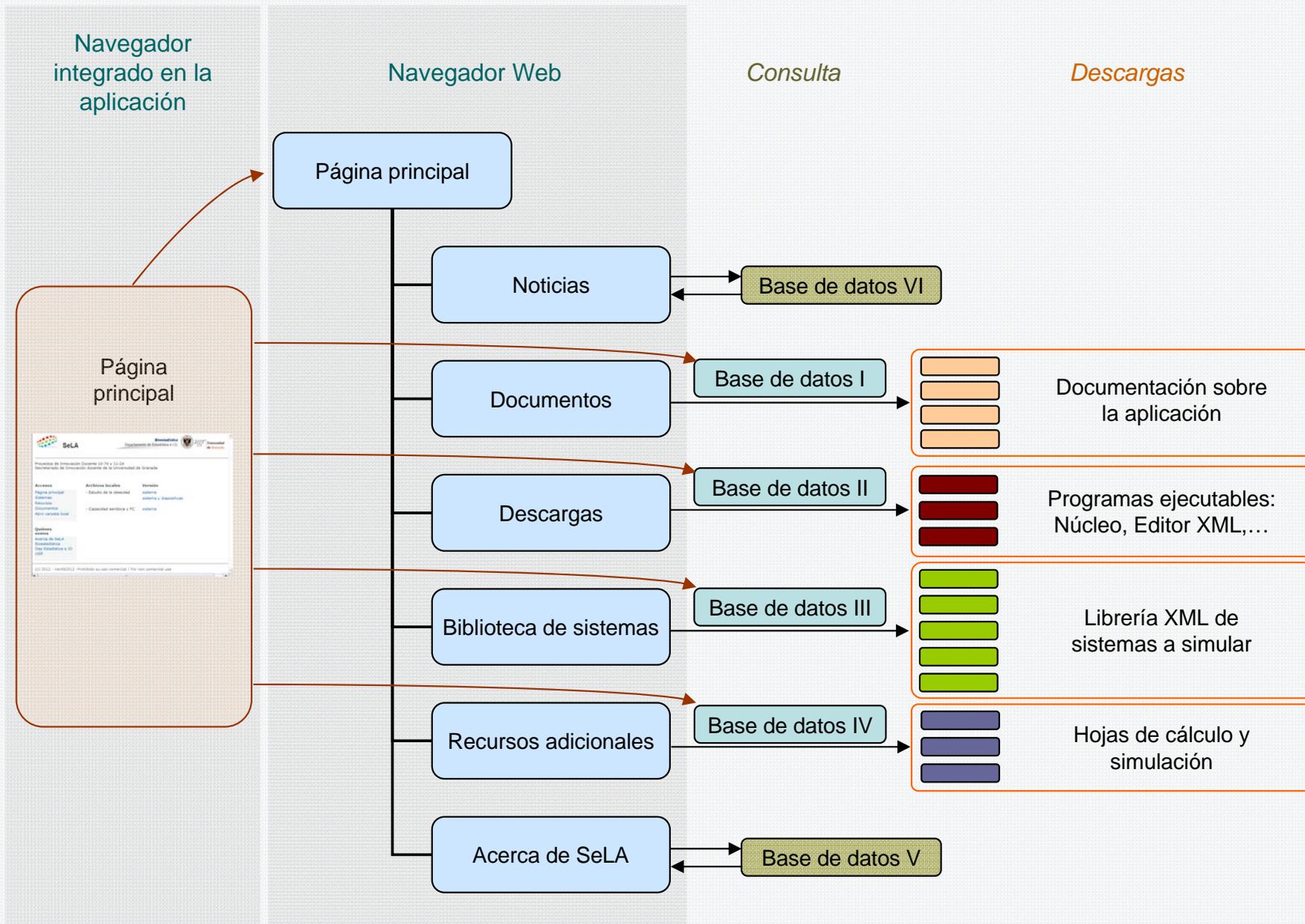
SeLA (*Statistical e-Learning Application*) es un proyecto desarrollado en el marco del programa de ayudas a proyectos de innovación docente por parte de la [Universidad de Granada](#) a través de su [Secretariado de Innovación Docente](#).

SeLA pone de manifiesto la dificultad implícita en el estudio de una población y permite reproducir la secuencia de actividades que, desde el punto de vista estadístico, permiten abordar dicho estudio.

SeLA es una aplicación que pretende servir de apoyo para la práctica del *aprendizaje activo*, así como de plataforma de *e-learning*, respondiendo de este modo a la demanda de los nuevos planteamientos educativos por parte del [Espacio Europeo de Educación Superior](#) (EEES).

Actualmente la aplicación SeLA y todo su entorno web se encuentra en fase activa de desarrollo. Este proyecto se elabora en el seno del grupo de [Bioestadística](#) del [Departamento de Estadística e I.O.](#) de la [Universidad de Granada](#). El equipo responsable del mismo desea expresar su agradecimiento a la Universidad de Granada por la confianza depositada en esta idea.

Granada, 01-10-2012



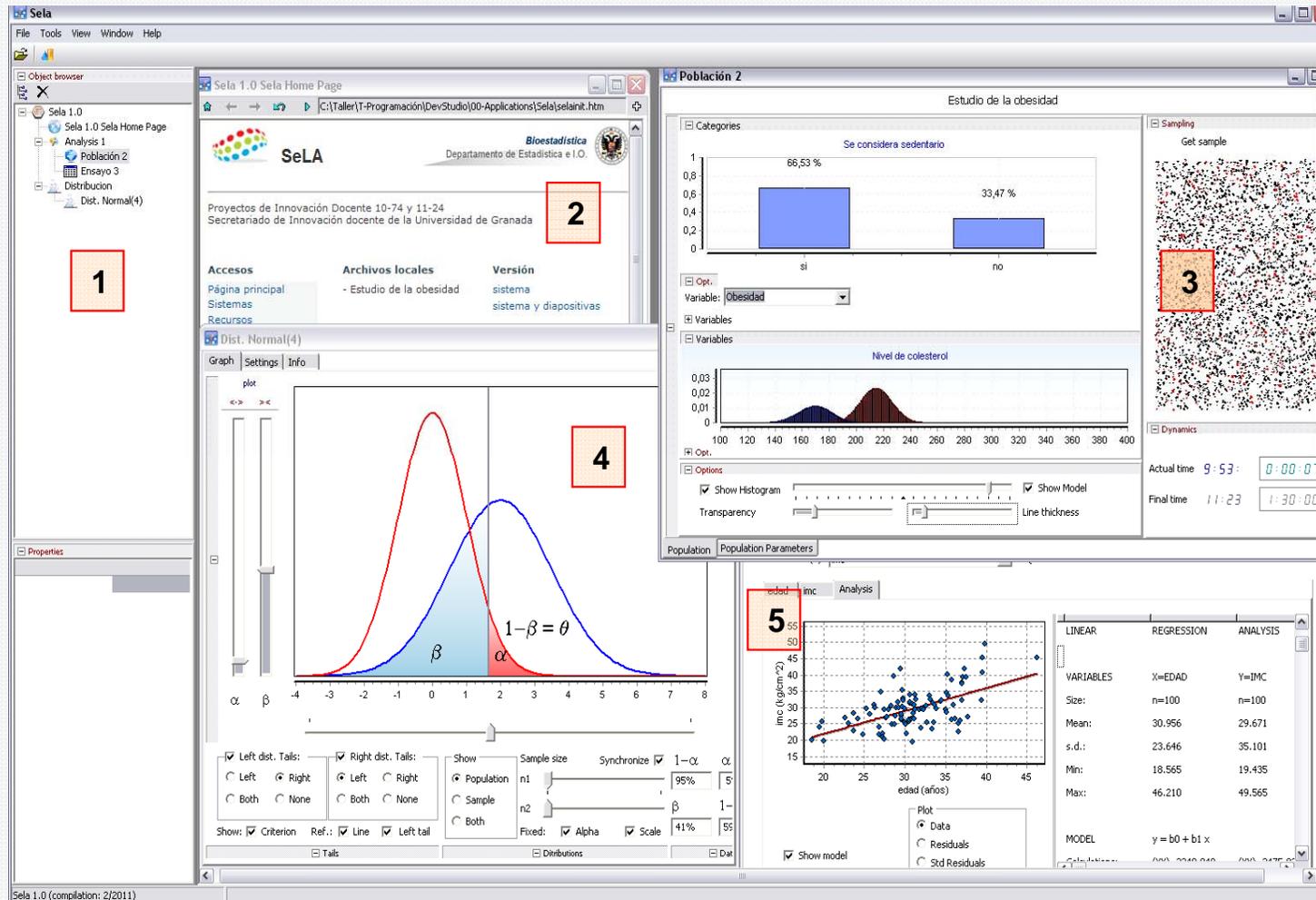


The screenshot displays the SeLA application interface with five numbered callouts:

- 1**: Content tree (Árbol de contenidos) in the left sidebar.
- 2**: Home page (Web de inicio) showing the SeLA logo and navigation options.
- 3**: Population simulation (Población) window showing a bar chart for 'Se considera sedentario' (66.53% si, 33.47% no) and a scatter plot of 'Obesidad'.
- 4**: Utility graphs (Utilidades) showing two overlapping normal distribution curves with shaded areas for  $\beta$  and  $1-\beta = \theta$ .
- 5**: Sampling and analysis (Muestreo y análisis) window showing a scatter plot of 'lmc (kg/cm<sup>2</sup>)' vs 'edad (años)' with a regression line, and a table of statistics.

VARIABLES	X=EDAD	Y=lmc
Size:	n=100	n=100
Mean:	30.956	29.671
s.d.:	23.646	35.101
Min:	18.565	19.435
Max:	46.210	49.565

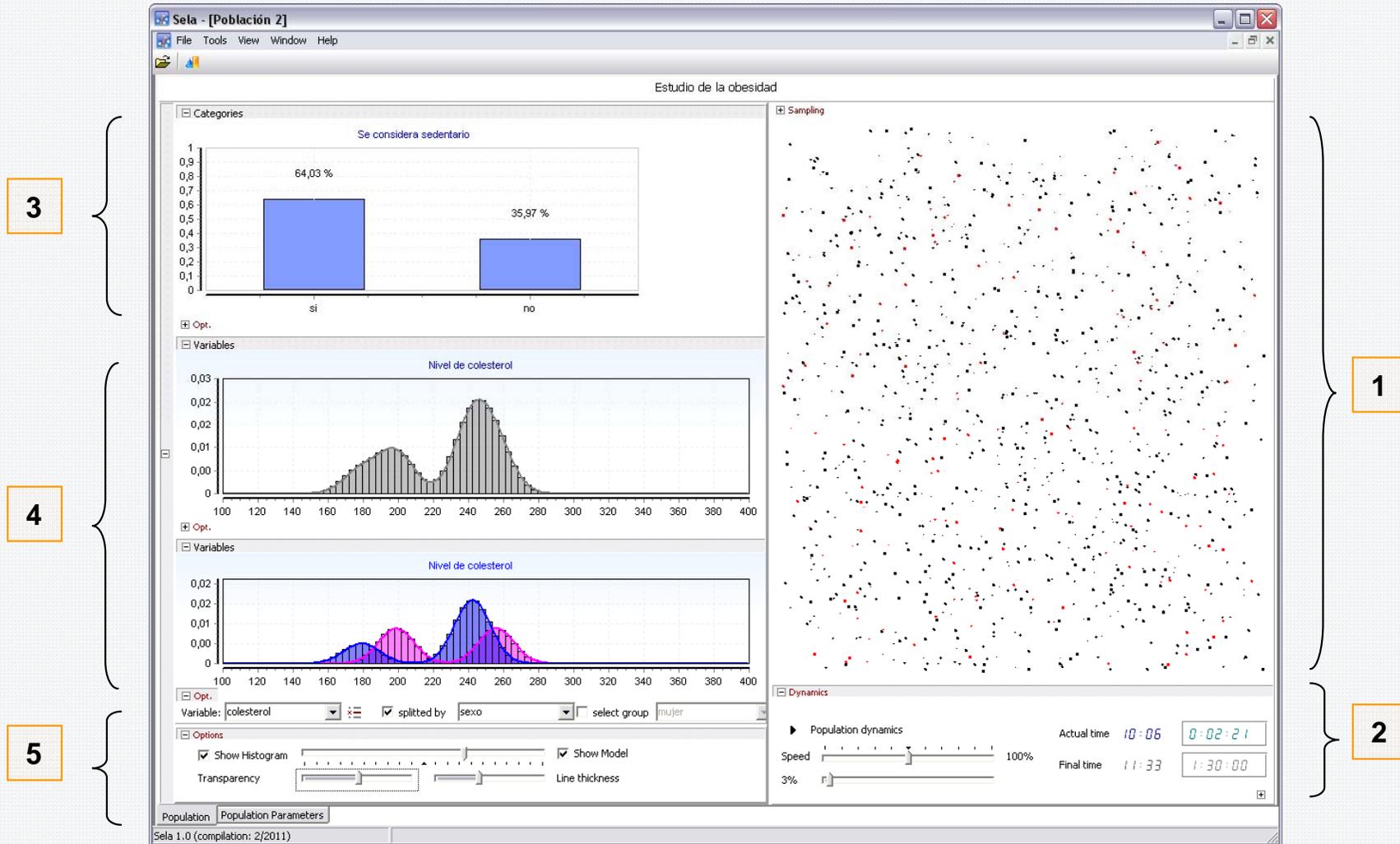
1. Árbol de contenidos
2. Web de inicio
3. Población (simulación de los componentes y distribuciones de las variables implicadas)
4. Utilidades.
5. Muestreo y análisis de la muestra obtenida



1. Árbol de contenidos
2. Web de inicio
3. Población (simulación de los componentes y distribuciones de las variables implicadas)
4. Utilidades.
5. Muestreo y análisis de la muestra obtenida



La Población a estudiar



1. Representación de la dinámica de la población
2. Controles del tiempo de simulación y velocidad de la dinámica de la población
3. Descripción gráfica de las variables cualitativas definidas en el problema
4. Descripción gráfica de las variables cuantitativas definidas (se pueden simultanear dos paneles con fines comparativos)
5. Control de la amplitud de los intervalos de los histogramas y de la presentación del modelo de distribución



1

2

3

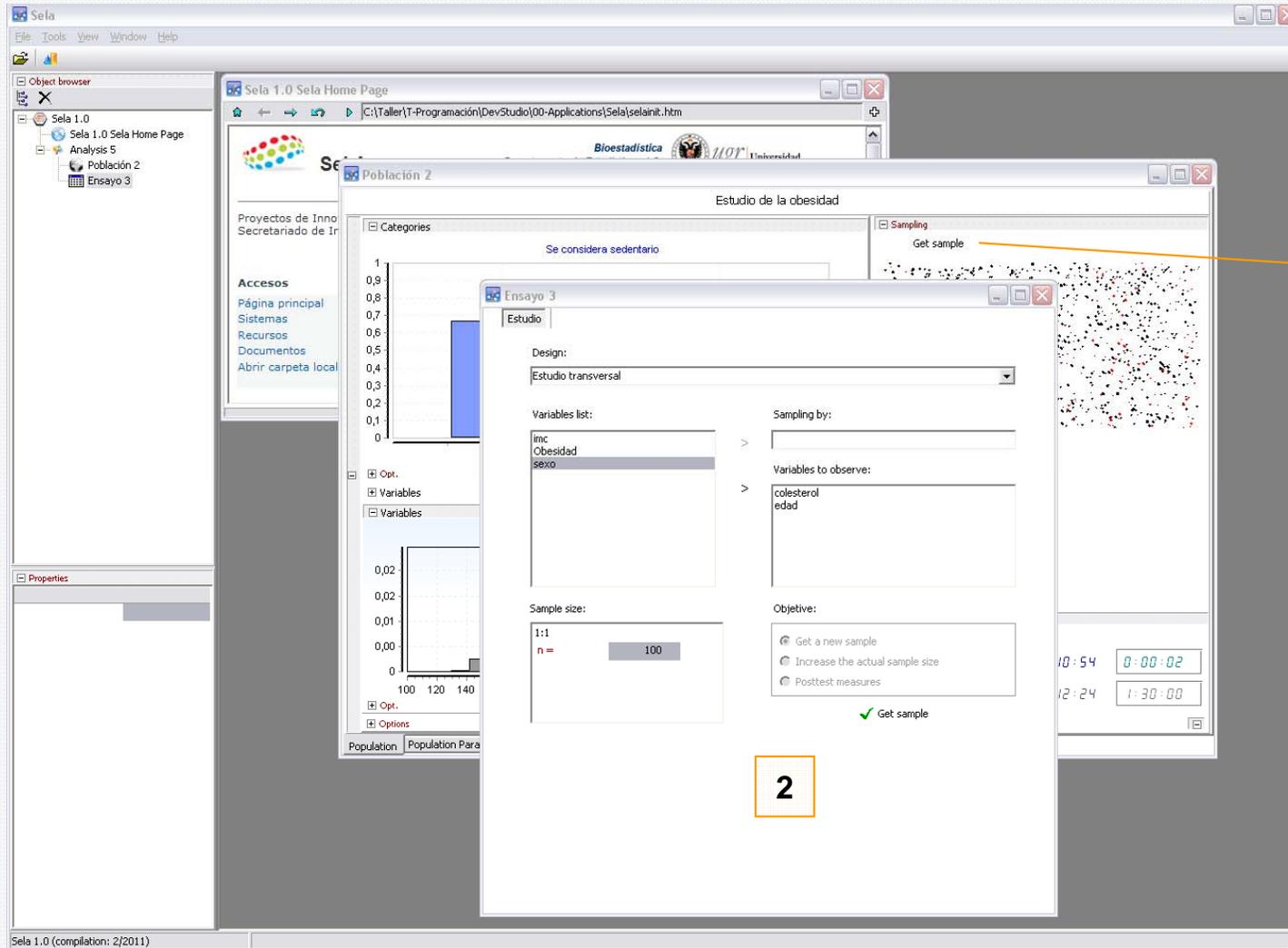
	estimación	SE	IC(-)	IC(+)	texp	g.l.	P
Varianza de regresión	69.946						
Ordenada en el origen	1.271	16.334	-32.303	34.846	0.078	26.000	0.939
Pendiente (Y/X)	5.386	0.707	3.933	6.839	7.619	26.000	0.000

Por cada unidad que aumenta la variable B, la variable A aumenta en promedio  $b=5,386$  unidades.

1. Simulación que ilustra la ley de azar
2. Áreas notables bajo la distribución normal (aplicación como soporte a la teoría de los contrastes de hipótesis)
3. Integración de hojas de cálculo en Excel (accesibles a través del menú web de la aplicación)



## Selección de la muestra. (I) Elección de variables a muestrear



1. Control para muestrear la población
2. Formulario de elección de
  - El tipo de estudio,
  - Las variables a muestrear
  - El tamaño de muestra



Selección de la muestra. (II) Medidas de cada individuo

1. Ficha de cada individuo de la muestra
2. Variables que se observan
3. Variables que se miden
4. Simulación del proceso de medición
5. Medidas repetidas de la misma variable en el mismo sujeto (ilustración del error de medida)
6. Controles para incorporar el sujeto a la muestra o muestrear (al instante) los sujetos restantes



## Aplicación principal (sela.exe)

### Selección de la muestra. (III) Datos muestreados y opciones de exportación

The screenshot shows the SPSS 'Datos' window for 'Ensayo 3'. The window has tabs for 'Estudio', 'Observaciones', 'Datos', and 'Análisis'. The 'Datos' tab is active, displaying a data matrix with 19 rows and 6 columns: '1:0', 'sexo', 'Obesidad', 'colesterol', 'imc', and 'edad'. The data is as follows:

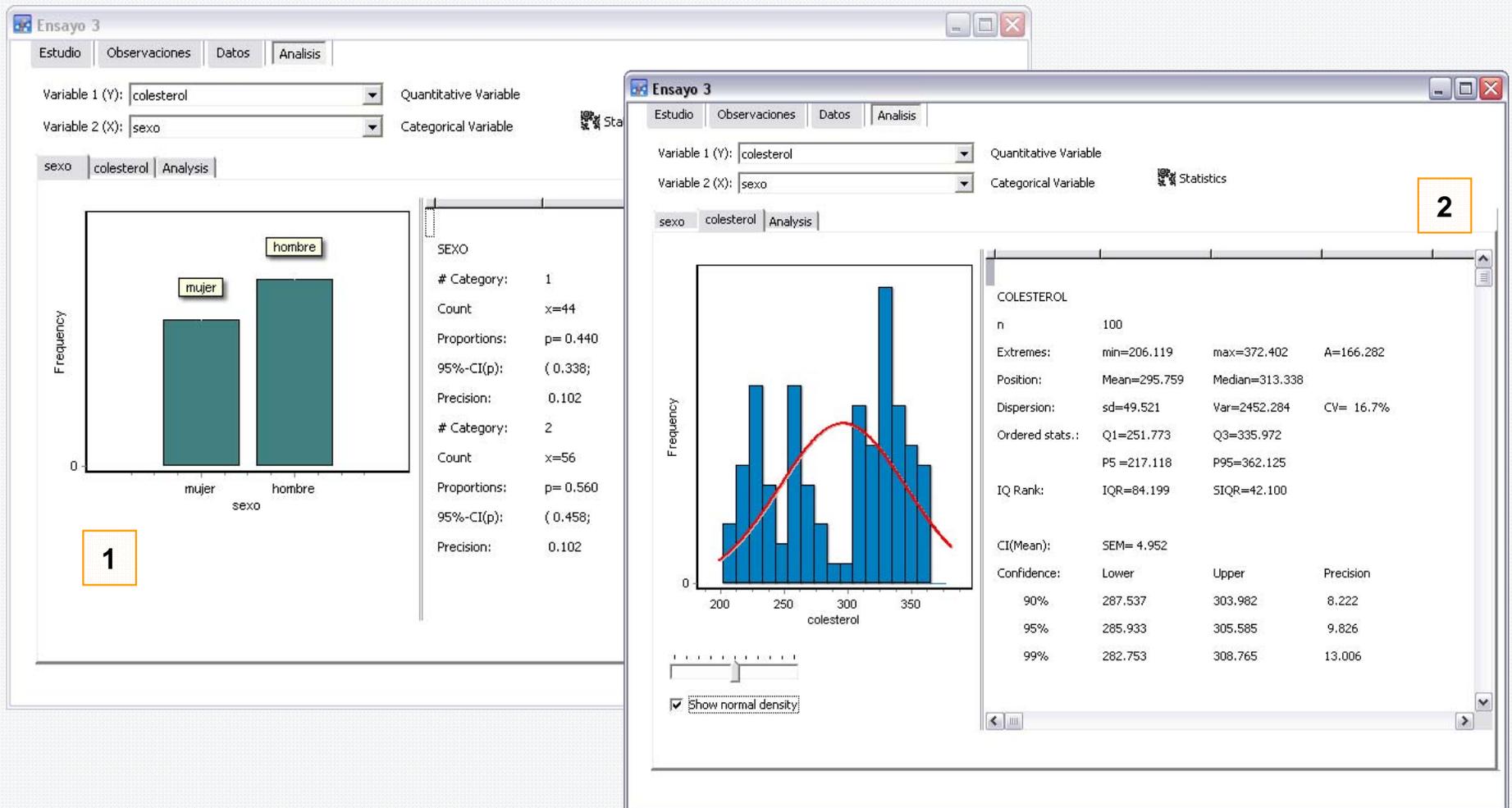
1:0	sexo	Obesidad	colesterol	imc	edad
1	2	1	316.82	49.377	51
2	2	1	331.58	50.219	49
3	2	1	306.17	43.712	53
4	2	1	331.89	56.138	52
5	2	1	308.19	44.422	50
6	1	2	262.70	37.595	38
7	2	1	334.86	52.249	51
8	2	2	231.39	55.921	44
9	2	1	331.65	48.242	52
10	1	1	357.72	50.321	49
11	2	1	332.70	53.009	51
12	1	2	275.89	40.602	38
13	1	2	261.09	44.043	37
14	2	1	346.09	57.736	55
15	1	1	352.68	49.465	44
16	2	1	334.03	52.545	53
17	1	1	341.45	49.938	51
18	2	1	313.23	46.794	56
19	2	2	240.07	59.920	48

On the right, the 'VARIABLES' list shows descriptions for 'sexo', 'Obesidad', 'colesterol', 'imc', and 'edad'. A cursor is positioned over the 'Obesidad' description, which is highlighted with a red box labeled '2'. The 'Obesidad' description is: 'Se considera sedentario - 1-> si - 2-> no'. The 'Export' icon (a floppy disk) is highlighted with a red box labeled '3', and its context menu is open, showing options: 'Spss syntax file (\*.sps)', 'Excel worksheet (\*.xls)', 'Text csv (\*.csv)', and 'Plain text (\*.txt)'. The 'Excel worksheet (\*.xls)' option is highlighted with a red box labeled '1'. Below the menu, it says 'Decimals are comma delimited'.

1. Matriz con la muestra de datos (tabla de casos x variables)
2. Descripción de las variables que componen la muestra
3. Menú de exportación de datos. Los formatos posibles son
  - SPSS (en forma de archivo de sintaxis, conserva etiquetas de variable y valor)
  - MS-Excel 2003
  - Texto plano (diferentes posibilidades de delimitación de campos)



Análisis de la muestra. (I) Estudio univariante



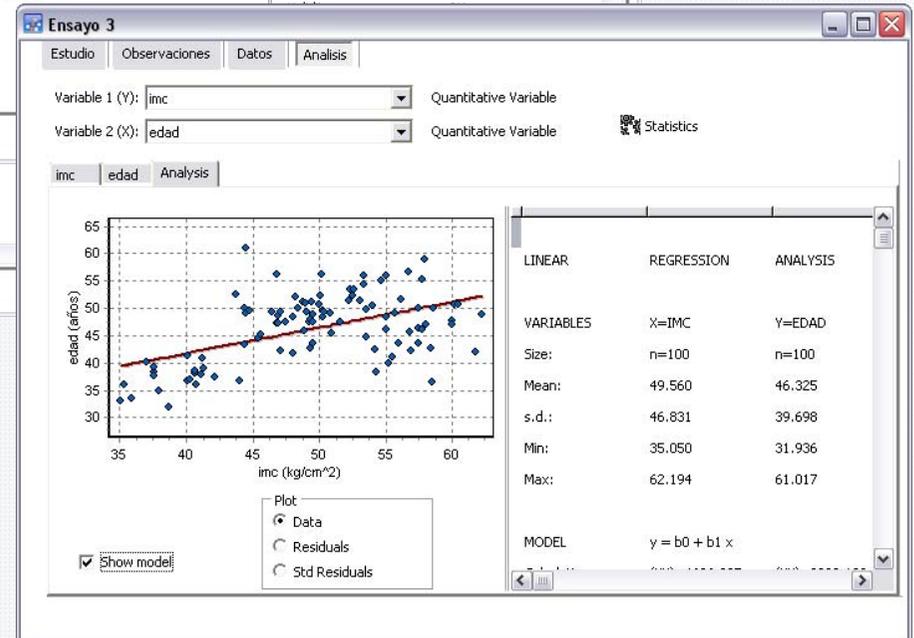
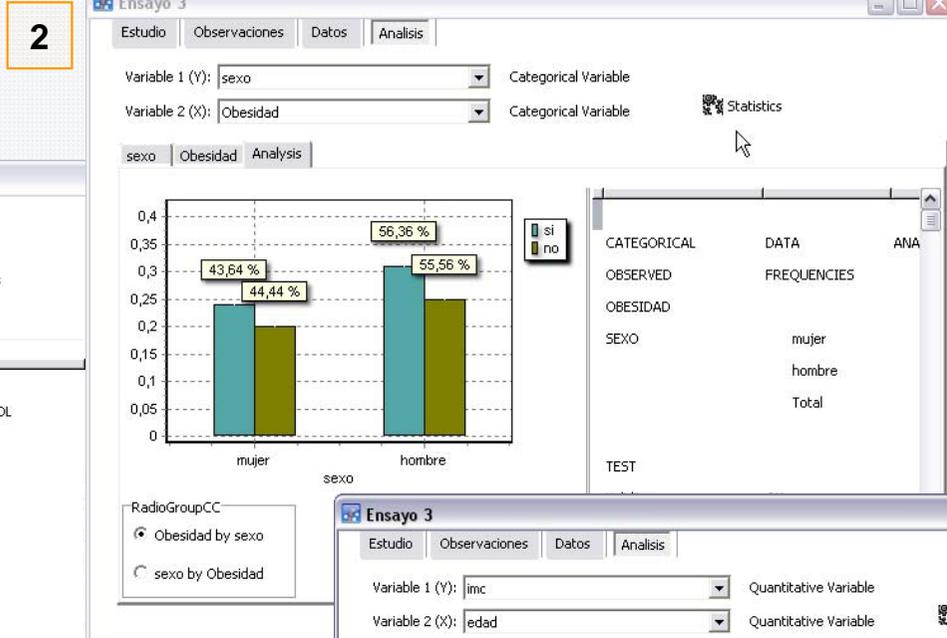
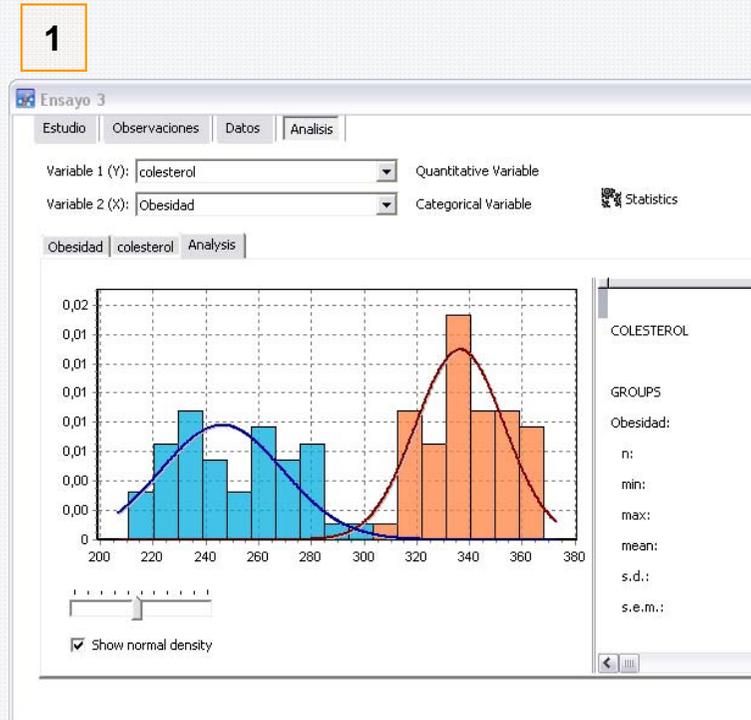
1. Panel de análisis de variables categóricas
2. Panel de análisis de variables cuantitativas

En ambos casos se presentan:

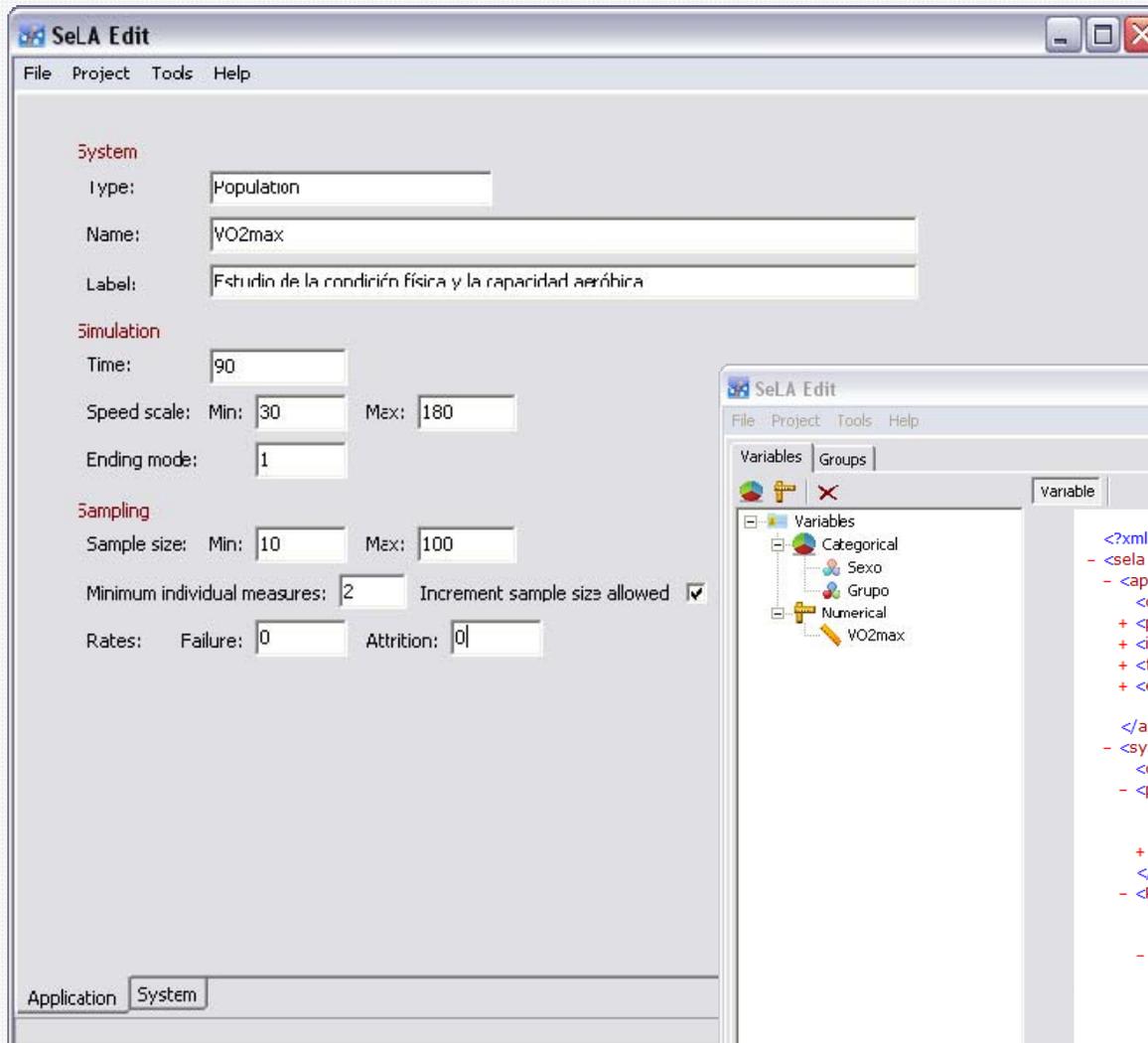
- Diagramas de frecuencias
- Medidas descriptivas
- Intervalos de confianza para los parámetros oportunos



Análisis de la muestra. (II) Estudio bivalente

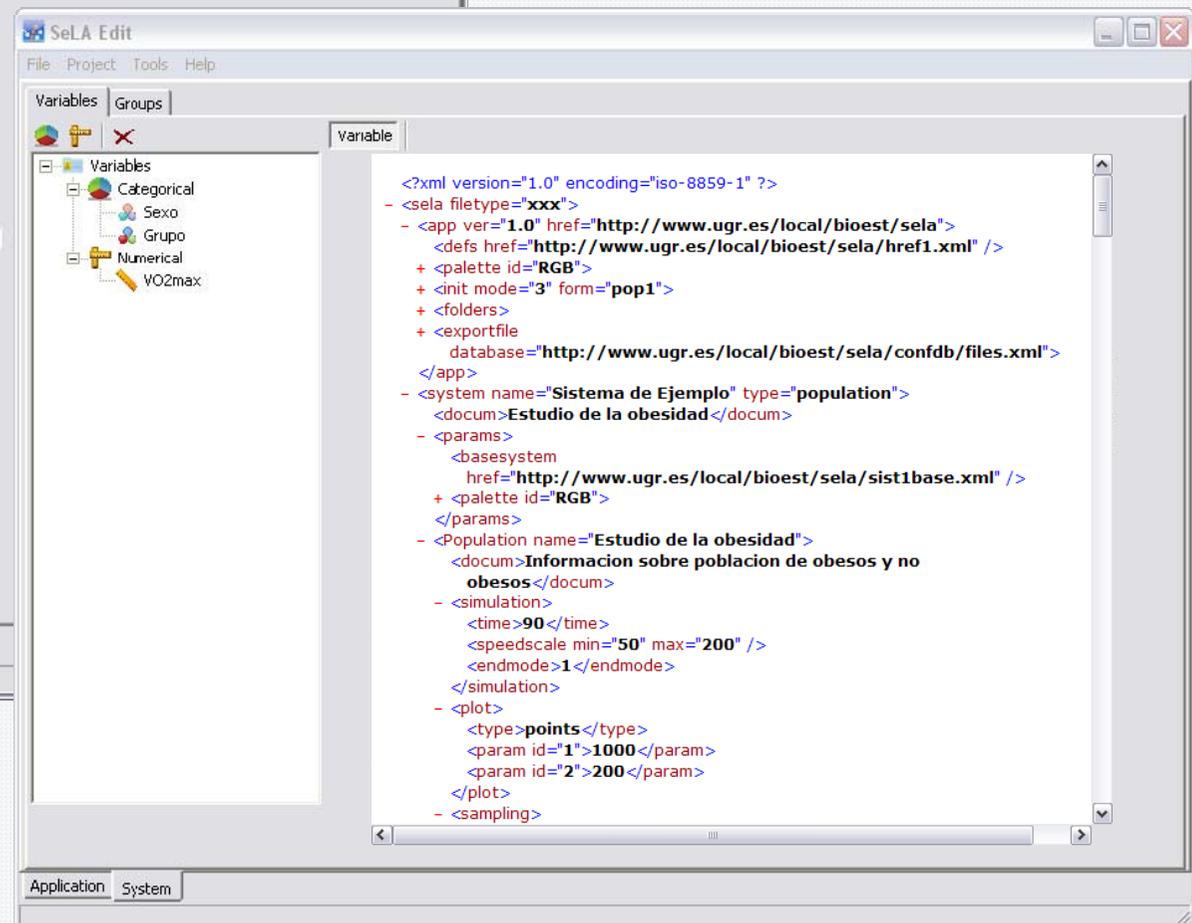


1. Panel de análisis de variable categórica x variable categórica (tabla de contingencia y test chi-2)
2. Panel de análisis de variable categórica x variable cuantitativa (pruebas de homogeneidad para medias)
3. Panel de análisis de variable cuantitativa x variable cuantitativa (análisis de regresión lineal)



1

Panel de configuración de las características generales del sistema

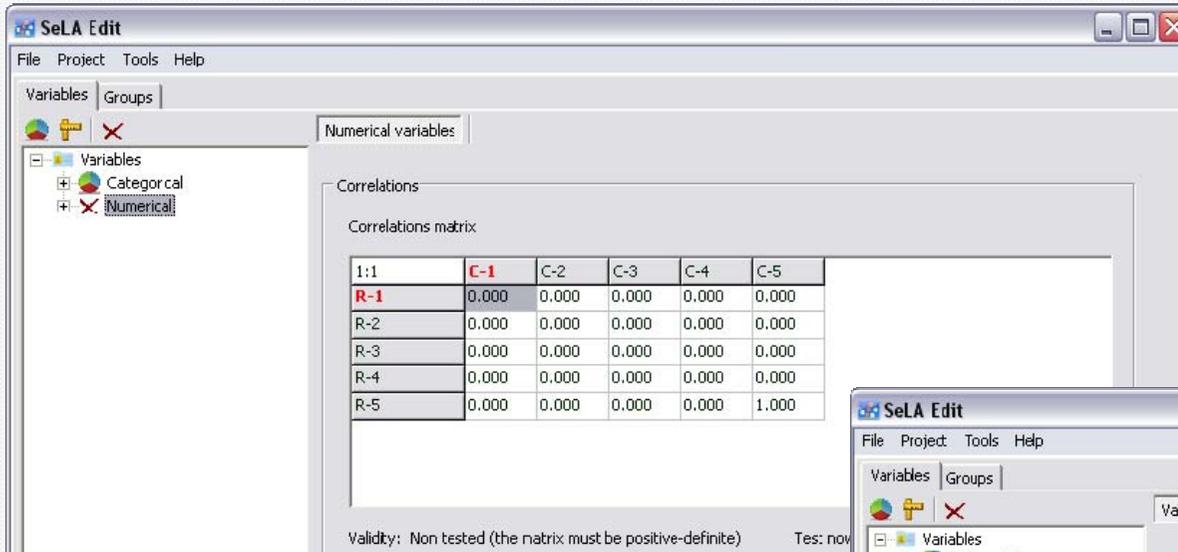


2

Editor directo de código XML



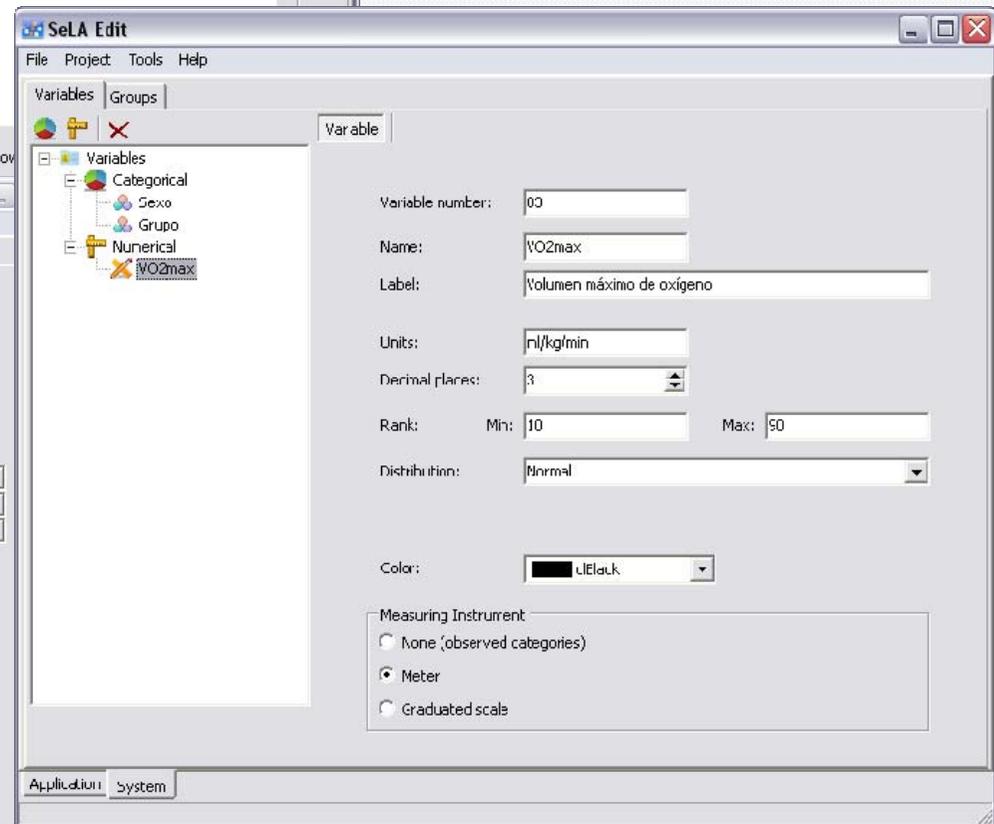
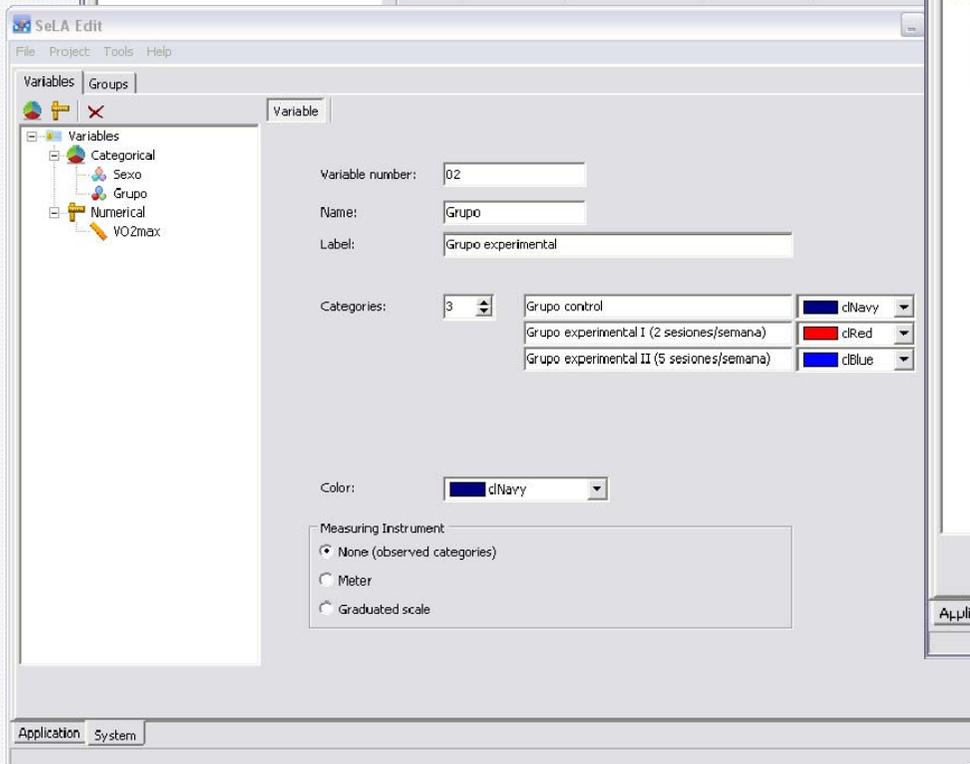
1



1. Especificación de la correlación multivariante
2. Especificación de variables categóricas
3. Especificación de variables cuantitativas

3

2





1. Gestor de grupos
2. Herramienta de perfilado de distribuciones (calibración de los parámetros para la simulación)



Utilidades de cálculo interactivas con SeLA – Variables cualitativas

Inferencia con tablas de contingencia (2x2)

BIOESTADÍSTICA Departamento de Estadística e I.O. Facultad de Medicina UGR Universidad de Granada

UTILIDADES ESTADÍSTICAS // Tablas 2x2 Test chi cuadrado y medidas de asociación Pedro Femia Marzo

1 Problema: Asociación E-FR  
 Factor: Exposición al factor de riesgo  
 Respuesta: Desarrollo de la enfermedad

Exposición al factor de riesgo	Si	No	
Desarrollo de la enfermedad	Si	No	% por filas
	63	47	78.97% 21.03% 100%
	8	32	20.00% 80.00% 100%
	63	47	100%

2 Datos

Asociación E-FR	Exposición al factor de riesgo		% por columnas
	Si	No	87.00% 31.9%
	No		12.70% 68.09%
			100%

3 Test  $\chi^2$

Validez del test  $\chi^2$ : min(E) 17.091 > 5 (el test es)

Estadístico	Situación	Significancia
35.663	-- Estudio transversal	0.00000
34.670	-- Estudio prospectivo	0.00000
34.820	-- Estudio retrospectivo	0.00000
35.685	-- Sin opo	0.00000
33.331	-- Corrección de Yates	0.00000

4 Medidas de asociación

Diferencia de Berkson  $\kappa = 0.95387$  ES 0.07989 IC (R) = 0.95387

Riesgo relativo  $R = 2.7354$  IC (R) = 0.4985

Razón de producto cruzado  $O = 14.6867$   $O^* = 13.6907$  IC (O) = 0.9954

Riesgo atribuible  $Rb = 0.4985$  IC (Rb) = 0.7321

Aproximación a los riesgos relativo y atribuible  $R = 14.6867 (= O)$   $Rb = 13.6907 (= O^*)$  IC (R) = 0.9954

1 Entrada de datos

n	202
x	100
n-x	102
Confianza (%)	95
Z $\alpha$	1.960

2 Estimaciones puntuales

p	0.4950	49.50%
q	0.5050	50.50%

3 Intervalo de confianza y precisión obtenida

3a Intervalo exacto

IC (95%)	Lim. Inf.	0.42415	42.41%
	Lim. Sup.	0.56610	56.61%
	Precisión:	0.071	7.10%

3b Intervalo basado en la aproximación a la normal (IC cuadrático de Fleiss)

IC (95%)	Validez:	< IC corto	
	Lim. Inf.	0.42363	42.36%
	Lim. Sup.	0.56647	56.65%
	Precisión:	0.071	7.14%

3c Tamaño de muestra (sin cpc)

Precisión	0.04	n	p	q
Sin información previa		601	0.500	0.500
Con muestra piloto		601	0.500	0.500

4 Test Ho: p=po

po	0.4
Colas	2
Zexp	2.686
Significación	0.0072

Inferencia con una proporción

Tendencia con proporciones

BIOESTADÍSTICA Departamento de Estadística e I.O. Facultad de Medicina UGR Universidad de Granada

UTILIDADES ESTADÍSTICAS // Tendencia con proporciones Pedro Femia (2010)

1 Observaciones

Periodo	Casos favorables	Casos totales	Proporción observada	Proporción estimada
0	100540	217613	0.46201	0.4323
3	65105	201998	0.32231	0.5037
6	268282	372844	0.71956	0.5750
9	262290	442481	0.59277	0.6464

2 Modelo lineal

Promedios	p	1-p			
	0.5638	0.4362			
Modelo	ordenada	Pendiente	txp	gl	P
	0.4323	0.0238	174.8093	64162	0.0000
Pronósticos	periodo	p estimada			
	12	0.7177			
	15	0.7891			

BIOESTADÍSTICA Departamento de Estadística e I.O. Facultad de Medicina UGR Universidad de Granada

UTILIDADES ESTADÍSTICAS // Aproximación a los Índices de validez y fiabilidad corregidos por azar en tablas 2x2 Pedro Femia Marzo

1 Datos Si hay un estándar, debe aparecer en filas

R1	C1	C2	46	0	46
R2	1	1117	1118		
	47	1117	1164		

Tabla analizada:

47	1118	1120
2	1118	1120
49	1119	1168

Problema tipo: Normal (2.2)  
Kappa = 0.989  
Kappa = 0.9677

2 Opciones

Muestreo tipo: 1 < Tipo I  
 R = estándar: 0 < NO  
 Asintótico tipo: 1 < Extra

Destacar fnto: 1 < SI  
 Se destacan las medidas válidas según el tipo de muestreo y la existencia de estándar

3 Medidas

Índice	Validez	Muestreo	Fila	Estimador	Varianza	S.E.
Conformidad Fi	R = Estándar	I	1	0.949703884	0.001361527	0.036763116
			2	0.996951595	0.000003315	0.001820666
Predictividad Pi	R = Estándar	I	1	0.949703884	0.001346812	0.036696934
			2	0.996951595	0.000003306	0.001818266
Acuerdo Ai	Siempre	I	1	0.039028927	0.000033697	0.005804948 <
			2	0.955980982	0.000037615	0.006133103 <
Consistencia Ci	R NO estándar	I	1	0.039028927	0.000002275	0.001508175
			2	0.955980982	0.000003040	0.001743662
D global	Siempre	I	1	0.939913122	0.001237754	0.035181722 <
			2	0.997396861	0.000002325	0.001524882 <
		II	1	0.995009908	0.000009523	0.002919479 <
			2	0.995009908	0.000008448	0.002906551

Acuerdo entre observadores



Utilidades de cálculo interactivas con SeLA – Variables cualitativas

Inferencia con una media

Comparación de 2 medias muestras independientes y apareadas

Regresión lineal simple

Test de normalidad

SeLA / Septiembre-2012  
Proyecto de Innovación docente (10-74 / 11-24)  
aprobado por el  
Secretariado de Innovación Docente  
de la Universidad de Granada

Programado en Borland Delphi (BDS 2006)  
Componentes utilizados:  
Steema TeeChart  
SDL-Epina Software

SeLA puede ser utilizado libremente con fines didácticos

SeLA NO puede ser utilizado con fines comerciales o de forma lucrativa

Todas las sugerencias son bienvenidas

Pedro Femia Marzo  
pfemia@ugr.es  
Bioestadística – Dep. de Estadística e I.O.  
Universidad de Granada